UNIVERSAL
LIBRARY

OU_160038

UNIVERSAL
LIBRARY

IMPERFECTIONS IN CRYSTALS

SERIES IN PHYSICS

General Editors:

J. DE BOER, Professor of Physics, University of Amsterdam
H. BRINKMAN, Professor of Physics, University of Groningen
H. B. G. CASIMIR, Director of the Philips Research Laboratories, Eindhoven

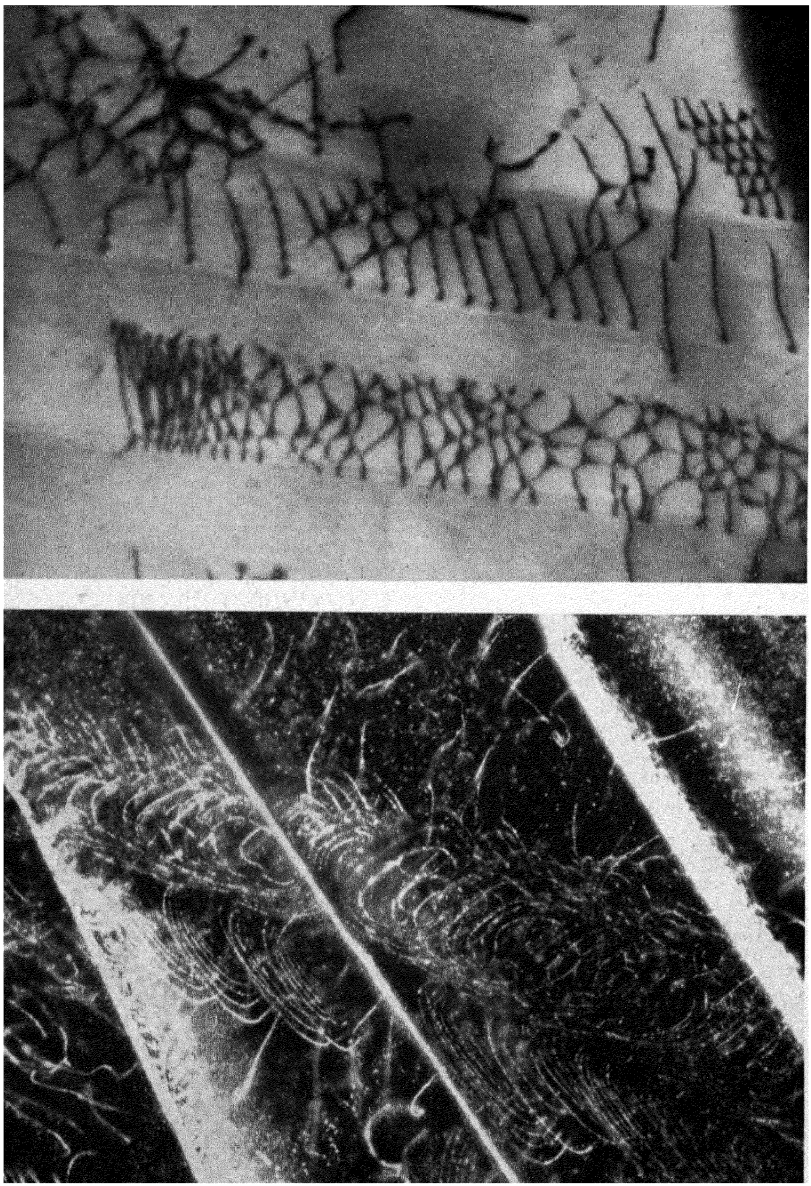
Monographs:

H. C. BRINKMAN, Application of Spinor Invariants in Atomic Physics
H. G. van BUEREN, Imperfections in Crystals
S. R. DE GROOT, Thermodynamics of Irreversible Processes
E. A. GUGGENHEIM, Thermodynamics
E. A. GUGGENHEIM, Boltzmann's Distribution Law
E. A. GUGGENHEIM and J. E. PRUE, Physicochemical Calculations
H. A. KRAMERS, Quantum Mechanics
H. A. KRAMERS, The Foundations of Quantum Theory
J. G. LINHART, Plasma Physics
J. MCCONNELL, Quantum Particle Dynamics
A. MERCIER, Analytical and Canonical Formalism in Physics
I. PRIGOGINE, The Molecular Theory of Solutions
E. G. RICHARDSON, Relaxation Spectrometry
P. ROMAN, Theory of Elementary Particles
M. E. ROSE, Internal Conversion Coefficients
J. L. SYNGE, Relativity: The Special Theory
J. L. SYNGE, The Relativistic Gas
H. UMEZAWA, Quantum Field Theory
A. VAŠÍČEK, Optics of Thin Films
A. H. WAPSTRA, G. J. NIJGH and R. VAN LIESHOUT, Nuclear Spectroscopy Tables

Edited Volumes:

J. BOUMAN (editor), Selected Topics in X-Ray Crystallography
J. M. BURGERS and H. C. VAN DE HULST (editors), Gas Dynamics of Cosmic Clouds. A Symposium
P. M. ENDT and M. DEMEUR (editors), Nuclear Reactions, Volume I
C. J. GORTER (editor), Progress in Low Temperature Physics, Volumes I–II
G. L. DE HAAS-LORENTZ (editor), H. A. Lorentz, Impressions of his Life and Work
J. KISTEMAKER, J. BIGELEISEN and A. O. C. NIER (editors), Proceedings of the International Symposium on Isotope Separation
J. KOCH (editor), Electromagnetic Isotope Separators and Applications of Electromagnetically Enriched Isotopes
Z. KOPAL (editor), Astronomical Optics and Related Subjects
H. J. LIPKIN (editor), Proceedings of the Rehovoth Conference on Nuclear Structure
N. R. NILSSON (editor), Proceedings of the Fourth International Conference on Ionisation Phenomena in Gases, Uppsala, 1959
K. SIEGBAHN (editor), Beta- and Gamma-Ray Spectroscopy
Symposium on solid state diffusion (Colloque sur la diffusion à l'état solide, Saclay, 1958)
Symposium on corrosion (3e Colloque de métallurgie, sur la corrosion, Saclay, 1959)
Turning Points in Physics. A series of lectures given at Oxford University in Trinity Term 1958
J. G. WILSON and S. A. WOUTHUYSEN (editors), Progress in Elementary Particle and Cosmic Ray Physics. Volumes I–V

P. EHRENFEST, Collected Scientific Papers



Examples of dislocation configurations.
Upper photograph: electron transmission micrograph (× 30 000; taken by HIRSCH, PARTRIDGE and SEGALL) of stainless steel foil, showing dislocation network associated with slipline. Lower photograph: optical micrograph (× 550; by AMELINCKX) of decorated KCl crystal, showing dislocation sources.

IMPERFECTIONS IN CRYSTALS

by

H. G. VAN BUEREN

Philips Research Laboratories, Eindhoven

SECOND EDITION



1961

NORTH-HOLLAND PUBLISHING COMPANY - AMSTERDAM

No part of this book may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the publisher

First Edition, 1960
Second Edition, 1961

PUBLISHERS:

NORTH-HOLLAND PUBLISHING CO. – AMSTERDAM

SOLE DISTRIBUTION IN U.S.A.:

INTERSCIENCE PUBLISHERS, INC. – NEW YORK

PRINTED IN THE NETHERLANDS

PREFACE TO THE FIRST EDITION

Some has been presented in the right sequence, some without order, and some but in part; much has been suppressed and left to the investigation by the inaugurated; these will, by devoting some thought to the context, discover therein not only a complete Theory..., but unfailing Experiments as well.

Henricus Cornelius Agrippa von Nettersheim ("Agrippa"):
"De occulta philosophia"
Cologne, 1510.

Solid state physics has taken its place among the major divisions of the physical sciences and the number of books concerned with this young and vital branch of science is steadily increasing. Most of these are devoted to what might be called quantumphysical aspects, like solid state electronics, optics or magnetism, where a theoretical description of the properties in the language of quantum mechanics and lattice vibrations is indicated.

At the same time it has become apparent that an equally important branch of solid state physics exists, which approaches the subject on a somewhat coarser scale: this is the study of *defects in the crystal lattice.* Crystal imperfections appear to influence the properties of solids to such an extent that solid state physics has proved in several instances to be equivalent to the study of the properties of the defect crystal, as contrasted to those of the perfect crystal.

Several monographs have been written dealing with particular aspects of the science of imperfections, but a book covering this whole field has not yet appeared. When it is realized that in the well-known series of review papers: Solid State Physics, until January 1959 30% of the contributions were predominantly concerned with crystal imperfections, and that for the planned contributions this percentage is even higher (40), it is clear that a general reference work would serve its purpose.

The present work does not claim such an ambitious designation; the field, however young, is already too wide for this. Nevertheless, in this book it has been conscientiously attemped to approach the subject of crystal imperfections from a wide angle, through dealing with the properties of imperfections in a diversity of solids, and pointing out

their influence on a wide variety of physical and mechanical properties. The purpose of the book is twofold:

1) to give the reader a *birdseye view* of the subject matter and to inform him about the research situation such as it was at the end of 1958; and
2) to interest him in the variety of *theoretical ideas* that have been proposed in this field of study, which might prove stimulating.

For the first purpose, a number of references has been included which might enable the reader to pursue a specific topic further; for the second purpose subjects have been treated and references added, just because of their theoretical interest only and not because of their general importance.

It will now be clear that the reference system is not complete, and that also the choice of subjects treated might be commented upon. To the authors' opinion, a strictly systematic treatment would have seriously impaired the readability. To stimulate the readers' mind, and to interest the student, is the first purpose of this book.

Acknowledgements:

I am greatly indebted to a number of friends and colleagues who have helped me by checking over parts of, or even the complete volume, and who have taken pains to correct a foreigner's use of the English language. Very sincere thanks are due in the first place to Dr R. W. CAHN of Birmingham University, Professor A. SEEGER of the Technische Hochschule in Stuttgart and Professor P. HAASEN from Göttingen University; their remarks have been of exceptional value in preparing the final version of the book. For the help of other members of the Stuttgart group, notably of Drs E. KRÖNER and S. MADER, I am very grateful indeed; their spontaneous aid as specialists in various fields has much improved the contents of the chapters concerned. To many of my colleagues in the Philips Research Laboratory I owe sincere thanks for their valuable criticism. Drs C. W. BERGHOUT, J. D. FAST, C. M. HARGREAVES, Y. HAVEN, J. HORNSTRA, J. L. MEIJERING and G. D. RIECK should be especially mentioned; to name all the others who have advised me is virtually impossible. The atmosphere of friendship and free exchange of information and criticism in the Philips Laboratory has been essential in the completion of this book.

Very important assistance in the preparation of the book has been given to me by Mr P. VISSERS and Mr J. LIESENGA, who have taken

care of all the line drawings in a most capable manner. I am seriously indebted to them for their cooperation.

Professor H. B. G. CASIMIR, director of the Philips Research Laboratories, has given me ample opportunity to make use of the facilities of the Laboratory and other services of the Philips concern in the preparation of the various versions of the manuscript and figures. For this help, which has been invaluable throughout the writing, and for the time liberally allowed to me during the final preparation of the book and for literature researches etc., I am sincerely indebted to him and to other members of the laboratory management. The kind and patient assistance of the publishers in the preparation of the printed version is gratefully acknowledged.

My wife has induced me to write the book, and has aided and sustained me in all possible ways. Without her, the work could not have been begun, and certainly not completed.

Eindhoven, March 1959

PREFACE TO THE SECOND EDITION

Owing to the short time elapsed between the appearance of this book and the need for a second edition, it did not seem necessary to the author—and indeed has not been possible to him—to prepare a *fully* revised manuscript. Only those errors and misprints that have become apparent have been corrected, a number of inaccuracies have been removed, and small additions to the text have been made here and there.

Nevertheless, the author realizes that appreciable progress has been made, experimentally as well as theoretically, in many subjects dealt with in this book. Especially the topics work-hardening, optical and rontgenographic visualization of dislocations, and plastic deformation of non-metals require at this time more attention than it was possible to give them when this book was first written. By inserting a few notes on the more salient points it has been tried to substitute for this omission.

At this place the author wants to express his gratitude to all those who have given their comments on text and presentation of the first edition.

Eindhoven, May 1961

ACKNOWLEDGEMENT
FOR USE OF ILLUSTRATIONS

Many diagrams have been derived with slight alterations from figures first published in various papers and books. In these cases reference has been made to the original diagram in the caption. A few drawings and most photographs have been reproduced in full with kind permission of the authors of the original publications in which they appeared. We gratefully acknowledge the help of Drs AMELINCKX, BOAS, CAHN, DASH, GILMAN, HAANSTRA, HEUMANN, HIRSCH, HORNSTRA, JAOUL, KOOY, LACOMBE, LOMER, MADER, MENTER, PARKER, PENNING, SEEGER, SUZUKI, WHELAN and WILSDORF. The reproduction has become possible by courtesy of the following Institutions and Publishing Houses: American Institute of Mining and Metallurgy Engineers, Inc., New York; American Institute of Physics, New York; Butterworths Scientific Publications, London; The Institute of Metals, London; McGraw-Hill Book Company, New York; Pergamon Press, London and New York; The Physical Society, London; The Royal Society, London; The Scientific American, Inc., New York; Stichting Physica, Amsterdam; Taylor and Francis, Ltd., London; Uitgeversmaatschappij Centrex, Eindhoven; Wiley and Sons, Inc., New York.

CONTENTS

PREFACE TO THE FIRST EDITION V
PREFACE TO THE SECOND EDITION VII
ACKNOWLEDGEMENT FOR USE OF ILLUSTRATIONS VIII
CONTENTS IX

INTRODUCTION 1

PART I

GENERAL PROPERTIES
OF LATTICE IMPERFECTIONS

CHAPTER I

THE PERFECT CRYSTAL

1. The crystallographic lattice 7
2. Physical properties 9
3. The vibrating lattice 16
4. Electronic structure of crystals 20

CHAPTER II

POINT DEFECTS IN CRYSTALS

1. Geometrical considerations 28
2. Energy of formation of point defects in crystals. 31
   2.1. Simple defects 31
   2.2. Association of point defects; crowdions and interstitialcies 33
3. Point defects in thermal equilibrium. 36
4. Formation of vacancies and interstitials in non-equilibrium concentrations 40
   4.1. Effects of impurities 40
   4.2. Effects of physical treatment 41

CHAPTER III

GEOMETRICAL DESCRIPTION OF DISLOCATIONS

1. Structure of dislocations 45
2. Displacements and stresses around dislocations 52

3. Interaction between dislocations 57
4. Interaction between dislocations and point defects 59
5. Motion of dislocations 62
5.1. Glide motion. 62
5.2. Climb motion 67
5.3. Motion of jogged dislocations 69
5.4. High-speed motion 70
6. Formation of dislocations 71
7. Types of dislocations occurring in various crystal structures 76

CHAPTER IV

MATHEMATICAL THEORY OF LATTICE IMPERFECTIONS

1. Introduction 80
2. Classical theory of elasticity 81
2.1. Stress, strain and their equations. 81
2.2. Differential equations for stresses and strains 86
2.3. Two-dimensional problems 88
3. Singularities in three-dimensional strain 90
4. Singularities in two-dimensional strain. 95
4.1. Plane strain 95
4.2. Anti-plane strain 102
5. Continuum theory of dislocations 104
5.1. Dislocations as sources of internal stress. 104
5.2. Description of imperfections 106
6. Classification of dislocations in elastic continua 110
6.1. Somigliana and Volterra dislocations 110
7. Force on a dislocation 112
8. Defects in crystalline solids. 113
8.1. Point defects 113
8.2. Dislocations 116

PART II

IMPERFECTIONS IN METALS

CHAPTER V

SPECIAL PROPERTIES OF LATTICE IMPERFECTIONS IN METALLIC CRYSTALS

1. Electrical resistance of lattice imperfections in metals 123
1.1. Vacancies and interstitials. 123
1.2. Dislocations 124
1.3. Magnetoresistivity of dislocations. 125
2. Thermal conductivity and lattice imperfections 126
3. X-ray and electron scattering by dislocations. 127

4. Rendering lattice imperfections visible 128
4.1. Etch pits and dislocations 129
4.2. Other methods for detecting individual dislocations 132

CHAPTER VI

PLASTIC DEFORMATION OF METALS CONSIDERED AS MOVEMENT OF DISLOCATIONS

1. Visual characteristics of plastic deformation 135
2. Critical shear stress for slip in metals and stress-strain relation 141
3. Dislocation theory of glide 144
4. Work-hardening 148

CHAPTER VII

PRODUCTION OF IMPERFECTIONS DURING PLASTIC DEFORMATION

1 Sources under stress in single crystals 153

CHAPTER VIII

WORK-HARDENING OF FACE-CENTRED CUBIC CRYSTALS

1. Interaction between dislocations 160
2. Stress-strain curve of face-centred cubic crystals 165
3. Work-hardening in easy glide 168
4. Work-hardening in region II 170
5. Third region of stress-strain curve. Influence of temperature on critical shear stress and flow stress 173
6. Deformation at different temperatures: work-softening 177

CHAPTER IX

THE INTERACTION OF DISLOCATIONS WITH IMPURITIES

1. Interaction between dislocations and impurity atoms 182
1.1. Yielding 182
1.2. Elastic interaction 184
1.3. Chemical interaction 188
1.4. Snoek ordering 189
2. Magnitude of interaction energy between dislocations and point defects 189
2.1. Elastic interaction 189
2.2. Electrical interaction 191
3. Locked dislocation under stress; microcreep 194
4. Thermal liberation of dislocations; delayed yield 198
4.1. Experimental 198
4.2. Theory of the release process 200

CHAPTER X

AGEING AND HARDENING OF ALLOYS

1. Strain ageing 208
2. Precipitation hardening 211
3. Solid-solution hardening, effect of order 216

CHAPTER XI

ANNEALING, POLYGONIZATION AND RECOVERY

1. Effects of annealing 220
2. Polygonization 221
3. Recovery as a thermally activated process 223
4. Grainboundary migration 227

CHAPTER XII

CREEP

1. Introduction 229
2. Transient creep 230
3. Steady-state creep 233
4. Creep in nearly perfect non-metallic crystals 235
5. Grainboundary creep and vacancy creep 236

CHAPTER XIII

SPECIAL TOPICS OF THE PLASTIC DEFORMATION OF METALS

1. Bauschinger effect 240
2. Size and surface effects 245
3. Plastic deformation of polycrystalline metals 246
3.1. Mathematical theory 246
3.2. Physical aspects of polycrystallinity 249
4. Deformation by mechanical twinning 251
5. Deformation by bending and kinking 255

CHAPTER XIV

NON-THERMAL GENERATION OF POINT DEFECTS

1. Generation of defects by plastic flow 259
2. Generation of point defects by irradiation 260
2.1. Nature of irradiation 260
2.2. Displacement theory 260
2.2.1. Neutrons 261
2.2.2. Protons, deuterons and other charged nucleons 263
2.2.3. Electron irradiation 268
2.2.4. Irradiation with photons 270

2.2.5. Replacement collisions and crowdions 271
2.2.6. Thermal and displacement spikes 272
2.2.7. Comparison of various irradiation procedures for the production of radiation damage. 275
3. Formation of defects during quenching 276

CHAPTER XV

PHYSICAL OBSERVATION OF DEFECTS IN METALS

1. Electrical resistivity. 283
1.1. General 283
1.2. Resistivity of cold-worked copper and other noble metals 284
1.3. Resistivity increase by irradiation 287
1.4. Resistivity change due to quenching 290
1.5. Change of magnetoresistivity during cold-work. 292
2. Thermal lattice conductivity 293
2.1. Theoretical 293
2.2. Experimental 296
3. Energy stored during cold work, irradiation and quenching. 298
4. Changes of volume and lattice parameter 299
5. Thermal recovery of influence of lattice imperfections 300
5.1. Survey 300
5.2. Interpretation of recovery in terms of diffusion processes 303
6. Effects of fast-particle bombardment and of quenching on the mechanical properties of metals 308

CHAPTER XVI

INFLUENCE OF LATTICE IMPERFECTIONS ON PHYSICAL PROPERTIES

1. Nuclear spin resonance in imperfect crystals 316
2. X-ray diffraction by lattice defects 322
2.1. Mosaic structure 322
2.2. Studies of line shape and line width. 325
2.3. Interpretation of X-ray diffraction phenomena in terms of dislocations 326
2.3.1. Measurable quantities 326
2.3.2. Deformed crystals 328
2.3.3. Microbeam experiments. 330
2.3.4. Cells 331
2.4. Comparison of X-ray dislocation densities with other determinations 331
3. Special kinds of X-ray diffraction 332
3.1. Small-angle scattering. 332
3.2. Anomalous transmission. 335
4. Neutron diffraction by lattice imperfections 337
4.1. General 337
4.2. Experimental method. 339
5. Dislocations and magnetic properties 343

CHAPTER XVII

INTERNAL FRICTION IN METALS

1. Mechanisms of internal friction 348
1.1. Relaxation and anelasticity 348
1.2. Relaxation spectrum 354
1.3. Mechanical hysteresis 355
1.4. Damped resonance 356
1.5. Viscous damping 358
2. Observations of internal friction 359
2.1. Experimental techniques 359
2.2. Dislocation damping 361
2.2.1. Dislocation relaxation (low-temperature or Bordoni internal friction) 361
2.2.2. High temperature dislocation damping 366
2.2.3. Dislocation resonance 368
2.2.4. Dislocation hysteresis 372
2.2.5. Transient cold-work internal friction 376
2.2.6. Deformation hysteresis 378
2.3. Internal friction not exclusively associated with dislocation motion 381
2.3.1. Visco-elastic after-effect and microcreep 382
2.3.2. Temperature-activated hysteresis 382
2.3.3. Grain boundary relaxation 383
2.3.4. Point defect internal friction 385
2.4. Internal friction related to the presence of foreign solute atoms in the metal 386
2.4.1. Snoek damping in body-centred cubic metals 386
2.4.2. Snoek damping in face-centred cubic metals and alloys . . 388
2.4.3. Stress-induced ordering in substitutional solid solutions . . 389
2.4.4. Köster damping in interstitial solid solutions 390
2.4.5. Internal friction due to phase transformations 393
2.5. Thermal, magnetic and electronic damping in metals 394
2.5.1. Thermoelastic damping 394
2.5.2. Magnetoelastic relaxation and magnetic aftereffect (in ferromagnetic materials) 394
2.5.3. Ferromagnetic (magneto-mechanical) hysteresis 395
2.5.4. Electron attenuation of ultrasonic waves and magneto-acoustic resonance 395
3. Survey 397

CHAPTER XVIII

DIFFUSION IN CRYSTALS

1. General theory of diffusion 404
2. Atomic theory of diffusion 407
2.1. Vacancy diffusion 408
2.2. Interstitial diffusion 411
2.3. Interstitialcy diffusion 411
2.4. Chemical diffusion, special mechanisms 412
3. Influence of pressure and temperature 414

4. Kirkendall effect 417
5. Diffusion along grain boundaries 422
5.1. Grain boundary diffusivity 422
5.2. Effect of orientation 426
6. Enhanced diffusion along individual dislocations, and owing to point-defects in non-equilibrium 429

CHAPTER XIX

GRAIN BOUNDARIES

1. Geometry of grain boundaries 435
2. Energy of a dislocation grain boundary 442
3. Motion of dislocation grain boundaries 446
4. Large-angle grain boundaries 446
5. Interaction between grain boundaries and other lattice imperfections . 447
5.1. Interaction with impurities and point defects 447
5.2. Interaction between dislocations and grain boundaries 450
6. Special types of grain boundaries; subboundaries; coherent and semi-coherent boundaries 450

CHAPTER XX

RECRYSTALLIZATION AND GRAIN GROWTH

1. Phenomenology 454
2. Nucleation and growth of grains—formal theory 460
3. Recrystallization and recovery 463
4. Recrystallization textures 465

CHAPTER XXI

PHASE TRANSFORMATION AND PRECIPITATION

1. Transformation 470
2. Formation of annealing twins and martensitic transformations 470
3. Diffusion-controlled transformations 476
3.1. General 476
3.2. Influence of plastic deformation and irradiation on precipitation and phase transformation 477
3.3. Guinier-Preston zones, age-hardening 479
3.4. Order-disorder transformations 481

CHAPTER XXII

FATIGUE AND FRACTURE

1. Fatigue 487
2. Fracture 492
3. Intercrystalline cracking 498

CHAPTER XXIII

CRYSTAL GROWTH AND IMPERFECTIONS

1. Growth of crystals as a dislocation mechanism 500
2. Special growth phenomena 506
2.1. Polytypism 506
2.2. Whisker growth 506

PART III

IMPERFECTIONS IN NON-METALLIC CRYSTALS

CHAPTER XXIV

FUNDAMENTAL PROPERTIES OF IMPERFECTIONS IN POLAR CRYSTALS

1. General distinctions between imperfections in metals and non-metals . 511
2. Point defects in polar crystals 512
3. Equilibrium of point defects in polar semiconductors 515
3.1. Polar semiconductors 515
3.2. Equilibrium reactions 517
3.3. Computation of concentrations 520
3.4. Impure crystals 523
4. Diffusion and ionic conductivity 525
5. Dislocations in polar crystals 528
6. Electronic conduction and colour centres 532

CHAPTER XXV

LATTICE DEFECTS AND THE ELECTRICAL PROPERTIES OF IONIC CRYSTALS

1. Intrinsic ionic conductivity 537
2. Dislocations and ionic conductivity 539
3. Impurity conduction 542
4. Diffusion 547

CHAPTER XXVI

OPTICAL PROPERTIES OF DEFECTS IN IONIC CRYSTALS

1. Colour centres in the alkali halides 553
2. Photoconductivity 563
3. Silver halides, photolytic darkening 563

CHAPTER XXVII

PLASTIC DEFORMATION AND IRRADIATION OF IONIC CRYSTALS

1. General features 567
2. Joffé-effect and surface sources 570
3. Mutual interaction between dislocations and point defects 572
3.1. Colorability by ionizing radiation. 572
3.2. Internal friction and dielectric loss 574
3.3. Mechanical properties of irradiated and quenched crystals 576
4. Optical absorption and plastic deformation (not colour centres) 577
4.1. Fundamental absorption 577
4.2. Microscopic evidence for dislocations in polar crystals. 578
4.3. Microscopic evidence for the generation and motion of dislocations 579
5. Electron and neutron irradiation effects in polar crystals. 580

CHAPTER XXVIII

THE ELECTRICAL PROPERTIES OF ELEMENTAL HOMOPOLAR SEMICONDUCTORS (GERMANIUM AND SILICON)

1. Electrical conductivity 583
2. Effect of dislocations on the electrical properties of germanium and silicon 592
3. Point defects 597

CHAPTER XXIX

DISLOCATIONS AND OTHER IMPERFECTIONS IN HOMOPOLAR CRYSTALS WITH THE DIAMOND STRUCTURE

1. Dislocations in the diamond structure 599
1.1. General 599
1.2. Simple dislocations 600
1.3. Composite dislocations 604
1.4. Partial dislocations and stacking faults 607
2. Point defects in the diamond structure 610

CHAPTER XXX

PLASTIC DEFORMATION AND IRRADIATION OF MATERIALS WITH THE DIAMOND STRUCTURE, ESPECIALLY OF GERMANIUM

1. General description of plastic properties of homopolar substances . . . 612
2. Low and medium-temperature deformation of germanium 614
2.1. Creep under constant stress 614
2.2. Deformation at a prescribed strainrate 617
3. Dislocations, sliplines and etchpits 619

4. Deformation at very high temperature. Thermal stresses 620
5. Dislocation theory of plastic creep of nearly perfect crystals 622
5.1. Theory of creep 622
5.2. Generation and motion of dislocations in the diamond lattice . . 625
5.3. Work-hardening in germanium and silicon. 629
6. Mechanical twinning in the diamond lattice 630
7. Plastic deformation and the electrical properties of germanium and silicon 633
7.1. Conduction of electricity 633
7.2. Formation of acceptors by deformation and quenching 636
7.3. Scattering by lattice imperfections 640
7.4. Dislocations and lifetime 641
8. Irradiation and quenching of diamond structure materials 642

CHAPTER XXXI

DEFECT DIFFUSION AND INTERNAL FRICTION IN GERMANIUM AND SILICON

1. Diffusion in germanium and silicon 646
2. Precipitation of copper and nickel in germanium 647
3. Diffusion and dislocations 654
4. Internal friction in germanium and silicon 656

NAME INDEX. 661

SUBJECT INDEX 669

INTRODUCTION

Since the middle of the thirties, the interest in the nature and properties of physical imperfections in crystals has risen rapidly. In those years it was properly realized that the theory of perfect crystals (the lattice theory as developed by Born and others) was not able to explain in a satisfactory manner the so-called *structure sensitive* properties of crystalline solids. Among these are counted, to name but a few diverse examples, the behaviour of metals and many other crystals under stress, the phenomena of crystal growth and recrystallization, diffusive properties, optical and dielectrical behaviour of insulating crystals, photoconductivity and luminescence and many other mechanical, optical and electrical properties. Even the simple transport phenomena that apparently yielded to a lattice-theoretical approach, such as the electrical and thermal conductivity of metals, were found not to be wholly explicable in all their features by taking account of lattice vibrations alone, although this in itself already constituted a departure from the idea of a perfect crystal. Crystalline imperfections of a completely different nature determine to a large extent the observed structure sensitivity.

The amount of literature on the subject of imperfections in crystals is very large, and what is worse, it is extremely diverse. Therefore, students of the subject can often only pay attention to a small fraction of it. For instance, investigators in the field of plasticity of metals are sometimes only dimly aware of the important effects on the properties of, say, ionic crystals, of precisely the same kinds of imperfections as they are themselves concerned with in metals. This is regrettable. Although the observational results of a certain experimental procedure may appear wholly distinct in different types of crystals, the under-

lying processes are often quite similar, and the study of them in one type of crystal may greatly facilitate the understanding of particular results obtained in another. To mention but one fairly typical example: whereas in most metals the low-temperature resistivity is moderately *in*creased by plastic deformation, the resistivity of ionic crystals is usually *lowered* by not less than two orders of magnitude after precisely the same treatment. Nevertheless, these opposite effects have both the same cause, viz. the formation of vacancies (and interstitial atoms) in the crystal lattice by the action of dislocations set in motion by the applied stress.

A number of excellent reviews, more or less covering the whole field of imperfections in crystals superficially, have appeared in literature. We only mention here Seitz's well-known paper on crystal imperfections (SEITZ [1952]), which has been followed by a number of Conference reports (e.g. FISHER *et al.* [1957]). Several authors have prepared extensive monographs dealing with a more restricted subject, of which we name here the books by COTTRELL [1953], READ [1953] and FRIEDEL [1956] on dislocations. Recently an interesting treatise dealing with the more practical aspects of the problem in all its diversity has been published (GRAY [1957]). Finally extensive review papers on one or two related aspects appear regularly in annual reports like Progress in Metal Physics, Solid State Physics, etc. The tremendous rate of progress of research in the field is well illustrated by the fact that most of these publications, however valuable, even after a few years already cannot any longer be regarded as reliable. This is especially so where work concerned with several aspects at once of the wide field of crystal imperfections is considered. One of the causes of this is the lack of crossreference already referred to.

The present author is well aware of this but, though realizing that the expansion proceeds at an ever increasing rate, he judges that there is a real need for a work of the latter character, discussing the influence and effects of crystal imperfections in a large diversity of substances and applications. Moreover, much substantial progress in the field has been made recently, which is not contained in any of the established reviews. To name but a few: in the last few years we have seen a rapid development of the study of the influence of physical lattice defects on the properties of homopolar semiconductors; a very large amount of work has been done in the field of internal friction in many types of solids; the atomic theory of plasticity of metals has been practically

completed and has been verified in several of its aspects by direct visual observations. It is therefore appreciably more reliable now than a few years ago, when only hypotheses existed, many of which have now been verified experimentally. The modern progress in reactor technology has stimulated fundamental investigations of the influences of nuclear radiation on material properties. Monographs on this subject have been published by DIENES and VINEYARD [1958] and by BILLINGTON and CRAWFORD [1961]. Many other instances could be mentioned. Knowledge of progress in one direction serves to promote work in other directions. It is to facilitate this that this book has been planned.

However, the author had necessarily to limit his scope.

The attention in this book is primarily focused on the phenomena related to the *physical* lattice imperfections such as dislocations, vacancies, interstitial atoms, and on their combinations, e.g. vacancy pairs and larger clusters of point defects, dislocation arrays such as grain boundaries, etc. It did not seem practical also to include in the book a discussion of *chemical* imperfections in crystals. Consequently, the variety of effects that are caused by the introduction of foreign atoms (in more than very small traces) are not systematically dealt with. For instance, the phenomenon of fluorescence remains undiscussed, the fluorescence of crystals being intimately related to the presence of foreign atoms. Again, the important electrical properties of donor and acceptor elements in the semiconductors germanium and silicium are only briefly mentioned and not treated in detail. Even those phenomena peculiar to metallic alloys, such as ordering processes, nucleation and growth of precipitates etc. are only very briefly described.

A second category of perturbations of internal crystalline order is related to the distribution of electrons in the crystal. These *electrical* imperfections, a systematic survey of which was first made by SEITZ [1952], we shall not discuss either. The reason for these omissions is a very simple one; the inclusion of these topics in any detail would mean another 500 printed pages, if not more.

The book consists of three parts. The first part deals with the *physical properties* of the three principal imperfections: dislocations, vacancies and interstitial atoms, in as far as these properties are inherent in such imperfections in all common types of crystalline solids. Those physical properties of imperfections that vary essentially as between different materials, are not yet treated; their discussion has been postponed to appropriate chapters in the following parts.

The second part is devoted to the influence of imperfections on the physical and mechanical properties of *metals*. Most of it concerns the plastic deformation of metals; attention is paid further to physical effects of imperfections, such as their influence on the electrical resistivity, on diffusion and on the phase transformations.

The third part contains a discussion of the effects of imperfections in some *non-metallic solids*. This is probably the least definite portion of the book, as the phenomena caused by imperfections in non-metallic solids are extremely complex and diverse and have not yet been studied as systematically as they have been in metals. By far the most important groups of materials in this class are, firstly, the binary ionic compounds, of which the alkali- and silver halides are the best investigated representatives, and, secondly, the diamond lattice semiconductors germanium and silicon. Characteristic evidence on defect properties of other kinds of crystalline non-metals are so scarce, that their incorporation in a work of this kind does not seem very opportune.

The study of crystal imperfections constitutes one of the frontiers of modern physics. Hence, the results obtained so far are hardly to be regarded as final. However, they are so interesting that their discussion may prove inspiring to the reader. And this is, as has been said before, the sole purpose of this book.

REFERENCES

BILLINGTON, D. S. and J. H. CRAWFORD (1961); Radiation Damage in Solids (Princeton Un. Press).

COTTRELL, A. H. (1953); Dislocations and plastic flow in crystals (The Clarendon Press, Oxford).

DIENES, G. J. and G. H. VINEYARD (1958); Radiation effects in solids (Interscience Publ., New York).

FISHER, J. C., W. G. JOHNSTON, R. THOMSON and T. VREELAND (1956); Mechanical properties of crystals (Wiley, New York).

FRIEDEL, J. (1956); Les dislocations (Gauthiers-Villars, Paris).

GRAY, T. J. a.o. (1957); The defect solid state (Interscience Publ., New York).

READ, W. T. (1953); Dislocations in crystals (McGraw-Hill, New York).

SEITZ, F. (1952); Advances in Physics **1**, 43.

PART I

GENERAL PROPERTIES OF LATTICE IMPERFECTIONS

CHAPTER I

THE PERFECT CRYSTAL

1. The Crystallographic Lattice

Crystalline materials are made up of one or more crystals, or *grains*, separated from each other by *grain boundaries*, and bounded elsewhere by the surface of the solid. Each individual crystal can be considered as a three-dimensional arrangement of identical *unit cells*, each containing the same number of atoms or ions. The unit cells are arranged on a periodic *space lattice*, that is, each particle finds an identical one after a translation over the distance

$$r = la + mb + nc\,, \tag{1}$$

a, b, and c being the three so-called primitive lattice vectors and l, m, n being whole numbers. There are 14 of these space or Bravais lattices. According to the distribution of symmetry elements over the space lattice one can distinguish between 230 *space groups*, which form a subdivision of the 32 possible *crystal classes*, each of which in its turn belongs to one of the 7 *crystal systems*: triclinic, monoclinic, (ortho)-rhombic, tetragonal, cubic, hexagonal and rhombohedral. All crystals belonging to a given crystal system possess a certain *minimum symmetry*. In the cubic system (three equal mutually perpendicular primitive vectors (axes)) we always encounter four 3-fold rotation axes (along the diagonals) and in the hexagonal system (two equal a-axes at an angle of 120° and a c-axis of different length perpendicular thereto) the minimum symmetry is one 6-fold rotation axis. Combination of these symmetry elements with other symmetry axes or planes gives the 32 crystal classes (or *point groups*). If we also consider axes and planes coupled with translations within the dimensions of the unit cell we come back at the 230 space groups. More on the fundamentals of

crystallography can be found in any textbook and will accordingly be omitted here (see also SMITH [1958]).

The cubic system contains three types of space lattice; simple cubic, body-centered cubic and face-centered cubic. Most metals crystallize in simple structures based on one of the two latter lattices (or allotropically in both), or in the close-packed hexagonal structure, where the axial ratio c/a is nearly 1.63, corresponding to a close-packed array of spheres. The semiconductors germanium and silicon, other materials of great technical importance nowadays, crystallize in the socalled *diamond structure*, as does diamond itself. This structure can be regarded as consisting of two interpenetrating face-centered cubic lattices, mutually displaced along a common space diagonal by $\frac{1}{4}$ of the length of the latter. The diamond structure is a characteristic structure to accommodate tetravalent atoms; the coordination number is 4. There exists an extended group of more or less homopolar binary compounds, in which the sites of a diamond lattice are occupied in such a way that both kinds of atoms are regularly distributed over the sites. This is called the *zincblende* structure after the prototype ZnS. Closely related to this is the *wurtzite* structure (prototype the other polymorphic modification of ZnS and ZnO), that stands in the same relation to the zincblende structure as does the hexagonal close packing of spheres to the cubic close packing. (As is well-known, this difference lies in the sequence of stacking of the various {111}-layers.) The wurtzite structure has hexagonal symmetry, yet everywhere preserves the tetrahedral configuration of the four nearest neighbours of any atom, and thus the 4-fold coordination is retained. There is a continuous transition between AB compounds crystallizing in these two structures and AB compounds of a more ionic nature. A typical transition between ionic and homopolar crystal is formed by AgI; this crystallizes in the hexagonal wurtzite structure. The more ionic the bonding forces are, the stronger is the tendency to crystallize in a system with a higher co-ordination number (but still cubic). The alkali as well as the silver halides (except AgI) crystallize either in the *sodium chloride structure* or in the *caesium chloride structure*. The former consists of a face-centred cubic lattice of positive (Na^+) ions combined with a face-centred cubic lattice of negative (Cl^-) ions, translated with respect to the former over half the distance of a primitive vector. Oppositely charged ions then occupy neighbouring sites. The caesium chloride structure is a combination of two simple cubic lattices occupied by oppositely charged ions and

translated relative to each other over half of the common body diagonal. Hence, a negative ion occupies the centre of a cube, the corners of which are occupied by positive ions, and vice versa.

The preference of metals for crystallizing in close-packed structures or in the only slightly less dense body-centred cubic lattice is due to the peculiar bonding character of metal atoms, that tends to reduce the *volume* as far as possible, without regard to the coordination number or the electrostatic forces (the latter being neutralized fairly completely by the electron gas surrounding each ion). Apparently there exists a close correlation between the nature of the crystal structure and the physical and mechanical properties of solids. This is once more illustrated in the following Table.

TABLE 1

Atomic forces, properties and lattice structure

Bonding force	*Coordination number*	*Crystal structure*	*Hardness and ductility*	*Electrical conductivity*
Metallic	as high as possible	close packed or nearly so.	ductile, little temp. dependent	large, decreases with increasing temp.; electronic.
Ionic	high	closest approach of oppositely charged ions.	limited ductility little temp. dep.	moderate, increases with increasing temp.; ionic.
Homopolar	low	fixed number of rigidly directed bonds	very brittle, except at high temp.	low, very dependent on purity, increases with temp.; electron and hole conduction

It is a matter of much interest, whether these different properties are directly due to differences in the nature of the bonding forces themselves, or only indirectly, e.g. by the intermediate of the tendency for the incorporation of foreign atoms or physical defects, which, in their turn, then affect the crystal properties in the manner described. This is one of the questions to which we shall often refer in this book.

2. Physical Properties

The perfect crystal is one in which the regular periodic arrangement of the atoms extends everywhere; such crystals do not exist. In the

first place, all atoms and ions vibrate about their equilibrium positions; this thermal vibration is a form of heat motion, modified by the periodic structure. As it is a quantum mechanical phenomenon, the energy remains finite, even at the absolute zero; the vibrations are then characterized by the *zero-point energy* of $\frac{1}{2}h\nu$ per oscillating atom. It is usually assumed that the potential energy governing the motion of the atoms (or ions) is nearly harmonic, that is, the interaction energy of an atom with its surroundings, regarded as a function of the displacement of the atom relative to its neighbours, contains primarily a quadratic term. With this assumption it becomes possible to develop a consistent theory of lattice specific heat and heat conduction. However, many phenomena, as for instance the thermal dilatation, depend primarily on the higher order (non-harmonic) terms, also present in the series development of the potential energy; these cause a shift of the mean equilibrium positions (the quadratic terms leave these unaltered).

BORN [1951] has pointed out that the occurrence of nonharmonic terms might be of a much farther reaching importance than was first thought. The thermal dilatation, for instance, is governed by an expression of the form (GRUENEISEN [1908]):

$$\frac{\delta l}{l} = \frac{\alpha}{k}\varepsilon(T), \tag{2}$$

where α is the coefficient of linear dilatation and $\varepsilon(T)$ the mean energy of an oscillator; k is Boltzmann's constant. At high temperatures $\varepsilon = kT$, thus $\delta l/l = \alpha T$, as must be. Now, δl will become equal to a, the atomic distance, when $l = na$, where $n = k/\alpha\varepsilon$. Even at low temperatures, $\varepsilon(T)$ remains finite; in the limit when $T = 0$ $\varepsilon(T)$ has the zero point value $\frac{1}{2}h\nu$, where we can put $h\nu = k\theta$, θ being some parameter like the Einstein temperature (see section 3). Since this is about 100 °K for all solids, n is easily found to be of the order 1000 at 0 °K.

At the melting point n is somewhat less, but not very much. This argument applies only to linear chains; in two- or threedimensional lattices the difficulty can be bypassed by considering closed chains of atoms. Born's original conclusion, that crystals of linear dimensions larger than about a thousand atomic distances could not be perfect, because the displacements due to the anharmonic terms integrate to an atomic distance or more, cannot therefore be maintained. This was pointed out by the same author in a later paper, in which it was argued that, although it was in any case impermissible to apply the

simple harmonic lattice theory to large crystals without due regard to anharmonicity, and the fundamental equations of lattice theory should therefore be changed, perfect crystals of large size might still exist. This is because the extended higher order differential equations that replace the harmonic equations might allow a possible solution identical with that of the harmonic approximation. In other words, this solution corresponds to an arbitrarily extended perfect (though vibrating) lattice.

The lattice of actual crystals is not only disturbed by thermal vibrations, but also by the presence of physical lattice imperfections.

An obvious classification of the possible imperfections is the following:

1) Zero-dimensional or point imperfections: vacancies, interstitial atoms, colour centres and their combinations: vacancy pairs, clusters, etc.
2) One-dimensional or line imperfections: dislocations.
3) Two-dimensional or surface imperfections: grain boundaries, twin boundaries, phase boundaries, stacking faults, the crystal surface.
4) Three-dimensional or volume imperfections: voids, inclusions of a second phase, etc.

Of these, all three kinds occur in large numbers in metals. In fact, it appears extremely difficult and up till now has not yet been possible, to obtain macroscopic pieces of crystalline material without both dislocations and point defects contained in it. In homopolar crystals the presence of all kinds of imperfections has been demonstrated as well as in metals although it is much easier here to decrease the concentration of one- or two-dimensional imperfections to very small values. The reason for this is not quite clear. In ionic crystals the neutrality condition poses certain restriction on the occurrence of defects of any specific kind; however, these crystals also usually contain very many imperfections.

The occurrence of point defects in crystalline solids is a quite obvious matter; by introducing a vacancy or an interstitial into a perfect lattice the possible number of distinguishable arrangements of the atoms increases suddenly from 1 to an enormous number, of the order of 10^{23} or more in most cases. That is, the introduction of point defects increases the *entropy*, and at high enough temperatures this compensates to a large extent for the energy of formation (section II–3).

Dislocations and grain boundaries, on the other hand, affect the crystal entropy only by a small amount and, moreover, represent themselves a kind of order, whereas their total energy of formation is very large. Therefore, they cannot exist in measurable concentrations as thermodynamically stable imperfections. The question of their actual origin is a most intriguing one.

Before treating this and other problems in more detail, we shall discuss a few properties of the perfect lattice which shall be needed here and there in the following. The discussion must necessarily be very brief; for more details we refer, for instance, to the book on lattice theory by BORN and HUANG [1954], and to LEIBFRIED'S paper [1955].

The cohesive properties of solids depend directly on the forces between the atoms in the perfect crystal. In ionic crystals these forces are mainly *electrostatic*. In the compound crystals of the type AB the charges per ion are $\pm ne$, and the electrostatic lattice energy per ion pair is of the form

$$E_i = -\frac{n^2e^2}{a} z_i$$

where a is the distance between two nearest ions of different charge, and z_i is *Madelung's number*, which depends on the crystal structure. It is of order unity, being 1.762 for the CsCl, 1.747 for the NaCl structure, and 1.639 and 1.638 for the zincblende and wurtzite structures respectively. Although it has the physical meaning of an effective number of neighbours, it is much smaller than the real coordination number. This is due to the "long range" electrostatic interaction with next-nearest and further neighbours. The electrostatic attractive forces are not the only ones active, there are also *repulsive forces*, that prevent the lattice from collapsing. These always depend very sensitively on the distance between the ions, as they arise from the interaction between the closed electron shells of the ions. This interaction only becomes appreciable when the wavefunctions belonging to the two shells overlap. As the radial parts of these wave functions have exponential character, the interaction rises very rapidly from zero to a large value. The interaction between ions is further complicated by the finite polarizability and the existence of excited states of the ions or atoms. These give rise to a third type of interaction (that is again attractive), namely the *van der Waals attraction*, which is inversely proportional to a high power of the distance (6 or 7). A fourth category

of forces acting between the ions is formed by the *exchange forces;* these are responsible for the valence-type bonding encountered in not purely ionic crystals. When the constituting atoms or ions have closed shells, as in the case of ionic crystals, the electrostatic and repulsive (overlap) forces are the most important of the four types mentioned and they determine the crystal structure. The repulsive forces being of narrow range, the electrostatic forces not, it is easily seen that the crystal structure will tend to be such that oppositely charged ions practically "touch" each other. It is therefore that one can speak of "ionic radii", although strictly speaking these depend on the crystal structure as well as on the nature of the ion in question. In the case of nearly equal ionic radii of the two partners in a binary crystal, the more close-packed caesiumchloride structure with larger coordinate number (8) is favoured over the sodiumchloride structure (6). However, if the radii differ, there is a greater chance to encounter somewhat more loosely packed structures such as the latter.

In metals and homopolar semiconductors, constituted of atoms that have incompletely filled outer shells, the interaction energy is made up by quite a number of contributions. Metals, at least the monovalent metals, are most easily described by the model of quasi-free valence electrons moving in an effective potential. The latter is built up by the positive ions with their filled inner electron shells and has the periodicity of the crystal. This schematic picture, if due regard is paid to the Pauli principle governing the electronic energies, leads to the well-known energy band scheme for the metal electrons, in which the conduction electrons occupy part of the numerous energy levels of the so-called conduction band. The energy per lattice cell, consisting of one or more ions and their surrounding electron clouds, is then approximately equal to the sum of: 1) the electrostatic energy of the system formed by the positive ions and the electron gas; 2) the kinetic energy of the electrons; 3) the repulsive energy between the closed inner ionic shells (this is most important in the noble metals and semiconductors); 4) the exchange energy between ions, and 5) that between electrons.

The exchange forces between the ions are most important in the homopolar semiconductors and also in ferromagnetics. The kinetic energy of the conduction electrons in its turn is the sum of the energy of the lower edge of the conduction band and the Fermi energy of the electrons.

Once the potential energy of the constituents of a crystal is known

as a function of the ionic coordinates, the elastic behaviour of the crystal can be computed, albeit only for small deformations (the ionic displacements must be small with respect to a lattice parameter). One distinguishes between homogeneous and inhomogeneous deformations. In the first case, as realized for instance in uniform compression, the lattice remains a perfect periodic structure and the stresses and strains in the material are independent of position. The components (u_i) of the displacement of each particle originally on the site (x_j) are then given by an expression of the kind

$$u_i = \sum_j v_{ij} x_j \qquad (i, j = 1, 2, 3),$$

where the v_{ij} form a symmetric tensor. The energy per unit volume connected with the deformation is some function of these displacements; in equilibrium the energy density must be minimum, thus linear terms in v_{ij} can not occur in it. Therefore, the energy density can in first approximation be written as

$$w = \sum_{i,j} \sum_{k,l} C_{ijkl} v_{ij} v_{kl} , \tag{3}$$

where the *generalized elastic moduli* C_{ijkl} depend on the interatomic forces.

The quantities v_{ij} are the 9 *strain* components $\partial u_i / \partial x_j$; the coefficients C_{ijkl} form the components of a tensor of the fourth rank, the tensor of the *elastic constants*. Usually this tensor is presented in an abbreviated form, with only 36 components $c_{11}, c_{12}, \ldots, c_{16}, \ldots, c_{66}$ instead of 81. This is possible owing to the symmetry of the strain tensor. By differentiating the energy density with respect to the elastic strains, we find the symmetric tensor of the *elastic stresses*:

$$t_{ij} = \frac{\partial w}{\partial v_{ij}} = \sum_{k,l} C_{ijkl} v_{kl} . \tag{4}$$

This formulates Hooke's law for crystals. The linear expression can also be written in the covariant form

$$v_{ij} = \sum_{k,l} S_{ijkl} t_{kl} , \tag{5}$$

where the S_{ijkl} form the tensor of the generalized *elastic coefficients*, hat is reciprocal to C_{ijkl}. Also for S mostly an abbreviated notation s_{ij}, with only 36 components, is used.

Since it is easily shown that $c_{ij} = c_{ji}$ ($C_{ijkl} = C_{klij}$), the number of *independent* elastic constants c_{ij} is 21 at the utmost. When the forces with which the atoms or ions interact are central forces, and when moreover each particle occupies a centre of symmetry, the so-called *Cauchy relations* hold:

$$\begin{aligned} &c_{23} = c_{44}, \qquad c_{31} = c_{55}, \qquad c_{12} = c_{66}; \\ \text{and} \qquad &\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6) \\ &c_{14} = c_{56}, \qquad c_{25} = c_{64}, \qquad c_{36} = c_{15}; \end{aligned}$$

and the number of independent constants is reduced to 15. In hexagonal crystals only 6 elastic constants are independent, because for reasons of symmetry:

$$\begin{aligned} &c_{11} = c_{22} = c_{33}; \\ &c_{12} = c_{23} = c_{31} = c_{21} = c_{32} = c_{13}; \\ &c_{44} = c_{55} = c_{66}. \end{aligned}$$

Usually, only five constants remain, that is, when the Cauchy relation $c_{12} = c_{44}$ applies. In cubic crystals even we need only three independent constants for a complete description. Isotropic materials finally are characterized by only two constants c_{11} and c_{44}. Here, the technical elastic constants: *Young's modulus* $E = c_{11} + 2c_{44}$, and *shear* or *rigidity modulus* $G = c_{44}$ are mostly used.

The Cauchy relation $c_{12} = c_{44}$ is rather well obeyed in the cubic crystals of binary ionic compounds. This proves that the central electrostatic forces dominate. In the monovalent metals, the elastic constants are principally determined both by the (non-central) Coulomb forces and the (central) ion-ion closed shell interaction. Therefore, the Cauchy relations do not apply. Neither do they apply to the crystals of homopolar materials, the exchange forces being certainly not central either.

In practical cases the binding and cohesive forces of a nonisotropic crystalline material are often characterized sufficiently completely by the already mentioned technical quantities: Young's modulus E and rigidity or shear modulus G, which then, of course, vary with the orientation of the applied stress, and by the scalar compressibility $\varkappa$. These can, for each crystal class, be expressed in the elastic coefficients s_{ij}, or, but usually less elegantly, in the elastic constants c_{ij}. In a cubic crystal

stressed homogeneously in one direction, making angles the cosines of which are α, β, γ with the cubic axes, one has:

$$\frac{1}{E} = s_{11} - 2\left[(s_{11} - s_{12}) - \tfrac{1}{2}s_{44}\right](\alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2),$$

$$\frac{1}{G} = s_{44} + 4\left[(s_{11} - s_{12}) - \tfrac{1}{2}s_{44}\right](\alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2), \tag{7}$$

$$\varkappa = s_{11} + 2s_{12}.$$

The quantity $(s_{11} - s_{12}) - \frac{1}{2}s_{44}$ forms a direct measure for the anisotropy of the crystal; it is zero in isotropic solids. For a detailed discussion of these aspects of crystal elasticity the reader is referred to HUNTINGTON [1958].

3. The Vibrating Lattice

A crystal is a set of N particles, held together by certain forces and distributed periodically in space. Such a system, acted upon by a disturbing force, performs vibrations. These can most conveniently be described by a set of $3N$ so-called *normal coordinates* q_i, which are linear combinations of the displacements of all particles and each vary harmonically in time, with the frequency ν_i. These motions are *normal modes* of the lattice. The set of normal modes of a crystal can be directly compared to a set of $3N$ independent harmonic oscillators with frequency ν_i. According to quantum mechanics, the eigenvalues of the energy of an oscillator are $\varepsilon_{i0} = \frac{1}{2}h\nu_i$; $\varepsilon_{i1} = \frac{3}{2}h\nu_i$, $\varepsilon_{i2} = \frac{5}{2}h\nu_i, \ldots$, etc., and the statistical partition function of the oscillator labelled i is thus:

$$Z_i = \mathrm{e}^{-\frac{1}{2}h\nu_i/kT} \sum_{n=0}^{\infty} \mathrm{e}^{-nh\nu_i/kT} = \frac{\mathrm{e}^{-\frac{1}{2}h\nu_i/kT}}{1 - \mathrm{e}^{-h\nu_i/kT}}. \tag{8}$$

The part of the free energy of the crystal that is due to this normal mode is given by

$$F_i^{\mathrm{vibr}} = -kT \ln Z_i,$$

thus by:

$$F_i^{\mathrm{vibr}} = \tfrac{1}{2}h\nu_i + kT \ln (1 - \mathrm{e}^{-h\nu_i/kT}). \tag{9}$$

The total free energy of the lattice is the sum of the contributions of

all normal modes and the energy W of the static lattice with each particle in its average position. Thus

$$F = W + \tfrac{1}{2}\sum_i h\nu_i + kT\sum_i \ln\left(1 - e^{-h\nu_i/kT}\right). \tag{10}$$

Regarding F as a function of T and V only, where V is the volume of the crystal, the following thermodynamic relations apply: For the energy of the lattice:

$$E = F + TS = F - T\left(\frac{\partial F}{\partial T}\right)_V;$$

and for the specific heat:

$$c_V = \left(\frac{\partial E}{\partial T}\right)_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_V.$$

One then easily finds:

$$E = W + \tfrac{1}{2}\sum_i h\nu_i + kT\sum_i \frac{e^{h\nu_i/kT}}{e^{h\nu_i/kT} - 1}, \tag{11}$$

$$c_V = k\sum_i \frac{(h\nu_i/kT)^2\, e^{h\nu_i/kT}}{(e^{h\nu_i/kT} - 1)^2}. \tag{12}$$

The second term in E represents the well-known zero-point energy of the lattice. From (11) and (12) it follows that at high temperatures ($kT \gg h\nu_i$) the temperature dependent part of E becomes simply kT (for each degree of freedom), and c_V approaches $3kN$, which is nothing but the well-known law of Dulong and Petit.

For quantitative application, a knowledge of the frequencies of all normal modes is required. In the so-called *Einstein model*, all frequencies are taken to be the same. This corresponds to the assumption of a set of equal and independent lattice oscillators. Then in all expressions only one simple parameter, ν_E, or the related parameter $\theta_E = h\nu_E$, the so-called *Einstein temperature*, occurs. Another, somewhat more realistic approach to the lattice dynamics was made by Debye. He assumed that the distribution function of the various frequencies can be approximated by

$$\begin{aligned} f(\nu)d\nu &= C\nu^2 d\nu \qquad && \nu \leqslant \nu_D, \\ f(\nu)d\nu &= 0 && \nu > \nu_D. \end{aligned} \tag{13}$$

The normalization condition

$$\int_0^\infty f(\nu)\mathrm{d}\nu = 3N$$

leads immediately to $C = 9N/\nu_D^3$. The quantity ν_D is called *Debye frequency*; often quoted is the related parameter $\theta_D = h\nu_D/k$, the *Debye temperature*. It will be clear that θ_E represents a kind of average of all θ_D pertaining to the various modes. To describe the observed crystal properties, the appropriate θ_E must therefore always be chosen smaller than the appropriate θ_D. In practice, the ratio amounts to about 0.7. Table 2 lists the Debye temperatures of the materials most referred to in this book.

TABLE 2

Debye temperatures of some crystalline materials†

Material	θ (°K)	*Material*	θ (°K)
Ag	229	AgBr	144
Al	375	AgCl	183
Au	164	CaF_2	474
Cu	343	KBr	177
Fe	355	KCl	227
K	100	KI	115–200
Li	430	LiF	607–750
Mo	360	NaCl	281
Na	160		
Ni	413	Ge	250–400
Pt	233	Si	660
U	200	C (diamond)	1860
W	270	ZnS	300
Zn	235		

† From the American Handbook of Physics.

In both models the specific heat per mole is a universal function for all crystals, when the temperature is measured in units of T/θ_E or T/θ_D respectively (reduced temperature). By comparing a number of measurements on different materials, and adjusting the characteristic temperatures suitably, one can thus check the applicability of either assumption. Except at very low temperatures, where the Einstein model fails, both models can be made to agree excellently with the observations. On the other hand, the Debye model is certainly correct in the limit of extremely low temperatures, when only the lowest frequencies

are excited. These correspond to elastic waves of very long wavelengths in the crystal. The discreteness of the atomic structure then becomes relatively unimportant and the solid can be more or less regarded as a continuous isotropic medium. The frequency distribution function for such a solid is evidently quadratic, owing to the constancy of the phase velocity of the waves. It is easily found from (12) and (13) that for $T \ll \theta_D$ the specific heat becomes proportional to T^3:

$$c_V = \frac{12\pi^4 Nk}{5}\left(\frac{T}{\theta_D}\right)^3. \tag{14}$$

At very low *reduced* temperatures T/θ_D this formula should thus hold quite generally in all crystals—whereas we have seen earlier that c_V becomes a constant, $3Nk$, at very high temperatures.—The Einstein model predicts an exponential decrease of c_V near the absolute zero.

A more precise analysis of the allowed lattice vibrations leads of course to much more intricate frequency distribution functions. A common feature of all distributions obtained is, that they become quadratic in the low frequency region, and contain one or more rather sharp maxima. In crystals with two kinds of atoms, two groups of frequencies exist, the *optical* branch, which is narrow and contains the higher frequencies, and the *acoustical* branch, which starts parabolically at low frequencies and approaches the optical branch at the other end. At high temperatures the narrow range of optical frequencies is excited, and the Einstein model becomes appropriate, the more so if the constituents differ more in mass. At low temperatures the Debye model is, of course, best.

For applications a modified Debye model is mostly used, in which the Debye temperature depends slightly upon the temperature of the crystal.

In insulators (e.g. purely ionic crystals at low temperatures) the conduction of heat takes place by the crystal lattice. The heating of a part of a crystal is equivalent to an increase of the vibrational energy in that part. In the case of independent harmonic oscillators, this excess energy radiates through the rest of the lattice in the form of lattice waves or *phonons* with approximately the velocity of sound. There is no dissipation, and thus an infinitely high heat conductivity. However, in practice some coupling always exists between the normal modes, and therefore some attenuation occurs. Debye introduced as a measure for this attenuation the ***mean free path*** of the lattice waves.

This mean free path should vary as $1/T$. It thus becomes very large at very low temperatures, and eventually exceeds the crystal dimensions. It then looses its meaning, and the conductivity becomes determined by the size of the crystal. The heat conductivity is proportional to the mean free path and the specific heat. As the latter is constant ($3kN$ at very high temperatures) the heat conductivity is proportional to $1/T$ in that region. At very low temperatures conductivity and specific heat follow the same temperature dependence, namely according to T^3. It would lead us too far to discuss the theory of lattice heat conduction here in any further detail.

The foregoing discussion does not apply to metals. There, the thermal and electrical conductivity as well as the specific heat and the optical properties are at low temperatures largely determined by the conduction electrons, the influence of which is proportional to T instead of to T^3. In the ionic and valence semiconducting crystals the electrons play an important part at elevated temperatures. In order to study this electronic contribution, it is necessary to investigate the electronic structure of metals and semiconductors.

4. Electronic Structure of Crystals

The behaviour of electrons in crystals can, as has been indicated before, be described as that of negatively charged particles in the periodic field formed by the ions and the other electrons (thermal vibrations neglected). As in most quantum mechanical problems there are stationary states, at various discrete levels of energy. Whereas in single atoms these levels are narrow and widely spaced, in crystals consisting of many atoms close together the levels coalesce into quasi-continuous groups, the so-called *energy bands*. Each band contains a large number of discrete levels equal to the number of unit cells in the crystal or a multiple of this. Energy E and momentum vector $\boldsymbol{k}$ of the electrons are no longer simply related as in the case of free electrons (Fig. 1*a*):

$$E = \frac{h^2}{2m} k^2, \tag{15}$$

but in the more intricate way illustrated in Fig. 1*b*.

In each band a relation of the kind (15) can be formally written down; however, the *effective mass m then depends on* $\boldsymbol{k}$. The existence of separate energy bands is a direct consequence of the modified form

of the electron wave functions, which now incorporate the periodicity of the lattice:

$$\psi_k(r) = u_k(r)e^{ik\cdot r} \tag{16}$$

in which $u_k(r)$ is a function with this periodicity. This kind of wave

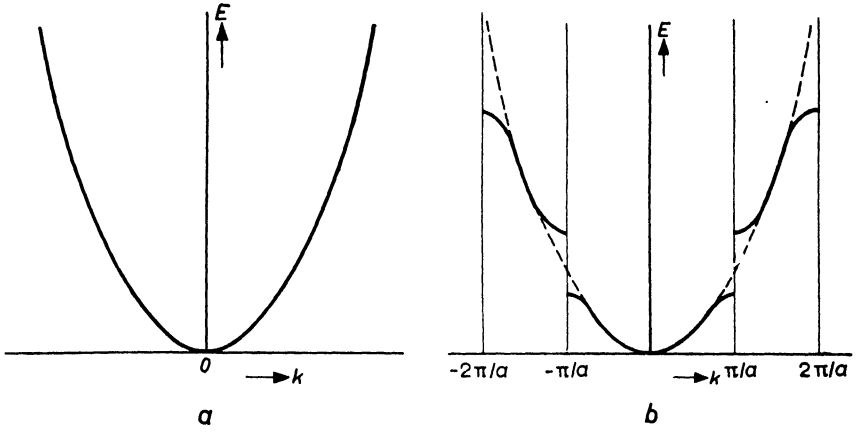


Fig. 1. *a*) Parabolic energy versus momentum relationship of free electrons: k denotes the wave number vector, the momentum vector of an electron amounts to $p = \hbar k$. The parabolic relation (15) is equivalent to the classical formula $E_{\mathrm{kin}} = p^2/2m$. *b*) Energy as a function of wave number for an electron in a periodic potential field. About the values of $k = n\pi/a$, where a denotes the lattice periodicity, discontinuities appear in the curve, which lead to allowed "bands" of energy separated by forbidden regions. Within each band, a parabolic relation of the form (15) sometimes more or less applies.

function is called a *Bloch function*. When only a slight influence of the crystal field on the electrons is present, $u_k(r)$ is practically a constant and the mass m too, and we have to do with nearly free electrons, with a quasi-continuous, parabolic energy spectrum. This case is realized in the conduction bands of the alkali metals, and to a certain approximation also in those of the noble metals. On the other hand, the electrons can be so tightly bound to the ions that $u_k(r)$ is nearly zero everywhere except close to the ion core. In that case the energy bands are very narrow, and correspond approximately to the discrete quantum states of the atom: m is a tensor quantity that varies strongly with k.

When a force, for instance caused by an electric field, acts on an electron, this tends to be accelerated, that is, it tends to increase its momentum and kinetic energy. To do this, it must occupy progressively higher energy levels. This is only possible if these levels are still empty.

Because of Pauli's principle, only two electrons can occupy the same state, that is, the same energy level. In the cubic monovalent metals the conduction band contains as many levels as there are conduction electrons and the band is therefore half filled (Fig. 2*a*). Acceleration will

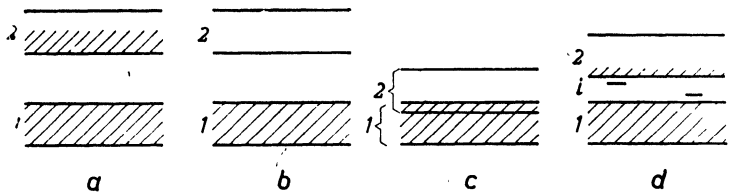


Fig. 2. Energy band schemes for: *a*) cubic monovalent metals; *b*) insulators; *c*) an alkaline earth metal; *d*) a semiconductor with impurity levels *i*. Two bands have been shown in each case, indicated by 1 and 2 respectively.

thus be very well possible, and it can be shown that in a perfectly periodic lattice the electrons will move in the direction of the applied force without any resistance at all. In actual crystals, where the structure is less perfect, some resistance is always present, but nevertheless these metals are always good electronic conductors. It may happen, however, in other materials, that the highest occupied levels just fill up completely an energy band (Fig. 2*b*). Then no acceleration is possible, except in extremely strong fields, in which the electrons jump to a higher, yet unfilled band. This is the case in insulators. It is also more or less the case in the alkaline earth metals; however, in these a higher band (the p-band) overlaps partly with the full s-band (Fig. 2*c*). These metals are, therefore, still good conductors. It may happen that, under the influence of the heat motion, electrons in a full band can jump relatively easily out of it into the next higher empty band. There they can of course be accelerated, and the solid has become an (intrinsic) semiconductor. The promoted electrons have left empty levels in the full band. These act as positive charge carriers, commonly named *holes*, and contribute also to the conductivity, in a manner quite analogous to that of the electrons.

The incorporation of foreign atoms in the crystal can also cause the presence of mobile electrons or holes, irregardless of temperature. Then one speaks of impurity semiconductors. Semiconducting properties of a material are usually encountered only when the width of the forbidden band gap is not very large, or when so-called "impurity levels"

exist within this gap that are not too far distant from one of the band edges (Fig. 2*d*).

It is plausible that the proportionality constant between force and acceleration for a charge carrier *in a crystal field* need not be equal to its mass, as is the case for free carriers. In many instances, e.g. in the monovalent metals, this is nevertheless approximately so; the conduction electrons are characterized by an "effective mass" that differs but slightly from the electron mass. In solids where a tighter coupling between electrons or holes and the lattice exists, and especially in solids with directed bonds (homopolar crystals), the "effective mass approximation" becomes less applicable and one has to resort to the detailed quantum mechanics of the system in order to describe it adequately. It would lead us too far to discuss this here.

In metals with their partly filled energy bands (single or composite, as the case may be), the equilibrium energy distribution of the conduction electrons is essentially governed by the *Pauli principle* (*degeneracy*); the resulting *Fermi-Dirac statistics* are characterized by completely full energy levels below a certain energy E_F, the *Fermi energy*, and completely empty levels above it. Only a narrow region of

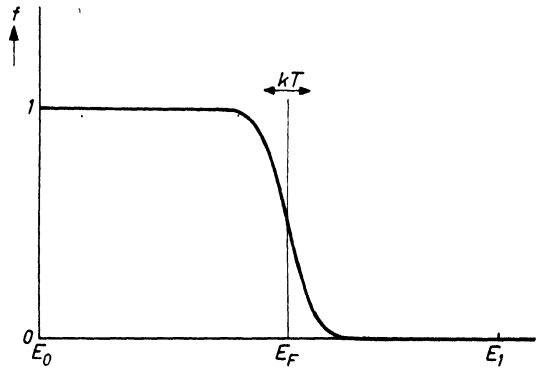


Fig. 3. Fermi-Dirac distribution function of electrons in the conduction band of a metal. The band extends from E_0 to E_1; at the Fermi energy E_F there is a fairly sharp transition, of width about kT, between filled and empty levels indicated by fraction number $f = 1$, and $f = 0$, respectively.

width of the order kT forms the transition between full and empty band halves (Fig. 3). The electrons occupying levels in the transition region can be accelerated, but those of lower energy cannot, owing to

the full degree of occupancy of the levels above their own. Therefore, only a fraction of the order of kT/E_F of the conduction electrons takes part in the transport of current through a metal. It is also only these electrons that can take up energy on heating, thus that contribute to the electronic specific heat. Usually E_F is of the order of a few electron volts. As kT amounts but to 0.1 eV at 1000 °K, at the utmost only a few per cent of the total number of electrons are available for these purposes.

In homopolar semiconducting materials containing a few mobile electrons and holes only, the distribution in energy of the carriers is no longer degenerate and *in each separate band* ordinary Boltzmann statistics apply.

Binary ionic compounds usually contain positive (metal) ions with closed shells, thus with firmly bound electrons. The negative ions possess one or more loosely bound electrons, however, which can jump from one negative ion to another. In a perfect ionic crystal these electrons just fill one complete energy band; this corresponds to all

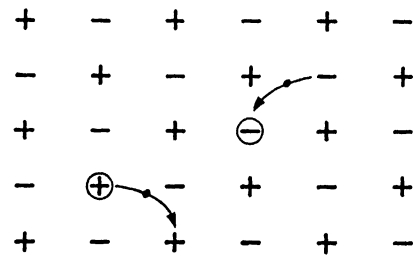


Fig. 4. Motion of electrons and holes in an ionic compound. Positive and negative ions are indicated by + and — respectively. A positive ion having acquired an electron is indicated by ⊕; the electron (dot) migrates by jumping to another positive ion (arrow). A once ionized negative ion ⊖ can be considered as the site of a hole. This migrates to a neighbouring negative ion when an electron from the latter jumps to the former.

ions having their appropriate charge. If, for some reason, an electron from this band is promoted to a higher band, both it and the corresponding *hole* are mobile. The motion of such a hole can be visualized as the jumping of an electron to a once ionized "negative" ion (actually neutral in a monovalent compound) from a neighbouring fully charged negative ion, leaving this ionized instead. The electrons travel from

positive ion to positive ion (Fig. 4). Therefore, the originally full band in such a crystal is often called the anion (chloride, sulphur, etc.) band, whereas the conduction band is then named the metal (silver, zinc) band.

We have already mentioned the electronic specific heat. Contributions to this only come from electrons that are able to increase their kinetic energy. In general, even in metals the electronic contribution to the specific heat is therefore but relatively small, except at very low temperatures, where the lattice contribution has decreased even more. The latter decrease, proportional to T^3 according to section 3, is much faster than that of the electronic specific heat, which (of course) only varies proportionally to T. In other materials the electrons contribute still less to the specific heat at normal temperatures.

In most solids the mobile electrons are of great importance for the transport phenomena; this is especially so in metals. One can show that in a perfectly periodic lattice, when a force F acts upon an electron, the expectation value of its "crystal momentum" $p = \hbar k$ changes according to

$$\frac{\mathrm{d}}{\mathrm{d}t}(\hbar k) = F\,, \tag{17}$$

exactly as in classical mechanics. In the absence of perturbations, the electron is thus simply accelerated. However, there exists a coupling with the thermal lattice vibrations (phonons): the electrons make *collisions* with the lattice during which vibration modes are excited. Then the energy and momentum of electrons and phonons *together* are conserved. Owing to the relative smallness of the energy of phonons of given momentum compared to the energy of an electron with the same momentum, the electrons lose but little energy in each collision, but the direction of their motion is greatly changed. This kind of interaction between electrons and phonons leads to a steady state, and thus to a limited steady current. In other words, a finite electrical resistivity is caused by the thermal vibrations: this is named the *thermal (contribution to) resistivity*. The steadying effect of collisions can in many cases be approximately described by the assumption of the existence of a *relaxation time*, τ_e. This denotes the mean free time elapsing between two collisions of the same electron. It enters in the expressions for the conductivity in the form:

$$\sigma = \frac{ne\tau_e}{c}, \tag{18}$$

where n is now the number of electrons per unit volume. Increasing the temperature leaves n approximately constant, but decreases τ_e, and thereby σ.

In semiconducting solids one can in principle follow the same line of thought, except for the fact that the statistics of the carrier distribution is wholly different. One still can formally write

$$\sigma = ne\mu ,$$

but μ, the *mobility*, that takes the part of τ_e/c according to an established custom in the description of non-metals, now usually depends in an intricate way on the temperature, impurity content, etc., as does n. In thermally activated (intrinsic) semiconductors n will undoubtedly increase with temperature, but the variation of μ depends on the band structure. In general, however, the semiconducting substances become more highly conducting as the temperature increases.

Analogous considerations apply to the electronic conduction of heat. Again a relaxation time can often be defined, which may, however, differ slightly from the relaxation time for electrical conduction. Often, instead of the relaxation time, or together with it, the concept of the *mean free path* of electrons is introduced. This is in the simplest case related to τ_e according to

$$l = v\tau_e ,$$

where v is the (thermal) *drift* velocity of the electrons. For a very detailed discussion of these matters we refer to a paper by SONDHEIMER [1952].

It will be clear that the electronic properties of solids are governed to a large extent by disturbances of the periodic structure, whether they be lattice vibrations or more localized defects. The same is true for the properties of the lattice as a whole. We shall therefore now proceed to the detailed discussion of these defects.

GENERAL REFERENCES AND TEXTBOOKS

BARRETT, C. S. (1952); Structure of metals (McGraw-Hill; New York).
GUGGENHEIM, E. A. (1955); Thermodynamics (North-Holland Publ. Comp.; Amsterdam).
MOTT, N. F. and R. W. GURNEY (1948); Electronic processes in ionic crystals (Oxford Univ. Press, Oxford, second impression).
MOTT, N. F. and H. JONES (1936); The theory of the properties of metals and alloys (Oxford Univ. Press; Oxford, Dover reprint 1958).
PEIERLS, R. E. (1955); Quantum theory of solids (Oxford Univ. Press; Oxford).
SEITZ, F. (1940); Modern theory of solids (McGraw-Hill; New York).
VOIGT, W. (1910); Lehrbuch der Kristallphysik (B. G. Teubner; Leipzig).
WILSON, A. H. (1953); The theory of metals (The University Press; Cambridge).
ZIMAN, J. M. (1960); Electrons and Phonons (The University Press; Cambridge).

SPECIFIC REFERENCES

BORN, M. (1951); Festschrift Göttingen Akad. Wiss., I, Math.-Phys. Klasse **I, 1**, and: Nachr. Akad. Wiss. Göttingen, Math.-Phys. Klasse **IIa**, No. 6.
BORN, M. and KUN HUANG (1954); Dynamical theory of crystal lattices (The Clarendon Press; Oxford).
GRÜNEISEN, E. (1908); Ann. der Physik **26**, 393.
HUNTINGTON, W. (1958); Solid State Physics **6**, 1.
LEIBFRIED, G. (1955); Encyclopedia of Physics, Vol VII, part 1, 104 (Springer; Berlin).
SMITH, C. S. (1958); Solid State Physics **6**, 175.
SONDHEIMER, E. M. (1952); Advances in Physics **1**, 1.

CHAPTER II

POINT DEFECTS IN CRYSTALS

1. Geometrical Considerations

The kind of imperfection of a crystal that is most apt to occur is the lattice site where an ion or atom is missing. This is called a *vacancy*. In metals the removal of an ion necessarily entails the removal of a conduction electron too, for the crystal as a whole must remain neutral. This neutrality condition is very stringent and is therefore (with some exception, see below) rigorously fulfilled. In pure ionic crystals the condition, in the absence of other types of defects, necessitates the presence of as many positive- as negative-ion vacancies; such a pair of vacancies of opposite sign has even obtained a name: *Schottky defect*. Often, however, the charge equilibrium is largely governed by the simultaneous presence of other kinds of lattice defects, e.g. divalent foreign ions in material containing normally monovalent ions. In the valence crystals electrons (holes) as well as foreign atoms contribute to the retention of neutrality. In the following we shall always assume the crystal to be neutral.

A defect which is more or less the counterpart of the vacancy is the *interstitial atom* or (short) *interstitial*. This defect forms when an atom or ion shifts to a position in the crystal which is not a lattice site. Again, the simultaneous presence of an extra electron, of a vacancy of an equally charged ion, or of a foreign ion is needed to ensure neutrality. Whereas vacancies, in principle, can occur in all crystal structures, the formation of interstitials is more easy in structures containing large interatomic voids than in close-packed structures where there is, properly speaking, no room for them. The diamond structure and the NaCl structure are examples of the first category, close-packed metal structures of the latter. It is therefore to be expected that interstitials in face-centred cubic metals will be a rare occurrence under normal

conditions. This is borne out by experiment. Actually, in all well-annealed metals (except possibly in the alkali metals) the only kind of interstitial atoms that can occur in appreciable numbers at normal temperatures are small foreign atoms, such as protons or boron, carbon and nitrogen atoms. In f.c.c. lattices these occupy the very small socalled tetrahedral positions of the type $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ or the somewhat larger octahedral positions $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ (Fig. 1). In the b.c.c. lattice

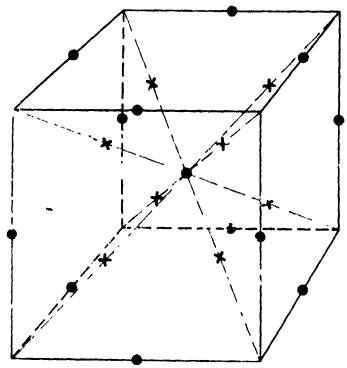


Fig. 1. Interstitial positions in the face-centered cubic lattice. Tetrahedral positions have been denoted by crosses, octahedral positions by black dots.

there are interstitial sites available of tetrahedral symmetry at the places $(\frac{1}{2}, \frac{1}{4}, 0)$ and somewhat smaller positions at $(\frac{1}{2}, \frac{1}{2}, 0)$ and the equivalent sites $(\frac{1}{2}, 0, 0)$ (Fig. 2). On the average the holes in the body-centered cubic lattice are relatively smaller than those in the close-packed cubic lattice with the same parameter; however, the former structure contains a much larger number of suitable interstitial sites (9 per atom) than the latter (3 per atom).

In the diamond structure, regarded from a purely geometrical point of view, there are but $\frac{2}{3}$ of the available lattice sites occupied (Fig. 3). The lattice could therefore in principle accommodate without geometrical difficulty not less than one interstitial atom (in a tetrahedral position) for each pair of normal atoms. These preferred positions are the sites corresponding to both types of sites in the f.c.c. lattice, but now they are all tetrahedrally surrounded. Analogous holds true for the wurtzite and zincblende structure and related structures (although the

symmetry is lower). Of course, the special bonding character prevents such a large number of interstitials, but a transient occupation of

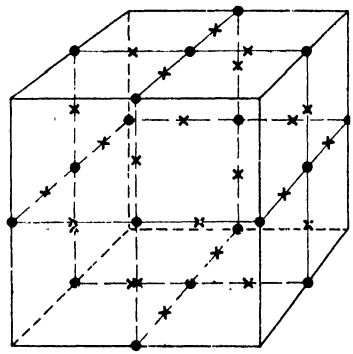


Fig. 2. Interstitial positions in the body-centred cubic lattice. "Large" sites have been indicated by crosses, "small" sites by dots.

many interstitial sites by the constituting atoms seems here quite well possible.

The ionic lattices occupy an intermediate position with regard to the

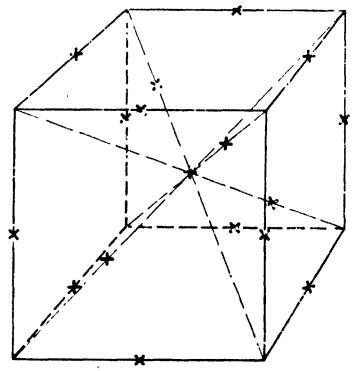


Fig. 3. Interstitial positions in the diamond lattice.

accommodation of interstitials. Whereas the geometry of the lattice usually leaves some place for them, on the other hand the constituting ions often differ so much in size that the packing is nevertheless very

effective. It varies therefore greatly between the various materials, whether considerable fractions of interstitials shall occur in them or not.

These geometrical arguments can of course only be considered as very approximate. They need to be replaced by quantitative energetic considerations in order to make numerical estimates of the expected concentrations possible.

In the first place one must consider the elastic stresses and strains set up in the medium, regarded as continuous, when a void is made in it or when material is added to it somewhere. These normally only contribute little to the energy of formation of a point defect. A detailed discussion of the elastic stresses around defects is presented in Chapter IV. Much more important is the energy associated with the disturbance of periodicity in the atomic structure of the solid, and with the disturbance of the typical binding forces holding the particles together. We shall deal with these "chemical" contributions in the next section. Once the energy of formation of a point defect is known, its thermal equilibrium concentration in a crystal can be found by applying the well-known methods of statistical mechanics.

2. Energy of Formation of Point Defects in Crystals

2.1. SIMPLE DEFECTS

Suppose an atom in a homopolar crystal, an ion in a metal, or a pair of ions of opposite sign in a polar material, is removed from the interior and brought to the surface. The lattice energy associated with the defect thus formed in the interior is determined by the various factors discussed in Chapter I. One usually finds this energy to be about 10 eV per ion pair in ordinary polar crystals, and one to a few eV per ion in metals or semiconductors. In order to remove the ion with its electron, or the ion-pair, and bring them to the surface, this energy has to be provided in the first instance. However, the lattice will relax around the defect thus formed and then part of the energy is restituted.

The contribution of the purely elastic relaxation in the case of a vacancy is nearly always very small, however. The elastic strains can in first approximation be found by considering the vacancy as a spherical void in a continuous medium (see Chapter IV), and are thus inversely proportional to R^3, where R is the radial distance. Thus only the nearest neighbours are appreciably affected. As the interionic forces themselves also rapidly decrease with increasing distance, then if the purely

elastic part of the relaxation is only very small, so is the contribution to the change of lattice energy. It can be shown that in none of the crystals that will be considered do the displacements of the ions surrounding a vacancy exceed a few percent of the atomic distance. The displacements may be either inwards or outwards, depending on the relative variation with the distance of the attractive and the repulsive force components respectively. When the atoms or ions can be regarded as hard spheres that are just touching each other (a case realized approximately in the noble metals), the average displacements are very small and inwards, whereas in a material in which the ions are held at a distance by electrostatic forces (as in purely polar crystals) the average displacements are usually in the radially outward direction. In both cases the elastic part of the energy of formation of the defect does normally not exceed a few tenths of an electronvolt.

For an interstitial atom the situation is different. For purely geometrical reasons the distortion of the lattice caused by the insertion of an interstitial is generally considerable. The displacements of the surrounding ions may even amount to 20% of the lattice parameter, and the corresponding strain energy to a few eV.

Another, very important contribution to the energy of both vacancies and of interstitials arises from the redistribution of charge around the defect. In polar crystals, in the neighbourhood of every vacancy the lattice becomes polarized; the nearest neighbours (all of one sign) are displaced through a greater distance than the next-nearest neighbours (of opposite sign) which may even move the opposite way. Owing to this difference, in the vicinity of the defect an electric field develops. Therefore, the energy to produce the defect is *decreased* by the amount $e\Phi$, where e is the absolute value of the charge on the ion removed and Φ the potential developed. Similar considerations hold in the case of an interstitial. In a metal the situation is far more complex, since now the effect of the whole electron gas needs to be considered. To illustrate this we shall (very briefly) outline one of the recent theoretical determinations of the energy of a vacancy in a monovalent metal, as carried out by FUMI [1955].

First to be computed is the change of the energy of the electron gas on the removal of one positive ion, together with its electron. The vacancy acts as a (screened) negative charge, and the conduction electrons are somewhat repulsed by it. The corresponding increase of energy is directly related to the Fermi energy of the electrons. Second-

ly, the effect on the electron gas of the change of density, due to the removal of an ion is considered; it is easily shown that this results in a decrease of the kinetic energy of the electron by an amount equal to $\frac{2}{5}$ of the Fermi energy. Thirdly, the effect of the ion-ion repulsion is taken into account. This is what we have previously called the elastic part of the energy and is thus not very important, especially not in the alkali metals where there is no repulsion to speak of. It is of some account in the noble metals where there is closed shell repulsion between ions. This contribution never amounts to more than a few tenths of an eV in the case of a vacancy and is then negative. It may amount to more than one eV for an interstitial in copper, however. By applying the formalism of Born's lattice theory and thus considering energy changes due to electronic and to ionic forces *simultaneously*, BENNEMAN and TEWORDT [1960] and SEEGER and MANN [1960] could increase substantially the reliability and accuracy of the energy calculations in the noble metals. The remaining uncertainty does not exceed more than 0.1 or 0.2 eV. The energy of formation of a vacancy in Cu was found as 1.0 eV, that of an interstitial (in interstitialcy formation, see below) as 2.7 eV. The volume expansion associated with the Frenkel-pair is 1.3 atomic volumes. Further discussion on this subject is postponed to Chapter XV.

For semiconducting materials there exist no reliable theoretical estimates for the energy of formation of point defects. Here the energy per valence bond plays a decisive role, and this is not known with accuracy at present. A comparison with other solids according to the known energies of sublimation, is unallowed. Only for vacancies in ionic crystals some correlation between formation energy and heat of sublimation does exist; in metals the Fermi energy should rather be considered and in semiconductors the situation probably lies in the middle.

2.2. ASSOCIATION OF POINT DEFECTS; CROWDIONS AND INTERSTITIALCIES

Point defects in crystals may associate into pairs of larger clusters, by reason of the effective charges they carry, or because the elastic strains around them favour the association. In most metals, the electrostatic interaction is unimportant, owing to the efficient screening of the charges by the electron gas, which makes the interaction poten-

tials depend exponentially upon the distance, with a mean range of less than one ionic distance. Also the purely elastic part of the interaction energy is small. There is only the circumstance that the necessary energy to remove an ion from a site adjacent to a vacant site is smaller by about $1/p$ of the binding energy per ion than that needed to remove it from an arbitrary site, where p is the number of nearest neighbours. Therefore, a positive energy is required to cause the dissociation of a pair of adjacent vacancies into two separate vacancies. This dissociation energy was estimated in copper to be about 0.3 eV.

In polar crystals there is no screening of effective charges, and pairs of vacancies of opposite sign may easily combine to form neutral defects, which will also be called vacancy-pairs. For NaCl the dissociation energy was, long ago, found theoretically to be about 0.9 eV (MOTT and GURNEY [1948]), but this is probably too large a value. About interstitial-vacancy interaction, and the mutual interaction between interstitials that undoubtedly exists in crystals with appreciable polar character not much quantitative is known at present. Very important is the interaction between point defects and impurity ions (Chapter XXIV). In the semiconducting solids germanium and silicon vacancy pairs seem to be a common occurrence, at least in deformed materials. The energy of association amounts probably to about 1.3 eV in germanium (LETAW [1956]).

Many details on the fundamental properties of point defects in solids are contained in the Handbook contribution by SEEGER [1956], to which the interested reader is referred.

It was suggested by PANETH [1950] that, at least in the body-centred cubic alkali metals, a possible configuration for an interstitial ion would be the *crowdion*. This is a configuration in which, along a ⟨111⟩ direction, over a distance of the usual number (say 5) atoms one additional atom is present. The crowdion can hence be regarded as an interstitial atom which has been "extended" along a ⟨111⟩ direction. Such a crowdion is expected to represent less energy than a concentrated interstitial. On the other hand, its mobility would be rigorously confirmed to its length direction. Owing to the smaller stresses around it, a crowdion has less effect on the physical properties (resistivity) of the material than an interstitial. TEWORDT [1958] has computed the energy of formation of such a crowdion along a ⟨110⟩ direction in face-centred cubic copper, and found it to be *larger* by about 25% than that of an ordinary interstitial, which contradicts the

original suppositions, made for a b.c.c. metal. However, apart from the normal interstitial position the *interstitialcy* position should be considered, as was pointed out by VINEYARD [1960] on the basis of his machine calculations (see Chapter XV), and simultaneously by SEEGER *et al.* [1960]. This socalled dumb-bell or split-interstitial configuration, as it is also called, is formed when a normal interstitial atom on an octrahedral site and its neighbouring atom in the cube direction both shift somewhat in the direction from the former to the latter, so that two "semi-interstitials" are formed, which are symmetrically placed with respect to a normal position and which give rise to a tetragonal deformation pattern (Fig. 4).

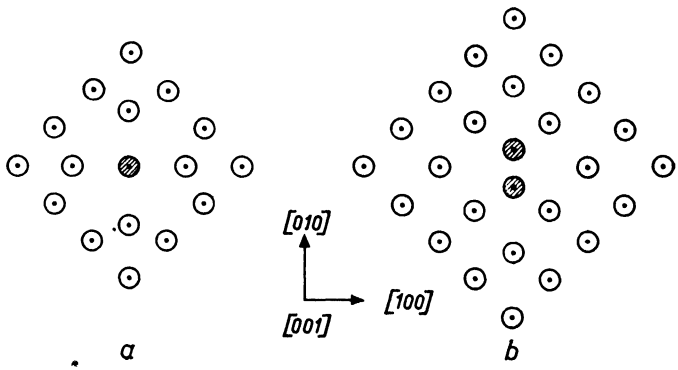


Fig. 4. *a*) Interstitial in the face-centered cubic lattice, normal position; *b*) Idem, in the interstitialcy or dumb-bell position, with tetragonal symmetry along a cube direction.

The combination of defects into clusters of more than two is a very interesting problem, but so complex that numerical calculations are only possible by the aid of electronic computing machines, after assuming some sort of approximation to the interatomic forces. Dienes and Vineyard, in studying the possible stability of a planar configuration of three adjacent vacancies in a face-centred cubic lattice, found such an arrangement not to represent an equilibrium configuration. Instead, they could show a definite preference for a cluster consisting of *four vacancies* in tetrahedral positions, with an *interstitial atom* in the middle (thus again equivalent to a *trivacancy*). In the approximation used the binding energy of this defect would be no less than 2.9 eV in copper. The unstable planar trivacancy transforms into this tetrahedral cluster very easily; the gain in free energy by the relaxation of the atom

on top to the interstitial position somewhat lower is 2.3 eV, and makes up for a large part of the binding energy.

To move such a cluster, what must happen is that the central interstitial is pushed back to its original lattice site, driving a new atom into the interstitial position so vacated. Then the whole tetrahedron is shifted one atomic distance. Because of the first process, the energy of migration should be very high (1.9 eV in copper) and consequently the cluster should be practically immobile. These clusters can, and will act as nuclei for larger *voids* at high temperatures.

About the equilibrium of clusters of two or more *interstitials* less quantitative computations exist. Various authors refer to the possible existence of *di-interstitials*, which should be stable at least at very low temperatures (see also Chapter XV). The situation with regard to *interstitial-vacancy pairs* is the same: experimental evidence, not unambiguous, exists to their occurrence, but computations of interaction energies are scarce and inaccurate. Huntington has considered various configurations of a vacancy and an interstitial in the f.c.c. lattice. Numbering the various positions by the square root of the odd number which expresses the distance in units of half a lattice distance (that this is possible was first pointed out by J. S. Koehler), only the number 7 does not correspond to a possible configuration. Huntington has showed the positions $\sqrt{1}$ and $\sqrt{5}$ to be unstable, but $\sqrt{3}$ might very well be stable (connection line along ⟨111⟩).

3. Point Defects in Thermal Equilibrium

The presence of point defects increases the configurational entropy of a crystal to a large measure; therefore, although their introduction requires a considerable positive amount of energy to be supplied, at elevated temperatures the *free energy* of the crystal may be reduced. Considering mutually non-interacting defects, with an energy of formation E, and taking account only of the configurational entropy, we arrive at the following estimate of this change of free energy. In a crystal containing N ionic sites, the removal of n ions can be done in $N!/(N-n)!n!$ distinct ways. The increase of configurational entropy is therefore:

$$\Delta S = k \log \frac{N!}{(N-n)!n!} . \tag{1}$$

In metals and homopolar crystals this formula applies to the formation of vacancies as well as to the generation of interstitials; in polar crystals, however, to conserve charge neutrality, the vacancies must occur either in pairs of opposite sign (so-called *Schottky defects*), or together with an equal number of interstitials of the same kind of ions (these combinations are called *Frenkel defects*). If these pairs are *associated,* that is, if the constituents form a unit that belongs together, formula (1) can be taken over without change (except for a possible numerical factor of order unity, arising from a certain freedom in orientation), when we read for N the number of ion *pairs* (of opposite sign), and for n the number of associated Schottky or Frenkel defects. In the case of *free* defect pairs, the two components of which behave independently, the discussion must be slightly changed. Let N as before be the number of available sites for vacancies of one kind, and let N' be the number of available sites for vacancies of the other kind (Schottky defects) or for interstitial ions (Frenkel defects) respectively. Then the n defect pairs can be arranged in

$$\frac{N!}{(N-n)!n!} \times \frac{N'!}{(N'-n)!n!}$$

ways. Compared to (1), there now exists a second contribution to the entropy:

$$\Delta' S = k \log \frac{N'!}{(N'-n)!n!} . \tag{1a}$$

The increase of internal energy is nE, where E stands either for the energy of one vacancy or one interstitial in a metal, or for the energy of a vacancy pair, or a vacancy-interstitial *combination* in a polar crystal. The change of free energy is in the first case and in the second case when the defects are associated:

$$\Delta F = nE - T\Delta S . \tag{2}$$

If in the second case the defects are free, one has:

$$\Delta F = nE - T\Delta S - T\Delta' S . \tag{2a}$$

By inserting (1) or (1a) and minimizing ΔF with respect to n, while applying Stirling's approximation for the factorial of a large number, it

is easily found that the thermal equilibrium number of either vacancies or interstitials in a metal, or of associated defects in a polar crystal, is given by:

$$n_{eq} = N e^{-E/kT} . \tag{3}$$

For free defects pairs the corresponding expression becomes:

$$n_{eq} = \sqrt{NN'}\, e^{-E/2kT} . \tag{3a}$$

These formulae have been derived under the tacit assumption that $n \ll N$ or N'.

The equilibrium concentrations of defect associations of a more general nature than the Schottky and Frenkel defects, such as di-vacancies in metals, can be found immediately from the foregoing expressions when only the energy of association of the defects is known. Let us consider the case just mentioned, viz. that of divacancies in metals. When the energy of association is denoted by U, that is, when the energy associated with the presence of two non-interacting separate vacancies is lessened by the amount U when these vacancies associate according to the reaction:

$$2v \rightleftharpoons d ,$$

then the concentration c_d of divacancies is, in equilibrium, related to that of vacancies c_v, by:

$$c_d = \alpha c_v^2 e^{U/kT} , \tag{4}$$

where α is a numerical orientation factor.

From this it follows that the equilibrium number of divacancies is given by

$$m_{eq} = \alpha n_{eq}^2 e^{U/kT} = \alpha N^2 e^{-(2E-U)/kT} . \tag{5}$$

Here α takes account of the various possible spatial orientations that a given vacancy pair can take; $\alpha = 6$ in f.c.c. and monatomic h.c.p. crystals, 4 in b.c.c. crystals; 6 in the sodium chloride and 8 in the caesium chloride structure. If, in ionic crystals, association occurs with foreign ions, the number of associated defects is, of course, no longer governed solely by the thermodynamic considerations described above, but primarily by the number of impurity ions present.

The above derivations of equilibrium concentrations of point defects are based on the assumption that the presence of these defects only affects the crystal entropy as far as the distribution over the available

lattice sites is concerned (entropy of mixing). However, near a point defect the vibrational frequencies of the atoms or ions are slightly changed, and moreover, on changing the temperature of the crystal, this expands and the activation energy of formation of defects will vary correspondingly. Further, the latter energy may depend on the temperature anyway (HAVEN [1955]). In the Einstein model of the vibrating lattice, every particle vibrates with the same frequency ν_E, and as can be found in any textbook, the part of the entropy of the crystal that is due to the vibrations is approximately given by

$$S_{vibr} = 3Nk \log \frac{h\nu_E}{kT}$$

(N being now the number of atoms or ions).

Suppose that x neighbouring atoms change their frequency to ν', owing to the introduction of a vacancy. This will only apply to vibrations along the direction between atom and vacancy, and therefore, to this frequency change there corresponds a variation of entropy:

$$\Delta S_{vibr} = xk \log \frac{\nu'}{\nu_E}. \tag{6}$$

In formulae (1) and (1a) this term must be added on the right hand side. This implies the introduction of a *pre-exponential entropy factor* γ in formulae (3), (4) and (5) of order of magnitude $(\nu'/\nu_E)^x$.

The influence of the thermal expansion causes the introduction of another, in general temperature dependent pre-exponential factor, which we shall call B, following MOTT and GURNEY [1948]. VINEYARD and DIENES [1954] and HAVEN and VAN SANTEN [1952] have pointed out that, by computing the entropy change due to the vibrational modes of a crystal properly from the point of view of the Born- von Karman lattice theory, the effects of lattice expansion and temperature dependence of the energy of formation have automatically been considered too. The factor B then becomes exactly equal to unity.

In most crystals x is at least 6 or 8. In polar crystals the change of vibrational frequency may be very great, and even amount to a factor of about 2, because the electrostatic binding forces are very seriously disturbed near a defect. MOTT and GURNEY [1948] estimated the pre exponential factor γ for Schottky pairs in alkali halides to lie between 10 and 10^4. When interstitials and vacancies are simultaneously formed (Frenkel defects) the factor is smaller by several orders of magnitude,

because the change of the vibrational frequency of the ions adjacent to an interstitial has opposite sign to that of the ions adjacent to a vacancy. In metals, and probably in homopolar crystals too, the changes of the vibrational frequency caused by the introduction of a defect are very small, of the order of a few per cent, especially where vacancies are concerned with their small elastic strains. Therefore the factor γ is of the order of unity for vacancies and does not exceed 10 for interstitials in a monovalent metal. As normally interstitials are not very important thermal defects, we can for practical purposes, neglect the presence of an entropy factor altogether in the case of the noble metals.

On raising the temperature of a crystal, the concentrations of the vacancies and interstitials that occur in thermal equilibrium increase. That is, point defects are provided from certain *sources*. Reversely, on decreasing the temperature, part of the defects disappear into *sinks*. The nature of these sources and sinks has only recently been clarified. It appears that other lattice imperfections, namely *dislocations*, play an important role in this respect. We shall return to this in the next chapter.

4. Formation of Vacancies and Interstitials in Non-Equilibrium Concentrations

Although the thermal equilibrium concentrations of point defects can only reach appreciable values when the temperature is not much below the melting point, often in practice concentrations are encountered at low or moderate temperatures that are higher by many orders of magnitude than the corresponding equilibrium values. This excess can be due to several causes.

4.1. EFFECT OF IMPURITIES

Even the most carefully prepared rocksalt crystals usually contain still 0.01 to 0.1% Ca^{2+}-ions, and these readily associate with positive ion vacancies: the energy of association is 0.3 eV (ETZEL and MAURER [1950]). It is therefore not strange that at moderate temperatures the number of vacancies present in such crystals is determined rather by the number of Ca-ions than by the thermal equilibrium value. At room temperature about 10% of these ions is expected to be associated with vacancies (SEITZ [1954]).

In non-polar crystals also some affinity might be expected to exist between foreign atoms and point defects; this follows for instance from certain observed influences of vacancies on the diffusive properties of impurities (Chapter XVIII). In compounds of non-stoichiometric composition the occurrence of point defects is common, whether they are of an ionic nature or not (Chapter XXV).

4.2. EFFECTS OF PHYSICAL TREATMENT

Point defects can be introduced in large numbers into solids by plastic deformation and by high-energy irradiation. Non-equilibrium concentrations may also be frozen into the solid by rapid quenching from a high temperature. Quantitative considerations as to the degree of freeze-in by quenching have been published for the case of vacancies in the noble metals (SEITZ, KOEHLER and BAUERLE [1957]). In these metals, owing to their relatively high activation energy of formation other defects than vacancies do not occur in appreciable amounts in thermal equilibrium. The quenching process is only effective when the quenching velocity is large enough to prevent the establishment of temporary equilibrium at the intermediate temperature during the quench. The rate of approach of a new equilibrium state on a change of temperature is governed by the diffusivity at that temperature. We shall return to this question in Chapter XIV.

The effect of cold-work will be bypassed in this Chapter as it will be treated in extenso in Chapters VII, XIV and XV.

A fast particle impinging on a crystal loses speed owing to the elastic and inelastic collisions it makes with the crystal particles. A particle with mass M and kinetic energy T in an elastic collision with a crystal particle, of mass m, transfers an amount of kinetic energy equal to:

$$\Delta T = -\frac{4Mm}{(M+m)^2} T \sin^2 \zeta/2 , \tag{7}$$

where ζ is the angle of deviation (in the centre of gravity system of coordinates). The number of collisions per unit length, or the *effective collision cross-section*, also depends on T and of course on the nature of the particle and the crystal. For moderate values of the kinetic energy and mass of the particles the energy transferred may easily be large enough to knock the crystal particle from its place. Then a vacancy is created together with an interstitial, i.e. a Frenkel defect.

In general $M \ll m$, and thus

$$\Delta T \approx -\frac{M}{m} T . \tag{8}$$

It is therefore advantageous to irradiate with heavy particles to obtain as large as possible an energy transfer per collision and thus a large number of defects. Electron irradiation will obviously be rather ineffective (the reason for this being that even a high-energy electron carries only a relatively small momentum).

Another important factor that determines the action of the radiation on the crystal is the *charge* of the impinging particles. The collision cross-section is much smaller for uncharged particles (i.e. neutrons) than for charged ones, but the former are much more effective in producing displaced atoms. This difference is largely due to the fact that charged particles easily *ionize* the ions or atoms of the crystal and thereby loose a large fraction of their energy without knocking any atoms from their site. Hence, and also for more sophisticated reasons which we shall not discuss here, the charge causes an increasing tendency for *in*elastic collisions. Charged particles of moderate energy do not penetrate very far into the material.

It would be very efficient, according to (8), to produce structural defects by irradiation with heavy atomic nuclei. This case can be realized when the material to be treated is fissionable. Then in the interior of the solid heavy fission fragments can be formed (for instance by bombarding with neutrons), that possess very large kinetic energy.

For the formation of Frenkel defects in non-fissionable materials neutron irradiation remains the most effective method. Practically homogeneous distributions of the defects in not too thick layers of material can be obtained in this way. Very energetic electrons, accelerated by voltages of the order of a megavolt or higher, can in practice also produce a homogeneous defect distribution, but only in rather thin samples. Low-energy electrons cause only ionization, and heavier charged particles do not penetrate appreciably into the material.

In principle the transferred energy of an impinging particle can either be used in the production of further Frenkel defects, or it can cause local heating in the crystal; in this case the energy of the collision is shared by a group of atoms, which do not leave their sites but only temporarily vibrate much more intensely. This local heating is usually denoted by the term *thermal spike*. It can have several conse-

quences, depending on the nature of the crystal and the particles, and other imperfections apart from Frenkel defects may be generated in it. Neutrons cause relatively many thermal spike regions, electrons are presumably not active in this respect. To obtain a *well-determined* state

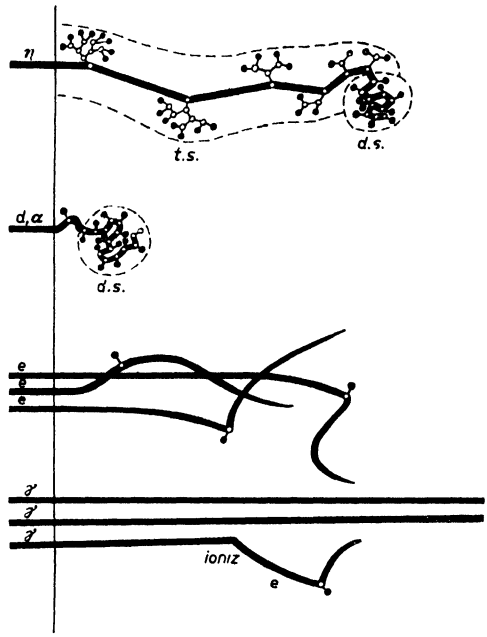


Fig. 5. Schematic comparison between the results of different sorts of fast-particle irradiation of solids. Neutrons (n) produce localized clusters of many (about 100) Frenkel defects, together with thermal spikes of great relative importance for the total damage. Charged particles such as deuterons (d) or alpha particles (α) cause less damage than neutrons but equally localized and concentrated in spikes, which are, because of the more difficult penetration, less homogeneously distributed through the material. Fast electrons (e) only produce one Frenkel defect each, at least if they do not possess too much kinetic energy, and these defects are randomly distributed. γ-rays occasionally cause ionization, thereby releasing a fast electron, which then in turn effects one displacement. The mean penetration is very large.

of imperfection, electron irradiation is therefore to be preferred above neutron irradiation.

An indirect method to produce a homogeneous and well-determined distribution of Frenkel defects only, is by irradiation with highly energetic γ-rays, such as are produced, for instance, by a Co^{60}-source. Occasionally the γ-rays generate a fast electron in the interior of the

material, which in turn causes displacements. The range of γ-rays is very much larger than that of electrons, and homogeneous defect distributions in much thicker samples of material can therefore be obtained. On the other hand, the efficiency of γ-ray irradiation for defect production is very low, and the method can only be used with advantage in those cases where the *initial stage* of radiation damage is investigated, for instance by the method of internal friction (see Chapters XIV and XVII). In Fig. 5 the nature of *radiation damage* caused by various kinds of irradiation is schematically reviewed. A discussion of this depends on the quantitative description of irradiation phenomena, which will be postponed to Chapter XIV.

There are many other properties of point defects that might be discussed in this Chapter, such as e.g. their influence on the physical properties. Since these vary, however, greatly in different kinds of material, the discussion can be carried out more properly in the appropriate chapters in Parts II and III of this book.

REFERENCES

BENNEMAN, K. H. and L. TEWORDT (1960); Z. Naturf. **15a**, 772.
CORBETT, J. M. and R. M. WALKER (1958); Phys. Rev. **110**, 767.
ETZEL, H. W. and R. J. MAURER (1950); J. Chem. Phys. **18**, 1003.
FUMI, F. G. (1955); Phil. Mag. **46**, 1007.
HAVEN, Y. (1950); Rec. Trav. chim. **69**, 1471.
HAVEN, Y. (1955); Rep. Conf. Defects in crystalline Solids, Bristol (The Physical Society, London).
HAVEN, Y. and J. H. VAN SANTEN (1952); Philips Research Reports **7**, 474.
LETAW, H. (1956); J. Phys. Chem. Solids **1**, 100.
MOTT, N. F. and R. W. GURNEY (1948); Electronic processes in ionic crystals (Oxford Univ. Press).
MOTT, N. F. and M. J. LITTLETON (1938); Transact. Faraday Soc. **34**, 485.
PANETH, H. (1950); Phys. Rev. **80**, 708.
SEEGER, A. (1956); Encyclopedia of Physics, Part VII–1, Crystal Physics I.
SEEGER, A. and E. MANN (1960); J. Phys. Chem. Solids **12**, 326.
SEEGER, A., P. SCHILLER and H. KRONMÜLLER (1960); Phil. Mag. **5**, 853.
SEITZ, F. (1954); Rev. Mod. Physics **26**, 7.
SEITZ, F., J. S. KOEHLER and W. BAUERLE (1957); Phys. Rev. **107**, 1499
TEWORDT, H. (1958); Phys. Rev. **109**, 612.
VINEYARD, G. H. and G. J. DIENES (1954); Phys. Rev. **93**, 265.
VINEYARD, G. H., J. B. GIBSON, A. N. GOLAND and M. MILGRAM (1960); Phys. Rev. **120**, 1229.

CHAPTER III

GEOMETRICAL DESCRIPTION OF DISLOCATIONS

1. Structure of Dislocations

To describe the characteristics of the linear imperfections called dislocations, we shall follow FRANK [1951] and introduce the concept of the Burgers circuit in a lattice. Consider two crystal lattices, one belonging to an *actual* crystal, which contains various kinds of imperfections, and another the frame of an hypothetical *ideal* crystal, without any fault whatsoever.

When it is tried to establish a one-to-one-correspondence between the atoms of the actual crystal and those of the ideal crystal, this is found to be possible without difficulty when only external strains, thermal vibrations, etc. are present. These distort the actual lattice somewhat, but leave no doubt as to the origin of the displaced atoms, compared to their sites in the perfect crystal. Even when point defects are present, the one-to-one-correspondence is not destroyed, only an atom is lacking here and there, or an extra one incorporated. Each region of the actual crystal where such a correspondence exists, is called a region of *good* crystal. Where this is not the case, one speaks of a region of *bad* crystal. Bad crystal regions can evidently be formed by inclusions of a second phase either of the same or of a different material. We shall exclude these trivial cases from the following discussion, and consider only a homogeneous phase (which need not, however, be a single crystal).

A closed circuit of arbitrary shape, running from atom to atom entirely through good crystal regions in the actual crystal (Fig. 1*a*) is called a *Burgers circuit*. We can draw the *associated circuit* in the ideal crystal by making use of the one-to-one-correspondence between atoms. The associated circuit need not necessarily be closed, even if local elastic displacements are corrected for. When it is not closed the

associated circuit is said to encircle a *dislocation* (Fig. 1*b*).

Any bad region must have a narrow dimension in at least one direction, otherwise it would constitute a second phase, which case we have excluded. Even so, a point imperfection does not transform an originally good region into a bad one, as it is still possible to establish the

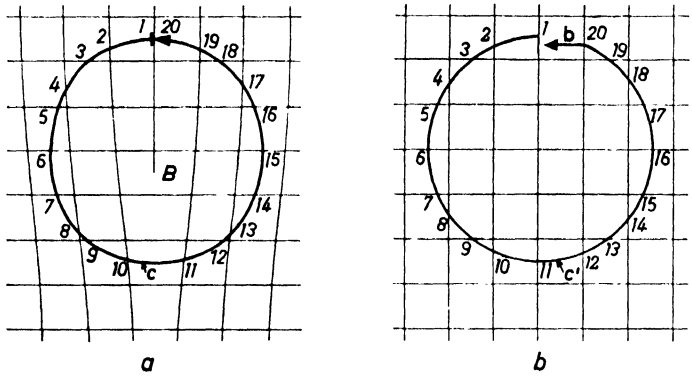


Fig 1. *a*) Cross-section of real crystal containing the region of bad crystal *B*. A Burgers circuit *c* has been drawn and the intersections with the lattice planes enumerated. *b*) Corresponding part of an ideal crystal. In the central part a one-to-one correspondence between atom sites cannot be established. The circuit *c'* is associated to circuit *c*; it has a closing failure equal to a lattice vector (vector *b*).

one-to-one-correspondence with the atoms of a perfect crystal in the presence of point defects. Defects of linear form are thus needed to procure a bad region. The simplest way in which the one-to-one correspondence can be destroyed is to construct a line defect with the property that the atomic displacements associated with it are *multiple-valued*. This is what actually has been done in Fig. 1*b*.

Then, when describing a Burgers circuit in the actual crystal that encircles such a defect over a full 360°, the end point never coincides with the starting point. In the terms used previously: the Burgers circuit has a closing failure equal to the residue of the multivalued displacement function. Such a *dislocation* cannot be eliminated simply by substituting good material for the bad material that forms its axis; in other words, it can not be compared to a row of vacancies of interstitial atoms, because such a row has no multivalued displacement around it. A dislocation is an imperfection of a wholly different character.

Assuming that, apart from the thin core of bad material along the dislocation axis, the crystal is good, and thus that the Burgers circuit, as it must do, runs entirely through good material, the closing failure is necessarily a lattice vector, and the line imperfection is called a *perfect dislocation*. The closing failure vector is called its *Burgers vector*. The sign of this vector can be fixed as soon as we have defined once and for all the direction in which any Burgers circuit has to be traversed. In the absence of other dislocations or faults destroying the goodness of the crystal, the Burgers circuit can be continuously displaced around and along the dislocation while retaining the same Burgers vector. A change in this quantity can only occur when part of the circuit enters a region of bad crystal. This leads to the following theorem: a dislocation has everywhere the same Burgers vector, and therefore cannot end within the crystal, except at the surface, a crystal boundary or at another dislocation or more general imperfection. Thus generally dislocations in the interior of a crystal will form closed loops or interconnecting networks. It is easily proved that the vector sum of the Burgers vectors of all dislocations meeting at a node of such a network is zero. This follows from the obvious corollary to the aforementioned theorem: when a dislocation line branches, is the sum of the Burgers vectors of the branches necessarily equal to the Burgers vector of the unbranched dislocation.

The existence of a closing failure $\boldsymbol{b}$ of the Burgers circuit entails that, in order to remove the core of bad crystal along the dislocation, one has to introduce certain changes through the whole region of good crystal encircling the dislocation line. The simplest way to do this is to make a cut into the crystal that ends on the dislocation axis, to displace both halves of the cut surface by the amount $-\boldsymbol{b}$, and to weld the two crystal parts together again (Fig. 2). It is usually necessary, as in the example of Fig. 2, for enabling the required displacement to add or take away atoms from the crystal. Conversely, one can imagine a dislocation to be generated by cutting the crystal halfway and inserting or removing some material to obtain a certain displacement. From this reasoning, a certain analogy to the so-called *Volterra dislocation* occurring in the theory of elasticity becomes apparent. More is said about this in Chapter IV.

As will be discussed in that chapter, there are six ways, differing in the relative orientations of cut surface and displacement vector, in which a Volterra dislocation in a continuous medium can be construct-

ed; accordingly, there are several distinct types of crystal dislocation also. J. M. BURGERS [1939] applied the theory of Volterra dislocations extensively to crystalline materials. It can be shown that here only the first three Volterra dislocations are of importance; crystalline defects corresponding to the other types can be imagined but they are not of much interest (see Chapter XIX).

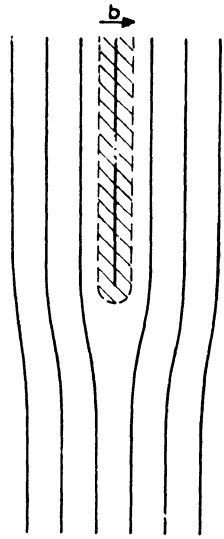


Fig. 2. A region of bad crystal along a dislocation axis can be repaired by cutting the crystal in half as far as the dislocation line, removing or adding some material (hatched region) in order to ensure a relative displacement of the cut faces opposite to the Burgers vector $\boldsymbol{b}$ of the dislocation, and joining the cut surfaces again.

The three "allowed" dislocations belong to two physically distinguishable groups:

a) The displacement corresponding to the Burgers vector is directed perpendicularly to the dislocation axis. Such dislocations are called *edge dislocations.*

b) The displacement vector is parallel to the dislocation axis. This dislocation is called a *screw dislocation.*

These two principal types of dislocations have been illustrated in a less abstract manner in Figs. 3 and 4. It will be clear from these figures

that both represent borderlines between perfect regions of the crystal, i.e. regions which have been displaced relative to each other over a distance that is equal to the Burgers vector. By moving an edge dis-

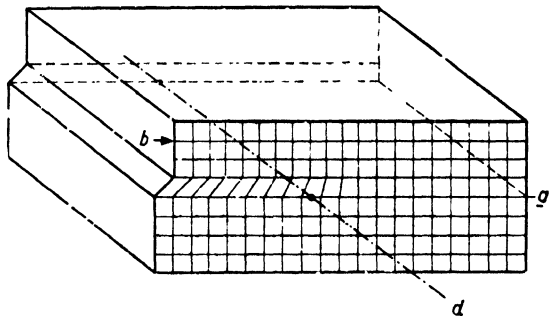


Fig. 3. Edge dislocation with Burgers vector **b**, perpendicular to the dislocation line *d*. This can move on the glide plane *g*.

location along the plane defined by the dislocation axis and the Burgers vector, the "slipped" region grows or diminishes in size relative to the unslipped regions. The same is true for the screw dislocation,

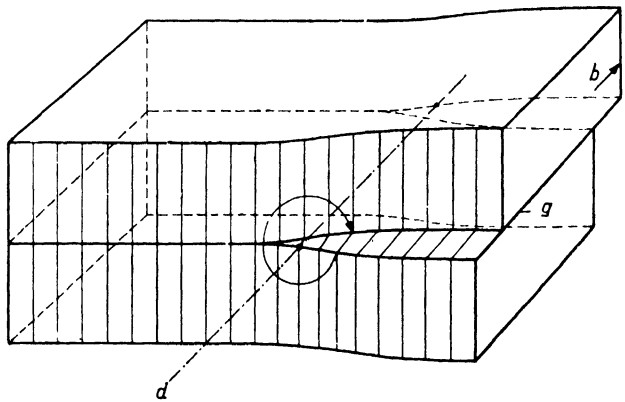


Fig. 4. Screw dislocation with Burgers vector **b**, parallel to the dislocation axis. This can move in any plane, that is, the glide plane *g* as indicated can be any other plane through the dislocation line *d*.

only here the *glide plane* can be any plane through the axis. It is because of this ability to produce crystal glide, that dislocations derive their importance in the theory of plastic flow of crystalline solids. This flow, as is well known, takes place predominantly by the kind of slip just described.

A general dislocation line has neither pure edge nor pure screw character; that is, its Burgers vector inserts an arbitrary angle with the dislocation axis. Moreover, this axis need not be straight. Thus we arrive at the picture of a dislocation line of arbitrary shape. The type of the dislocation changes continuously along its length. This has been depicted in Fig. 5.

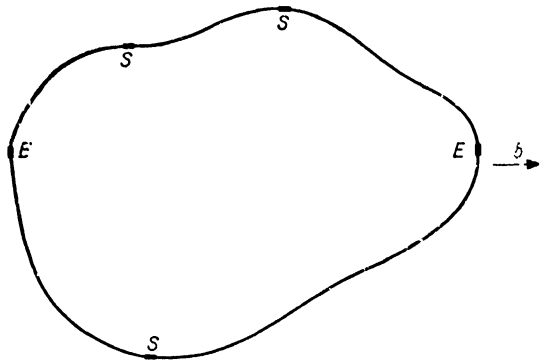


Fig. 5. Dislocation line of arbitrary shape. The Burgers vector **b** is everywhere the same, thus the type of the dislocation changes along its length. Pure screw and pure edge parts have been indicated by S and E respectively.

The two parts of the crystal on both sides of the dislocation differ only in having suffered a relative translation **b**, which is a lattice vector. The lattice is transformed into itself when a perfect dislocation has passed. It is possible, of course, to imagine a dislocation with a Burgers vector that *differs* from a lattice vector. Then necessarily the Burgers circuit can not have run entirely through a region of good crystal, because such a dislocation does *not* transform the lattice in itself. The circuit must have intersected at least one surface defect on which the crystalline registry changed, e.g. where the crystal changed to its twin orientation. This surface defect ends on the dislocation line. One calls such a dislocation a *partial dislocation*, and the surface defect associated with it is named a *stacking fault*. The stacking fault need not necessarily run right through the crystal, it may end on another partial dislocation, the Burgers vector of which adds up to that of the first to a complete lattice vector. Such a system of two partial dislocations connected by a stacking fault is called an *extended dislocation*.

In crystal physics interest is centered on those partial dislocations that transform the lattice into another closely related to it, e.g. into the

twin structure. A well-known partial dislocation in the face-centred cubic lattice is shown in Fig. 6; its stacking fault consists of four "hexagonally" stacked layers inserted between the "cubically" stacked octahedral layers (HEIDENREICH and SHOCKLEY [1947]). The hexagonal

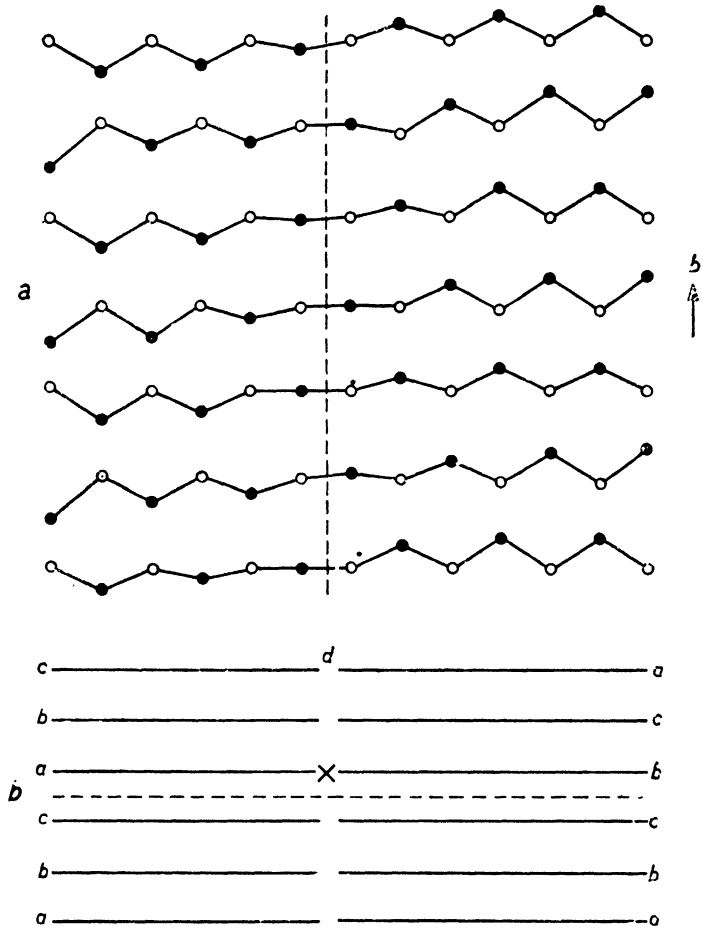


Fig. 6. *a*) Heidenreich-Shockley half dislocation of screw type in the f.c.c. lattice. The dislocation line *d* and the Burgers vector *b* are both along the same [11$\bar{2}$] direction. The dislocation is seen perpendicular to its (111) glide plane. Atoms in the plane of drawing are denoted by open circles, atoms one layer above it by dots.
b) On both sides of the dislocation the stacking is different. On the left hand side it is normal ...*abcabc*..., on the right hand side it has become ...*abcbca*... Thus a stacking fault has been produced. This can end in another half dislocation, with a Burgers vector inserting an angle of 120° with the former one. The two together with the stacking fault, form an extended dislocation.

stacking differs from the cubic in that the sequence of the three possible angular orientations (differing by a 120° degrees-rotation about the common ⟨111⟩ normal) of the octahedral planes is ...*abababab*..., rather than ...*abcabcabc*... .

Where the two partial (or "half") dislocations of this "Heidenreich-Shockley"-type extended dislocation are brought together, a perfect dislocation is formed with the Burgers vector along a [110] direction (those of the two partial dislocations were along [112] directions). Conversely, a perfect dislocation of this kind in a close packed structure can split up into an extended pair. Whether this splitting will actually occur and how wide the stacking fault will be, depends entirely on the energy of the system; the two partials together generally represent less elastic energy than the original dislocation alone, but the stacking fault also contributes considerably to the total energy of the system. We shall return to this question later.

The nature and properties of a dislocation segment are determined completely by the length and orientation of its Burgers vector, and by the orientation of its axis (represented by a unit vector). Both vectors together define the *slip plane* of the dislocation; only in the case of a straight screw dislocation this plane is arbitrary. A dislocation can move fairly easily in its slip plane since then only shearing motions of the atom rows occur (see section 5); however, to leave that plane, dilatational changes in the stacking of the atoms must be produced. Unless these are corrected by atomic diffusion, either voids or interstitial agglomerations occur. Consequently, the latter kind of motion (*nonconservative* or *climb motion*) is much more difficult than the *glide motion* in the slip plane. The second kind of movement is of the greatest importance in the *plastic deformation* of a crystal, the former kind plays a certain role in the phenomena of *recovery* and *recrystallization* of cold-worked metals, and of the thermally activated kind of deformation that is called *creep*.

2. Displacements and Stresses around Dislocations

In this section we shall approach the problem of elastic stresses and displacements around a dislocation along the lines set out in section 1, that is, from the point of view originally adopted by J. M. BURGERS [1939] (the „inventor" of the screw dislocation). We shall only touch on the most important features of the problem, postponing the detailed

mathematical treatment of the dislocation theory to Chapter IV. Moreover, the considerations of this section apply only to dislocations in isotropic and continuous media. The formulae presented in it, however, show clearly the properties common to any dislocation in any crystal, and can therefore be regarded as representative. These are the formulae that are used nearly exclusively in this book (and in other books as well).

Starting from the definition of an elastic dislocation mentioned in the previous section, J. M. Burgers showed that the displacements around a closed dislocation loop could be represented by the *multivalued* expression

$$u = \frac{1}{4\pi}\left[b\Omega + \oint_c \frac{b \times \mathrm{d}r}{r} + \frac{1}{2(1-\nu)} \operatorname{grad} \oint_c \frac{b \times r}{r} \cdot \mathrm{d}r\right]. \quad (1)$$

Here b stands, as usual, for the Burgers vector of the dislocation, ν is Poisson's ratio, r is the radius vector measured from an arbitrary origin unto a position on the dislocation line and Ω is the solid angle subtended by the dislocation loop as seen from this origin. The integration is to be carried out over the whole dislocation loop c. The occurrence of the term $b\Omega$ causes the multivaluedness; it is by far the most important term, for it contains the mathematical formulation of the theoretical definition of a dislocation. The other terms were added by Burgers to ensure equilibrium of forces and compatibility with the equations of the theory of elasticity. PEACH and KOEHLER [1950] transformed the right hand side of equation (1) into the sum of a number of rather lengthy line integrals, which are well suited for numerical computation, however. In some cases, notably that of a straight dislocation in an infinitely extended continuum, these integrals can be evaluated in closed form. Thus, for example, one derives the following displacement functions connected with a straight edge dislocation along the z axis, with its Burgers vector, of magnitude b, along the x axis:

$$\begin{aligned} u &= \frac{b}{2\pi}\left\{\arctan\frac{y}{x} + \frac{1}{2(1-\nu)}\frac{xy}{x^2+y^2}\right\}; \\ v &= \frac{b}{2\pi}\left\{\frac{2\nu-1}{4(1-\nu)}\ln(x^2+y^2) - \frac{1}{2(1-\nu)}\frac{x^2}{x^2+y^2}\right\}; \\ w &= 0. \end{aligned} \quad (2)$$

For a screw dislocation along the z-axis, with Burgers vector $\boldsymbol{b}$, we have simply:

$$w = \frac{b}{2\pi} \arctan y/x \; ; \tag{3}$$

$$u = v = 0 \, .$$

In these formulae the multivaluedness of the displacements is clearly expressed.

From (2) it follows that around an edge dislocation a *dilatation* occurs. The density change of the material can be expressed by the formula

$$\frac{\Delta D}{D} = \frac{b}{4\pi} \frac{1 - 2\nu}{1 - \nu} \frac{\sin \theta}{r} \, , \tag{4}$$

when r and θ denote the cylindrical coordinates around the dislocation axis ($\theta = \arctan y/x$, $r = (x^2 + y^2)^{\frac{1}{2}}$). No dilatation occurs around a pure screw dislocation.

The displacements associated with a mixed dislocation along the -axis, the Burgers vector of which makes an angle φ (Burgers angle) with this axis, can be found by adding expressions of the type (2) and (3), weighing the screw displacement function by the factor $\cos \varphi$ and the edge displacement by $\sin \varphi$.

From the displacements follow the stresses σ_{ij} ($i, j = x, y, z$) by differentiating and introducing the shear modulus G. One derives for the edge dislocation:

$$\begin{aligned}
\sigma_{xx} &= -\frac{Gb}{2\pi(1-\nu)} \frac{y(3x^2 + y^2)}{(x^2 + y^2)^2} \; ; \\
\tau_{xy} &= \frac{Gb}{2\pi(1-\nu)} \frac{x(x^2 - y^2)}{(x^2 + y^2)^2} \; ; \\
\sigma_{yy} &= \frac{Gb}{2\pi(1-\nu)} \frac{y(x^2 - y^2)}{(x^2 + y^2)^2} \; ; \\
\sigma_{zz} &= \nu(\sigma_{xx} + \sigma_{yy}) \; ;
\end{aligned} \tag{5}$$

and for the screw dislocation:

$$\tau_{xz} = -\frac{Gb}{2\pi}\frac{y}{x^2+y^2}\,;$$
$$\tau_{yz} = \frac{Gb}{2\pi}\frac{x}{x^2+y^2}\,; \tag{6}$$

all other stress components are zero.

When the solid containing the dislocation has finite dimensions, additional terms that ensure vanishing tractions at the surface have to be added; these are not very important for a general understanding of the problem (see also Chapter IV). At distances very close to the dislocation axis the expressions (2) – (6) become invalid, owing to the fact that the formulae from the theory of linear elasticity represent only approximations for the case of small displacements. Actually, in crystals the formulae should not be used for quantitative purposes within the first two atomic distances from the dislocation axis. The properties of the dislocation *core*, thus defined, are not very well known. A good description requires the methods of higher-order elasticity theory or, better, electron theory of metals. A start has been made in this direction recently (SEEGER [1956a]). In the absence of detailed knowledge about the dislocation core, the elastic *energy* of such an imperfection can only be estimated. Fortunately, this difficulty does not pertain to the long-range stressfield of a dislocation. The dislocation stresses and strains decrease at moderate and large distances *inversely proportionally* to the radial distance from the dislocation axis; a dislocation has thus a much more extended stress field than a point defect, for the stresses due to the latter decrease proportionately to the inverse third power of the distance (see next chapter). As a natural conclusion it follows that the elastic energy of a dislocation in an infinite crystal diverges logarithmically. Only by putting an effective limit to the range of stresses (e.g. the size of the crystal, or the average distance to other dislocations) can this divergence be removed. One then finds for the elastic energy of a dislocation of mixed type (Burgers angle φ), with the exception of the core energy, per unit length of the imperfection:

$$E_{el} = \frac{Gb^2}{4\pi} P \left\{\frac{\sin^2\varphi}{1-\nu} + \cos^2\varphi\right\}, \tag{7}$$

where P stands for a factor of the form $\log(r_e/r_i)$, r_e being the limiting

radius of the stress field, r_1 the radius of the dislocation core. The core energy can only be estimated approximately; it depends sensitively on the position of the dislocation axis with respect to the atomic layers, and only of course on the type of lattice. The core energy usually amounts to less than one eV per atomic plane; that is, it is of order $Gb^2/_{10}$. Taking $r_e = 10^{-4}$ cm, $r_1 = 3 \times 10^{-8}$ cm, one finds the "exterior" elastic energy to be about $\frac{1}{2}Gb^2$ per unit length. Thus the latter contribution to the energy of the dislocation is greater than the former. In view of the uncertainties in the computation we shall simply put the *total elastic* energy of a dislocation in an arbitrary crystal equal in order of magnitude to Gb^2 per unit length or Gb^3 per atomic distance along the axis. The quantity Gb^2 is often called the *line tension* of the dislocation. The line tension of a circular dislocation loop with coplanar Burgers vector was computed by NABARRO [1952]; again an expression practically equal to Gb^2 was obtained.

The *force* on a dislocation—a not very easily defined notion—exerted by an exterior stress field is directed perpendicular to the dislocation line. Per unit length of dislocation, it is equal to the product of the Burgers vector $\boldsymbol{b}$ and the component τ_b of the shear stress lying along this vector in the glide plane:

$$F = \tau_b \cdot b \,. \tag{8}$$

This can be illustrated by the following argument due to MOTT and NABARRO [1948].

Consider a cube of crystal of unit volume. When a straight edge dislocation of unit length, with Burgers vector $\boldsymbol{b}$, moves forward over the small distance $\mathrm{d}\boldsymbol{s}$, this results in the average relative displacement $|\,\mathrm{d}\boldsymbol{s}\,| \cdot \boldsymbol{b}$ of the two crystal halves on both sides of the slip plane. The applied stress on the crystal being σ, the work done thus amounts to $\sigma \cdot (\boldsymbol{b} \cdot |\,\mathrm{d}\boldsymbol{s}\,|) = (\tau_b \cdot \boldsymbol{b})\,|\,\mathrm{d}\boldsymbol{s}\,|$, which means that the "force" on the dislocation is logically defined as $F = \tau_b \cdot \boldsymbol{b}$ per unit length, irrespective of the orientation of that length. It is only this force component, acting in the glide plane, that interests us when the glide process is considered; the other force components affect the motion of dislocations appreciably only under "abnormal" circumstances (see below).

An important class of forces on dislocations arises from the mutual interaction between these imperfections, and from the interaction with other defects. In the following two sections we shall briefly discuss these forces, at least as far as the purely *elastic* interactions are con-

cerned. These are common to all solids; other contributions to the interactions such as those that arise from electrical causes, will be discussed in the chapters dealing in a more detailed way with the properties of specific crystals.

3. Interaction between Dislocations

This interaction can be computed with the aid of formula (8), by considering the component of the shear stress in the direction of the Burgers vector of dislocation 1, caused by dislocation 2. The latter is given by formulae of the kind (5) and (6). Thus we find only the force component in the glide direction. Another way to derive the interaction forces is to calculate the energy necessary to introduce dislocation 2 into a crystal that already contains dislocation 1, and subtracting from this the "self energy" of dislocation 2, according to (7). In the latter way we obtain the total energy of interaction, and thus by differentiating, all force components. It is easily proved that (apart from second order corrections) the interaction energy between an edge dislocation and a screw dislocation parallel to it vanishes (COTTRELL [1953]), as the stresses caused by the edge dislocation are everywhere perpendicular to the shear strain caused by the screw, and no work is done in introducing the latter into a crystal containing the former. The force between two parallel edge dislocations with the same Burgers vector is also easily derived. Its components are

$$F_b = \frac{Gb^2}{2\pi(1-\nu)} \frac{x(x^2-y^2)}{(x^2+y^2)^2}, \qquad F_p = \frac{Gb^2}{2\pi(1-\nu)} \frac{y(3x^2+y^2)}{(x^2+y^2)^2}; \tag{9}$$

where F_b denotes the mutual force component in the common glide direction and F_p the component of the interaction force perpendicular to the glide planes; x and y are the components of the distance between the two dislocations in these directions respectively. The sign of the forces is not included in (9); it is positive in the case of two dislocations of the same sign, negative for dislocations of opposite sign. Thus the latter attract each other when they lie in the same glide plane. This follows immediately when it is realized that the dislocations can annihilate each other when they come together, which process is, of course, accompanied by a very considerable decrease in energy.

However, when edge dislocations of opposite sign are situated on neighbouring parallel planes, they cannot annihilate completely. However, they can recombine to form a row of vacancies or of interstitial atoms, as depicted in Fig. 7. This kind of interaction removes the long-range stress field; presumably it plays a very important role in the generation of point defects during plastic deformation (sections 5.3 and 6). If the two glide planes are too far apart, however, attraction makes place for repulsion at small distances in the x-direction. Conversely, two identical parallel edge dislocations on the same glide plane repel each other, but they attract when their distance, as measured perpendicularly to the glide planes, is larger than that measured in the common glide direction, thus parallel to these planes. Hence, an array of parallel edge dislocations of the same sign, arranged vertically above each other in parallel glide planes, forms a stable configuration.

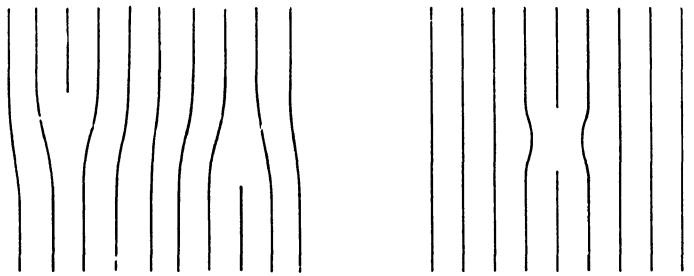


Fig. 7. Two edge dislocations of opposite sign (*a*) on neighbouring parallel glide planes can recombine into a row of vacancies (or interstitials). (*b*). This is usually an energetically favourable process, since the elastic strains are largely eliminated.

For two parallel screw dislocations of the same strength the situation is simpler. The only force component is directed along the distance between the two dislocation lines, and is given by

$$F_r = \frac{Gb^2}{2\pi}\frac{1}{r}, \tag{10}$$

where r denotes this distance (sign omitted). Now always repulsion exists between dislocations of equal sign, attraction between those of opposite sign. The case differs from that of two parallel edge dislocations because in the latter case there is a *tangential* component of the force which causes the interaction force in the glide direction to *change*

sign for $x = \pm y$. This tangential force is absent when only pure screw dislocations are considered.

The interaction between two unlike dislocations directed arbitrarily is more difficult to compute exactly. This interaction has some aspects in common with the electromagnetic interaction between two currents, running along the dislocation lines (PEACH [1951]). The current densities take the place of the Burgers vectors, the magnetic vector potential that of the displacement field (compare the "Kirchhoff" laws of conservation of Burgers vector, discussed in section 1). The interaction forces are governed by a kind of generalized Biot-Savart law:

$$F = \frac{\sigma \cdot \boldsymbol{b} \times \mathrm{d}\boldsymbol{l}}{\mathrm{d}l}, \tag{11}$$

where $\mathrm{d}\boldsymbol{l}$ is a line element of one of the dislocations of Burgers vector $\boldsymbol{b}$, acted upon by the stresses due to the other dislocation, that form the tensor σ. The simple expressions (9) and (10) form special cases of this general expression. This analogy, although not as perfect as it seems at first sight, may sometimes facilitate the discussion of dislocation interactions (PEACH and KOEHLER [1950], READ [1953]).

4. Interaction between Dislocations and Point Defects

The elastic interaction of dislocations with point defects was discussed in detail by COTTRELL and BILBY [1949], for the special case of an interstitial solute atom in an isotropic medium. These authors started from the very simple assumption that a solute atom with radius $r(1 + \varepsilon)$, where $\varepsilon \ll 1$ and r is the radius of the available hole in the parent crystal, interacts in first approximation only with the hydrostatic component of a stress tensor σ. The hydrostatic interaction energy amounts to

$$E_{\mathrm{h}} = -\tfrac{4}{3}\pi\varepsilon r^3 (\sigma_{xx} + \sigma_{yy} + \sigma_{zz}), \tag{12}$$

where σ_{xx} etc. denote the normal stresses at the site of the atom. To find the energy of interaction of the solute atom with a dislocation, the appropriate values for σ_{xx}, etc. must be inserted in (12). In this approximation, only dislocations with at least some edge character interact with point defects. This can also be understood physically: atoms with

a deviating radius will tend to move to those regions around a dislocation where the lattice is expanded or contracted, and the outsize atom can consequently be better accomodated. Screw dislocations, having no volume changes around them, are ineffective in this respect. For the interaction energy of a solute atom at a distance R of an edge dislocation, Cottrell finds:

$$E_h = \alpha G b \varepsilon r^3 \frac{\sin\theta}{R}, \tag{13}$$

where θ denotes, as in (4), the position angle with regard to the dislocation system; it is zero in the direction of the Burgers vector. The factor α is about 0.2 (CRUSSARD [1950]) for edge dislocations in metallic materials. BILBY [1950] and ESHELBY [1951]) have further extended the computations; the result was only a slight change of the numerical constant α.

Elastic interaction between solute atoms and a screw dislocation arises only when the atoms produce non-spherical distortions, and shear strains that can interact with the dislocation shear stresses. This case is realized for instance for interstitial atoms in body-centred cubic lattices, where the shear contribution to the interaction energy is of the same magnitude as the dilatational contribution given by (13). The difference between screw and edge dislocations in this respect is thus not so great as it seems; moreover, in most practical cases, interactions with both kinds of dislocations of a non-elastic nature play an important role too. These will be discussed in the appropriate chapters.

Solute atoms are thus in general attracted by a dislocation, and tend to cluster around it. The degree to which this occurs depends on the temperature (Chapter IX).

The elastic behaviour of vacancies and interstitials is analogous to that of centres of compression or expansion (see Chapter IV), and they are thus also attracted to edge dislocations. This applies especially to interstitials, as these have mostly large strains around them, which moreover are often even non-spherical, and can thus interact with screw dislocations also (section IX–1.2). Vacancies, on the other hand, do not have much elastic interaction with a dislocation, but the electrical and other effects associated with the absence of an atom or ion make up for this in most solids. Therefore, in general, there exists always some mutual interaction between dislocations and point defects.

A distinction from the case of a solute atom is that a physical point defect can be *absorbed* on a dislocation and then disappear completely. One way in which this elimination can easily take place is illustrated in Fig. 8. Consider an edge dislocation that jumps from one glide plane to the next, one atomic distance above it. Such a jump is called a *jog*.

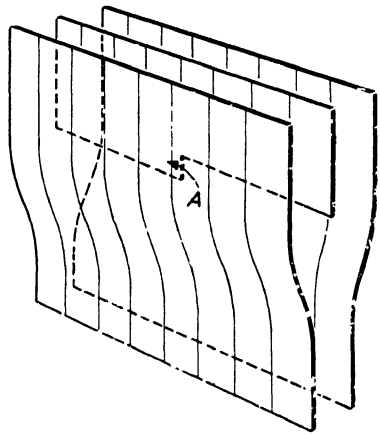


Fig. 8. Absorption of a vacancy by a jog in an edge dislocation. Motion of the vacancy to A results in the displacement of the jog over one atomic distance along the dislocation line to the left, and in the elimination of the vacancy.

For an edge-type dislocation the length of the extra half plane of atoms associated with it, differs by one atomic distance on opposite sides of the jog. When a vacancy arrives at the site of the jog, it can move to the place indicated by A, which results in a displacement of the jog by one atomic spacing along the dislocation axis, and in the simultaneous disappearance of the vacancy. Conversely an interstitial atom can be absorbed at the jog, thus causing it to shift one distance in the other direction. Jogs thus act as perfect sinks for point-defects. Point defects, once attracted to an arbitrary dislocation site, will migrate along it towards the nearest jog to be absorbed.

In thermodynamic equilibrium the reverse process—emission of point defects by jogs—takes place as frequently as absorption. It is generally accepted nowadays that jogs in dislocations, rather than the crystal surface or interior surfaces, are by far the most important sources for "thermal" point defects. SEEGER [1955] has reviewed the problem of the formation of point defects at jogs in metals, and has

studied the effect of motion of the dislocations on it. We shall return to this question as well as to the formation of jogs in the next section.

FRIEDEL and COULOMB [1957] have advanced arguments to show that in many non-polar substances the vacancies adsorbed at dislocations retain their identity, and are not eliminated at jogs. When present in large numbers, they should give rise instead to the formation of channels or voids along the dislocation lines, which are empty of atoms. The theory was proposed to explain several features of the work-hardening phenomenon in metals, to be discussed in Part II of this book.

Interaction of vacancies and interstitials with jogs in dislocations will be especially important in polar crystals. There, a jog can have an effective charge, either positive or negative, and therefore have not only a strong elastical, but also a strong electrical interaction with the point defects (which evidently also represent electric charges). More on this will be said in Chapter XXIV on polar crystals.

5. Motion of Dislocations

5.1. GLIDE MOTION

Two kinds of dislocation motions have been distinguished already: motion in the glide plane, also called *glide* motion or *conservative* motion, and motion in which the dislocation leaves its glide plane, so called *climb* or *nonconservative* motion. The first type of movement is, as has also been pointed out before, that which constitutes the macroscopic phenomenon of slip in crystals. It is an "easy" kind of motion, at least in solids with non-directional bonding forces between the atoms, thus in metals and polar crystals. There, its occurrence is practically independent of the temperature. In the valence crystals germanium and silicon glide motion is observed only at elevated temperatures, presumably because the severing of the directed bonding forces between the atoms needs to be aided appreciably by lattice vibrations. In metals and polar crystals the energy needed to bring a dislocation from one position to another in the same glide plane is small. It arises from the periodicity of the lattice structure: this entails the existence of discrete equilibrium positions of the dislocation axis, in which the distortion of the crystal around it agrees best with the lattice symmetry. In between the equilibrium positions the strain energy of the system: dislocation-crystal, is larger, as unsymmetric distortions repre-

sent a higher potential energy. PEIERLS [1940], NABARRO [1947] and DIETZE [1952a], see also LEIBFRIED and DIETZE [1951], have estimated the maximum energy difference between possible dislocation positions. They assumed a sinusoidal relation to exist between the displacement of a lattice particle from its equilibrium site and the restoring force, and considered moreover the crystal halves above and below the glide plane as elastic continua. These approximations constitute what is known in literature as the Peierls model, an extensive discussion of which has been given by SEEGER [1956b].

In the final expression for the energy (for the mathematical derivation the reader is referred to Chapter IV) an unknown parameter, viz. the *width* of a dislocation (expressed in terms of the lattice parameter), occurs in the exponent. This width measures the extension of the strained region around a dislocation; its value depends sensitively on the nature of the interatomic forces and on the actual dependence of the restoring force on the atomic displacement and on the temperature. Its magnitude can only be estimated, it may vary between 2 and 10 atomic distances in metals, which means an uncertainty of several factors of 10 in the "Peierls force" resisting dislocation motion. Calculations of DIETZE [1952b], in which the Peierls model was slightly modified, and of FOREMAN, JASWON and WOOD [1951] in which a more realistic, *an*harmonic function for the atomic restoring force was used, made it clear, however, that in the close-packed cubic metals the Peierls force should be small, corresponding to a shear stress of about $10^{-4}\,G$ or less opposing the dislocation motion. The latter value is comparable to the observed critical shear stress in well-annealed single crystals. Except under special circumstances, the Peierls force seems to play thus but an insignificant role in the plastic behaviour of metals. Direct experimental determination of the magnitude of this force is very difficult. One way of doing this will be described in Chapter XVII; the results seem to point to a Peierls force that is much larger than was expected, even in the face-centred cubic noble metals!

In valence crystals the situation is, seen from a theoretical point of view, still less satisfactory. There probably the fundamental dislocation friction determines the plastic properties of the material to a large extent. It is no longer of a purely elastic nature, however, chemical and electrical forces contributing appreciably to it. For a discussion we refer to Chapter XXX.

A dislocation that moves through a crystal contributes to the

macroscopic strain. A dislocation element of length l that covers a distance s sweeps out an area ls and thereby contributes to the relative displacement of the crystal halves on both sides of the glide plane the amounts $(ls/A)\boldsymbol{b}$, where $\boldsymbol{b}$ is the Burgers vector and A the total area of the slip plane. Thus the increment of plastic shear is lsb/V, if V stands for the volume of the crystal. Let the total length of dislocation per cm^3 of the crystal be $N\,(=l/V)$, then the macroscopic shear strain is Nsb. The quantity N is called the *dislocation density*, it has the dimension $[l^{-2}]$.

The glide motion of dislocations is not only obstructed by the "intrinsic" Peierls force, it is, at least in metals, usually far more seriously hindered by the presence of other lattice imperfections. The elastic interaction between point defects or solute atoms and dislocations causes in all solids moving dislocations to be attracted by these defects. They can even become more or less immobilized by them. By removing the deforming stresses, the dislocations come to rest in pinned positions, and renewed motion is only possible under appreciably larger stresses than were needed for the previous motion, because the dislocations must first be pulled loose again. This is the common reason for the existence of a sharply defined *upper yield stress* in many impure metals. Possibly similar effects play an important role in the behaviour of dislocations in the elemental semiconductors.

Even more efficacious than point defects in encumbering the motion of dislocations, are other dislocations. There are three ways in which the impediment takes place: *a*) by the elastic interaction (of long range, see section 3); *b*) by the action of dislocations intersecting the glide plane of the moving dislocation; *c*) by the action of dislocations more or less parallel to the moving one, but of such a type that on combination a dislocation of an immovable type (*sessile* dislocation) is formed.

Two dislocations with arbitrary Burgers vector that intersect each other, after the intersection each possess a discontinuity in their axis, already encountered earlier, and there called a *jog*. A jog can be described as a short segment of the first dislocation, of length equal to, and directed along, the Burgers vector of the second dislocation, and vice versa (Fig. 9). Since each jog forms part of the original dislocation, its Burgers vector is necessarily the same as that of its parent dislocation (theorem of the conservation of the Burgers vector, section 1), so that at the site of the jog the jogged dislocation has a different

character from elsewhere. The glide plane associated with the jog in general differs from that of the parent dislocation. This means that glide motion of the latter infers usually climb motion of the jog. As we shall see in the next section, the latter kind of motion requires an energy of activation in all kinds of solids, and thus the progress of a dislocation, even that in the original glide direction is strongly hindered by the presence of a jog, formed after intersection with another dislocation. In section 5.3 we shall describe in detail what happens when a jogged dislocation actually moves further.

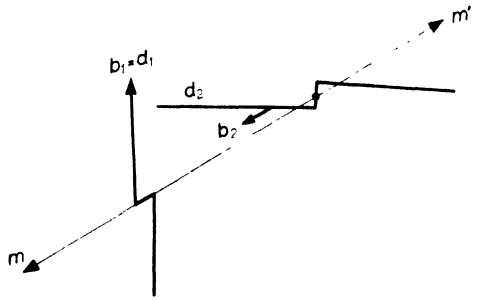


Fig. 9. Dislocation d_1 (screw type, Burgers vector b_1) has intersected with dislocation d_2 (nearly edge type, Burgers vector b_2). In both dislocations jogs have been formed parallel to the Burgers vectors of the other dislocation; mm' denotes the relative direction of motion.

Several distinct mutual orientations of the two dislocations are imaginable in which the jogs formed do *not* hinder the further glide motion of either of them. In a pure edge dislocation, jogs can never obstruct glide motion. Further, quite generally, if the intersecting dislocation has a Burgers vector coplanar with the glide plane of the other dislocation, the jog in the latter will have the same glide plane as that dislocation itself, and no obstruction occurs either.

Even more serious effects result from the intersection of two dislocations when these are split up into partials. Extended dislocations can only glide when the glide plane contains the plane of the stacking fault, or in other words, when the Burgers vectors of both partials lie in the plane of the stacking fault. A partial dislocation that has not this property can move either by performing a climb motion, leaving the plane of its stacking fault as it is, or by shifting the stacking fault to another crystallographic plane. This, of course, is a highly improbable process as it requires simultaneous movements of many atoms. Such a

dislocation is called *sessile*. A jog in an extended dislocation is accompanied by very large strains except when at its site the two partials have been brought together (Fig. 10). The formation of a jog, e.g. by the process of intersection necessitates therefore the local elimination of the stacking fault (*constriction*). As SEEGER [1954] has shown, (and will be further discussed in Chapter VIII), this usually requires a large amount of energy, especially in those f.c.c. metals such as aluminium where the stacking fault is normally wide. For all these reasons, considerable stresses are needed to force extended dislocations through each other.

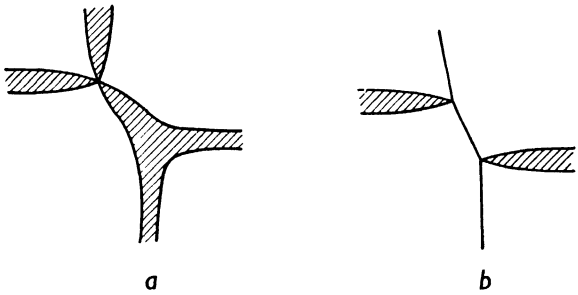


Fig. 10. *a*) If two extended dislocations attract each other and intersect, recombination occurs over a certain length and a structure is formed consisting of an *extended* and a *constricted node*. *b*) After passing, in both dislocations *jogs* remain, which requires *constriction* of the two stacking faults.

Either the two dislocations going to intersect attract each other or they repulse. In the former case a recombination reaction occurs over a certain length which was first observed by HIRSCH [1959] in his electron transmission photographs. This does not happen in the second case. If the (attracting) dislocations are extended, in the recombination structure socalled *nodes* are present: one distinguishes between *extended* and *constricted* nodes (Fig. 10a). The energy needed to form a constriction in an extended dislocation of width d can be estimated as about $0.1\, Gb^2d$.

Two extended dislocations on intersecting glide planes but with more or less parallel axes may also obstruct each others motion, if by combining two partials belonging to different dislocations a sessile (immobile) dislocation can be formed. This means that the whole system has become immobile, because the other two partials cannot move without extending the stacking fault belonging to them. Similar

locking phenomena can occur when two perfect dislocations combine and afterwards dissociate into sessile partials, as pointed out by LOMER [1951] and by COTTRELL [1952] respectively. This is of course only possible when such a separation into partials lowers the energy of the ensemble, that is, in special orientations. An important example in the face-centered cubic structure is formed by the mutual locking of two Heidenreich-Shockley extended dislocations on intersecting glide planes.

In the notation to be explained in section 7, this example is the following:

1st extended dislocation $\frac{1}{2}a\,[10\bar{1}]$, dissociated into $\frac{1}{6}a\,[11\bar{2}] + \frac{1}{6}a\,[\bar{2}11]$;

2nd extended dislocation $\frac{1}{2}a\,[011]$, dissociated into $\frac{1}{6}a\,[112] + \frac{1}{6}a\,[\bar{1}21]$;

Reaction: $\frac{1}{6}a\,[\bar{2}11] - \frac{1}{6}a\,[\bar{1}21] = \frac{1}{6}a\,[\bar{1}\bar{1}0]$.

The reaction product is a sessile dislocation, because its Burgers vector does not lie in either of the planes of its two stacking faults, (111) and $(11\bar{1})$. The three remaining dislocations form a typical triangular group, held together by these two stacking faults.

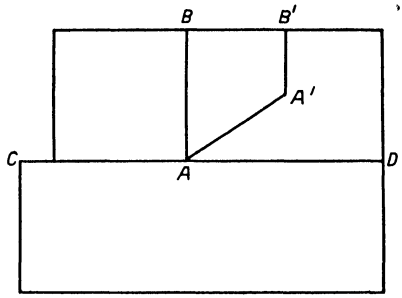


Fig. 11. Climb motion of a dislocation of not pure screw character (the axis is perpendicular to the plane of drawing). When the dislocation moves from A to A', the extra half plane BA becomes $B'A'$, that is, it becomes shorter. Some atoms have therefore to diffuse away or to become interstitial (after COTTRELL [1953]).

5.2. CLIMB MOTION

When a dislocation with some edge-character leaves its glide plane (a pure screw dislocation, whatever its motion, always moves in a glide plane) a defect state is produced that can only be repaired by diffusion, that is by transport of mass through the crystal. The extra half plane of atoms that ends on the dislocation axis (Fig. 11) becomes shorter or longer by one atomic distance, when the dislocation climbs

to a neighbouring glide plane. This means that one row of atoms has to be removed or added, respectively. This can only take place by diffusion, otherwise a row of interstitials or vacancies is formed.

In either case an energy of activation for the climb process is needed. When the climb continues the dislocation can be considered to act as a source or sink of vacancies or interstitials as the case may be and these have to be transported away. In all metals the formation and migration of point defects together also determine the rate of *self-diffusion*; therefore the activation energy for climb is usually equal to that for self-diffusion. Its presence implies that under ordinary stresses and at

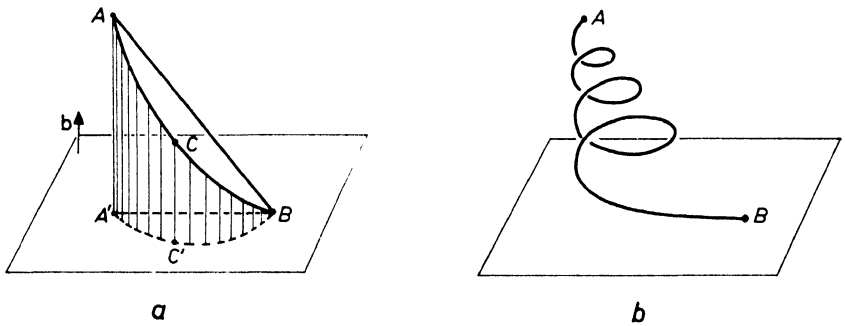


Fig. 12. *a*) The straight dislocation AB is locked at its end points. Its projection on the plane normal to its Burgers vector **b** is the broken straight line $A'B$. By *climb* motion, the pinned dislocation attains the spiral form ACB with the dotted line $A'C'B$ as a projection. Now, the curved dislocation can further *glide* along the cylinder generated by lines through its axis parallel to **b** (prismatic glide).
b) The dislocation may occasionally wind itself through several turns (when sufficient material is provided to make such excessive climb possible). It then forms a helicoidal dislocation, which again undergoes prismatic glide.

normal temperatures dislocation climb is an energetically unprobable process. In most metals it occurs only at elevated temperatures, where, for instance, it leads to the phenomenon of polygonization (section XI–2).

A special kind of controlled dislocation climb, which is probably not active in metals but is important in ionic crystals where ionic diffusion (electrolytic flow) is difficult, has been described and observed by AMELINCKX *et al.* [1957]. Dislocations of predominantly screw character, pinned at two ends, take up a curved form after a limited amount of climb, and thereby acquire an edge-type component (Fig. 12). Motion of the curved dislocation along the surface of a cylinder

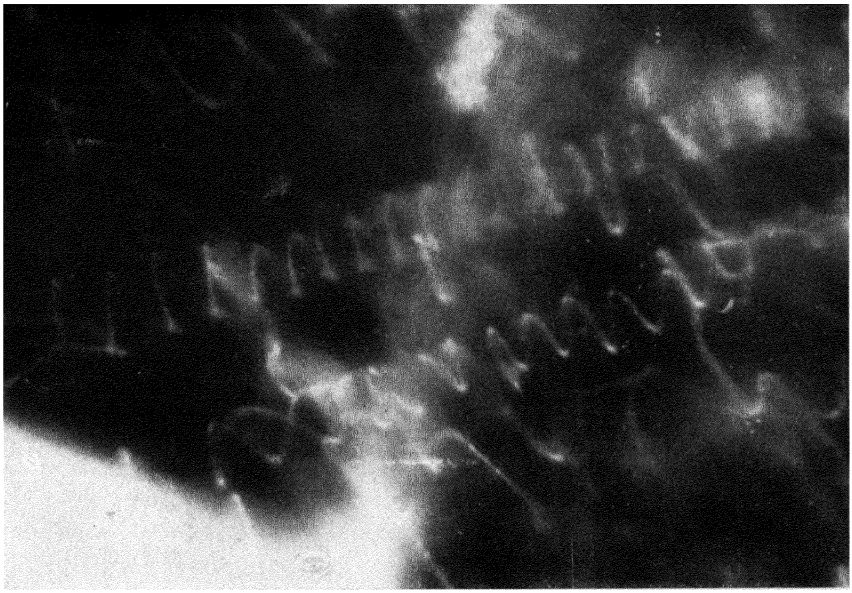


Fig. 13. Photomicrograph of decorated helicoidal dislocations in NaCl (AMELINCKX *et al.* [1957]).

generated by lines parallel to the Burgers vector can now occur without addition of material. This has been called *prismatic glide* by SEITZ [1950]. When the end points remain fixed, the combination of limited climb and prismatic glide can, under circumstances, result in the formation of so-called *helicoidal dislocations* (Fig. 13). Dislocation helices have been observed inter alia in calcium fluoride and the silver halides. It is predominantly the interaction between point defects and dislocations that controls the formation of helicoidal dislocations.

5.3. MOTION OF JOGGED DISLOCATIONS

The glide motion of a jogged dislocation in general includes climb motion for the jogs, as explained in section 5.1. Therefore, at temperatures not high enough for large-scale diffusion to take place in the solid, the moving jogs leave a trail, or string, of either vacancies or interstitial atoms behind them. The number of these depends on the type of the jog; if this has been created in a pure screw dislocation after the latter cut through a perpendicular screw, the spacing between the defects is as close as possible, that is, a continuous string of defects remains in the wake of the jog. Naturally, the applied stress must provide the necessary energy to form the defects.

SEEGER [1955] has studied in detail the action of moving jogs in metals. He considered the effect on the geometry of jog formation of the ratio between the defect energy and the line tension of the dislocation, and also that of temperature. At moderate temperatures, thermal activation will influence the rate of defect formation, and the diffusivity of the defects is usually large enough for the strings formed originally to scatter almost at once. It can be shown that the energy of formation of an isolated defect is somewhat higher than that of one that forms part of a string. Therefore the activation energy for motion of the dislocation increases with increasing temperature. The activation energy for jog motion moreover depends linearly on the stress τ (shear stress component in the glide direction). One has:

$$U = \alpha U_d - \tau b d l , \tag{14}$$

where U_d is the formation energy of an isolated defect, α a factor of order unity, d the width of the stacking fault (if present) and l the length of dislocation between two jogs. This formula can easily be understood, observing that the gain in potential energy of a dislocation

segment of length l, which contains one jog, when it moves the distance d is just τbdl.

An interesting aspect of the problem of the motion of jogged dislocation arises when the possibility of sidewise motion of the jog along its own dislocation is recognized. By this the efficiency of defect formation is of course lowered. However, the effective velocity of the jog then in most cases becomes very great and therefore its kinetic energy may rise to extremely high values (section 5.4). Moreover, the jog is more or less attracted to the string of defects which it forms. This also tends to prevent sidewise motion. Therefore we assume this motion to be severely restricted, although in actual cases a jog in a moving dislocation will nevertheless undoubtedly perform some kind of zig-zag motion.

It is certain that many point defects are formed during plastic deformation of all crystalline solids. Most authors believe this to be predominantly the result of the motion of jogged dislocations (SEITZ [1952], MOTT [1952], VAN BUEREN [1957]). However, according to FRIEDEL [1956], the generation of point defects by moving jogs should be an improbable process, even when sidewise motion is prevented. Instead, the point-defects should be formed by the combination of dislocation segments, as can occur for instance during the action of Frank-Read sources (section 6).

Then often dislocation segments of opposite sign pass each other closely on neighbouring glide planes. Such segments attract each other (see formula (9)) and can then easily recombine to form a row of vacancies or interstitials (Fig. 7). Friedel argues that this process must be preferred above the production of defects by moving jogs.

5.4. HIGH-SPEED MOTION

The stress field of a moving dislocation differs from that of a stationary one because of the existence of a finite upper limit to the speed of atomic motion, or, in other words, by reason of the finite velocity of sound, c. It is not possible for a dislocation to move more rapidly than this; mathematically this limitation is apparent in the following formula for the energy T of a dislocation segment, moving with velocity v:

$$T = \frac{T_0}{\sqrt{1 - v^2/c^2}}. \tag{15}$$

The formula bears a striking resemblance to that for the kinetic energy

of a moving mass in the special theory of relativity. This and further details have been mathematically discussed by, among others, FRANK [1950] and ESHELBY [1953]. The kinetic energy of a dislocation becomes comparable to its stationary ("rest") energy at about $\frac{1}{2}c$. But even before that already other factors begin to limit the speed of a dislocation, particularly the emission of sound waves. This is equivalent to a kind of "radiation damping" of the motion of the imperfection. Moreover, the interaction of the displacement field of a moving dislocation with the thermal vibrations of the lattice, as well as the temperature variations caused by its motion (through the thermoelastic effect) produce energy losses (NABARRO [1952]).

The whole complex problem of the energy dissipation of a moving dislocation is not yet completely solved. The work of ESHELBY [1949], NABARRO [1951], and LEIBFRIED [1950] cannot be regarded as more than a first approximation.

Measurements of the velocity of individual dislocations have recently been carried out by GILMAN and JOHNSTON [1959], and in a less direct way also by PENNING [1959] and VAN BUEREN [1959]. We come back to this in Chapters VI and XXX, but point out here already that the observed velocities are generally smaller by many orders of magnitude than the velocity of sound.

6. Formation of Dislocations

It will be clear from section 5.3 that point defects can be formed in great numbers during plastic deformation, either by the action of jogs in moving dislocations or by recombination of suitable dislocation segments. Estimates of the densities of point defects thus obtained agree moderately well with actual determinations (Chapter XV). No problem, therefore, exists in explaining the occurrence of point defects in crystals. The origin of dislocations, on the other hand, is as yet much more obscure. No matter how carefully one prepares a, not too small, metal crystal, it is always found to contain very many dislocations per cm^2, usually at least 10^6 cm^{-2}. The same density is encountered in most polar crystals. Homopolar crystals are usually far more perfect; common crystal-growing techniques lead to densities of about 10^4 cm^{-2}, whereas by very careful control of the temperature gradients during cooling, a dislocation density as low as a few dislocations/cm^2 can be realized. The same low density has been reached in a few care-

fully prepared ionic crystals, but must be considered as exceptional.

Dislocations are not thermodynamically stable lattice defects, (COTTRELL [1953]), as they do not appreciably lower the configurational entropy of a crystal through their presence, and their energy of formation is very large. Several theories as to the possible causes of their origin have been proposed. Mostly these apply to specific materials solidified under special conditions, and care must be taken in making generalizations.

For metal crystals (and *this* idea is probably applicable to other solids as well), drawn from the melt by one of the conventional techniques the hypothesis of TEGHTSOONIAN and CHALMERS [1951], elaborated by FRANK [1956] but criticized by several others, might perhaps apply. The existence of a thermal gradient in the newly crystallized

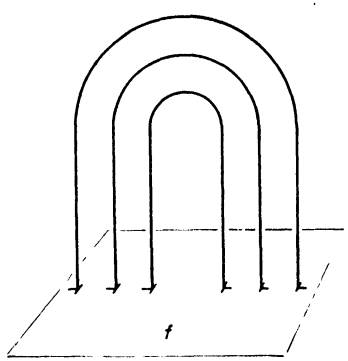


Fig. 14. Elongated loops of edge dislocations perpendicular to the growth front f of a solidifying crystal. These have been formed by the collapse of supersaturated vacancy disks, arranged in the plane of the loops (after FRANK [1956]).

solid implies that the thermal equilibrium vacancy content varies from a high value near the hot liquid-solid interface to a much lower value in the cooler parts of the crystal. Frank showed that this variation of the vacancy concentration with the distance and therefore, since the crystal is steadily growing, with time of the vacancy concentration at a given site is generally much too rapid for it to be continuously established by diffusion of the excess vacancies out of the solid in the time therefore available. The excess vacancies must therefore somehow condense within the solid. It is now regarded as very probable that the condensation products are flat disks parallel to the growth direction

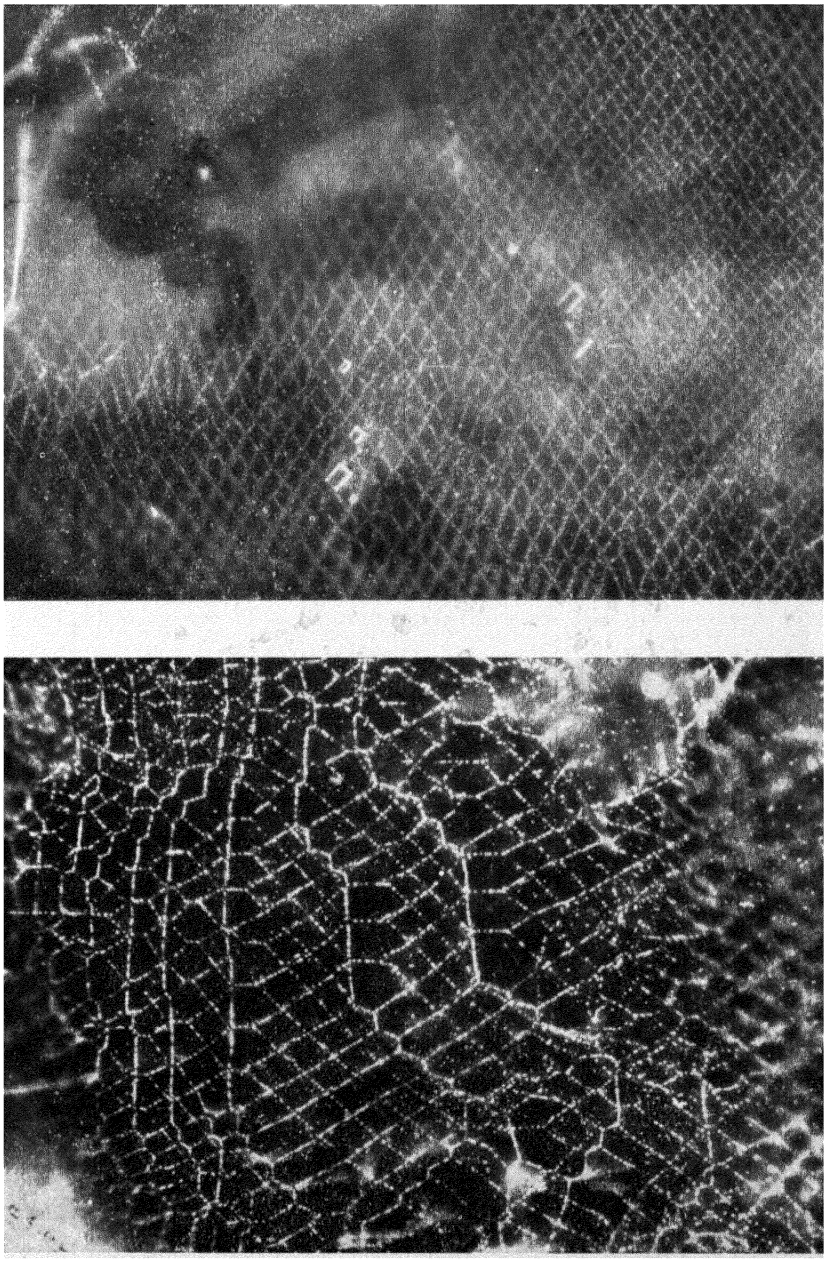


Fig. 15. Two examples of visually observed spatial dislocation networks in KCl crystals (after AMELINCKX [1958]). *a*) KBr(+AgBr); *b*) KCl (+AgCl); (600×). For a description of the technique leading to these micrographs we refer to Chapter V.

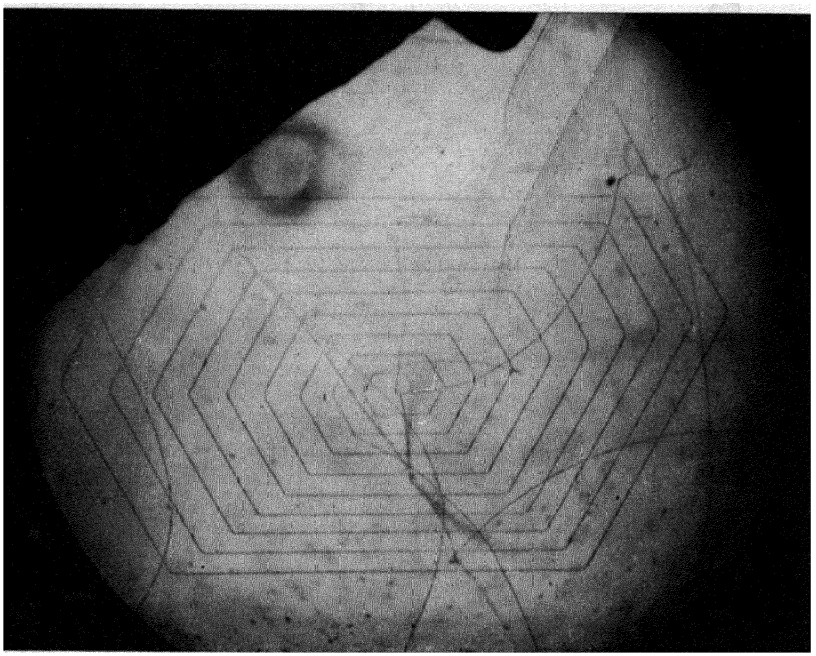


Fig. 17. Example of Frank-Read sources observed in silicon (after DASH [1956]).

of the crystal, which increase in size by the edgewise addition of new vacancies and thus follow the advancing growth front. Such a disk can collapse, and leave a U-shaped edge dislocation with a Burgers vector normal to the plane of the disk. Many edge dislocations may be formed in this way, and these can afterwards assemble, by the result of their mutual interaction, in configurations representing a state of lower potential energy. Systems of elongated loops, in intersecting planes, will form in such a way that minimal stresses remain, and something resembling a spatial dislocation network is generated (Figs. 14 and 15).

It is usually assumed that in unstrained crystals the dislocations are indeed present as a network, as the line tension tends to shorten their total length as much as possible. The situation can be somewhat compared to a raft of soap froth. Moreover, such a network, where three or more dislocations meet in so-called nodes, is not only elastically stable, but also geometrically possible in many crystal structures (see section 7).

KUHLMANN-WILSDORF *et al.* [1958] have recently proposed a theory in which it is assumed that dislocation formation by vacancy condensation takes place not only during growth but also during the plastic deformation of metals. The loops ("R-dislocations") formed would then subsequently act as Frank-Read sources. We shall come back to this interesting concept in Chapter XIV.

Thermal stresses set up by non-uniform cooling of a crystal, can also lead to the formation of considerable numbers of dislocations during crystal growth, as was shown by BILLIG [1956] and by PENNING [1958] for germanium. In fact, thermal stresses are to a large extent responsible for the dislocation density of 10^4 cm^{-2}, quoted above, which is usually encountered in crystals of this material. The discussion of the mechanism of generation of dislocations by thermal stress will be postponed to later in this book (Chapters XIV and XXX) as it is largely a matter of plastic deformation.

In crystals grown from the slightly supersaturated vapour or from a slightly supersaturated solution the so-called dislocation mechanism of crystal growth can operate, as will be described in Chapter XXIII. A crystal grown by this mechanism necessarily contains a certain length of dislocation. As to the origin of the dislocations in the small nuclei from which this growth starts, HOLLOMON [1954] has advanced arguments based on visual observation, showing that at least in some cases these imperfections form spontaneously during the accidental encounter of two nuclei.

After this treatise was finished, a conference report entitled Growth and Perfection of Crystals (Wiley, New York [1959]) appeared, in which a large amount of evidence concerning the formation of imperfections during growth has been collected.

Whereas it seems now at least not improbable that in some way or another, dislocations will almost certainly be generated during crystal growth, and will thus be present in crystalline solids, it remains to be explained why in plastically *deformed* crystals the dislocation

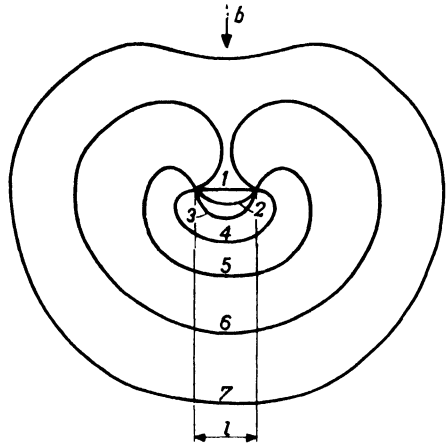


Fig. 16. Frank-Read source. A dislocation element 1 of length l and Burgers vector $\boldsymbol{b}$ is anchored at its ends. When a shear stress is applied along $\boldsymbol{b}$, the dislocation bends out and takes the forms 2, 3. At a certain stress the critical half-circle-form 3 has been reached. A further slight increase of the stress causes a spontaneous expansion 4, ... 6. In the latter situation the two near lying screw parts eliminate each other, and the closed loop 7 is formed together with the original piece of dislocation 1.

density is always many orders of magnitude higher than in the as grown crystals. The dislocation density in moderately deformed metals or polar crystals varies between 10^{10} and 10^{12} cm^{-2}; in deformed homopolar substances it ranges from 10^{6} to 10^{8} cm^{-2} (Chapter VII).

This increase is due to the existence of a dislocation *multiplication* mechanism that is activated by the deforming stress, and which contributes essentially in making plastic deformation such a relatively easy process in most crystals. Many different multiplication mechanisms have been proposed, but as they are all closely related to each other,

they may be illustrated generally by the classical Frank-Read source model (Fig. 16).

This is formed by an arbitrary piece of dislocation, of length l, anchored in some way at its ends, so that it cannot move as a whole but only stretch sideways under the action of a shear stress. Consider for example a pure edge dislocation in a shear stress field τ parallel to its Burgers vector b. The equilibrium shape of this dislocation is an arc of a circle, as can be easily shown by balancing its line energy against the potential energy gained in moving with the stress. The radius R of the arc is determined by the condition:

$$\frac{Gb}{2R} = \tau .$$

As soon as

$$\tau = \frac{Gb}{l},$$

the arc becomes a semicircle and the equilibrium critical: on any further increase of the stress, the dislocation loop will expand spontaneously, in the manner of Fig. 16. The lateral parts have screw character and move sideways. Behind the source, two screw parts of opposite sign (and opposite direction of motion) approach each other and can finally eliminate each other mutually. Then the original source dislocation is reformed, but in addition, a closed dislocation loop has been generated which expands freely onwards. The whole process can now be repeated in exactly the same manner. If the orientation of the source dislocation is not purely edge, the recombining rear portions of the loop need not be of purely screw character. When, moreover, these rear portions do not lie in the same glide plane, as will be the case when the dislocation has intersected another dislocation during its sweep around the anchor points, a complete elimination of both pieces of dislocation is impossible. Instead, a row of vacancies or of interstitials is formed, as was explained in sections 2 and 5.3. This mechanism of defect formation was first proposed by FRIEDEL [1956].

In principle, an unlimited number of loops can be formed by a Frank-Read source, if only the local stress does not decrease below the *critical* value

$$\tau_{cr} = \frac{Gb}{l}. \tag{16}$$

In practice, the number of loops emitted *is* limited, as will be discussed in Chapter VII. However, by the simple source mechanism described and by others closely related to it, the original density of dislocations in a solid may easily be increased by several orders of magnitude, under the action of stresses in the plastic range.

An original dislocation configuration in the form of a spatial network should be excellently suited to provide F-R sources, since each network element can in principle act as such. The nodes, where three (or more) dislocations meet, are relatively immobile, owing to the different kinematic properties of the composing dislocations. In silicon single crystals as well as in some polar crystals the behaviour of incidental network elements as Frank-Read sources has been observed visually (Fig. 17, DASH [1956], AMELINCKX *et al.* [1958]). In metals, although direct observations are scarce, source action is probably as common. KUHLMANN-WILSDORF and WILSDORF [1954], observing Cu-alloys, made visible parts of the loops of a system that is probably a Frank-Read source. However, it has become quite clear recently from a large number of various observations, optical as well as mechanical, that the dislocation network if present in a material is usually only of secondary importance in so far as the provision of sources is concerned. The surface seems to play an essential role in this, the nature of which is still unknown at present.

7. Types of Dislocations Occurring in Various Crystal Structures

There are several factors that limit the number of distinguishable types of dislocations in real crystals. First, restrictions exist as to the possible direction of the dislocation line. However, these are usually not very rigid, except for a preference for low-index crystallographic planes. Consequently, the position of a dislocation line in its glide plane, and hence the type of the dislocation, cannot be generally predicted. The line tension, which tends to shorten a dislocation as much as possible is not much different for different line directions (for some quantitative results, compare for instance a paper by FOREMAN [1955]). The preference for low-index planes results also from the effect of line tension: this depends on the actual atomic spacings along and across the dislocation line and reaches a minimum when the interplanar spacing perpendicular to the glide plane is as large as possible, which is just the characteristic of low-index planes. The line tension is also propor-

tional to the square of the Burgers vector (formula (7)), and this property restricts the choice between possible Burgers vectors very much. A dislocation with an arbitrary Burgers vector b, that can split up into two or more new dislocations with Burgers vectors b_1, b_2, etc., will only be stable if:

$$b^2 \leqslant b_1^2 + b_2^2 + \ldots\ldots . \tag{17}$$

Otherwise, the total line energy of the system will decrease by the splitting, and the latter process will occur spontaneously. Of course, in applying the criterion (17) due regard must be paid to slight differences from the simple formula (7) in the line tensions of the individual splitting components. This condition is only a criterion for stable Burgers vectors, and restricts little or not at all the possible types of dislocation.

A second stringent restriction on the Burgers vector is, that it must be either a lattice vector, or directly related to such a vector. The consequences of both conditions together have been worked out by FRANK and NICHOLAS [1953] and by THOMPSON [1953] for the body-centred and face-centred cubic and for the hexagonal close-packed lattice, and by HORNSTRA [1958] for the diamond lattice. Several complicated structures have also been investigated. In multi-atomic crystals additional conditions must be fulfilled, to prevent occurrence of configurations that disagree with the system of bonding forces (whatever that may be). Very severe restrictions can follow from the requirement of charge equilibrium, which arises in polar crystals. Each crystallographic structure has thus its own particular set of possible dislocations and general rules can not easily be laid down.

It is not the aim of this book to give many details on the geometry of dislocations, not even for mono-atomic lattices. For this see for instance the books of READ [1953] or J. M. and W. G. BURGERS [1956]. We only mention in the following Table 1 the most important types of Burgers vectors that can be encountered in practice in some simple lattices, as these follow from the work quoted above. The notation $a\,\langle i, j, k\rangle$ denotes a vector of length a, with (i, j, k) as crystallographic indices of its direction, the scale factor by which its indices must be multiplied to obtain the length of the vector, is put before the brackets.

TABLE 1

Burgers vectors of most important dislocations in some crystal lattices

Structure	*Perfect dislocations*	*Imperfect dislocations*
Simple cubic	$\langle 1,0,0\rangle$ $\langle 1,1,0\rangle$ $\langle 1,1,1\rangle$	none
Face-centred cubic	$\frac{1}{2}\langle 1,1,0\rangle$ $\langle 1,0,0\rangle$	$\frac{1}{6}\langle 1,1,2\rangle$ (H.-S. half disloc.) $\frac{1}{3}\langle 1,1,1\rangle$ (Frank sessile) $\frac{1}{6}\langle 1,1,0\rangle$ (Cottrell-Lomer disloc.) $\frac{1}{6}\langle 1,1,4\rangle$ (sessile) etc.
Hexagonal c.p.	$\frac{1}{3}\langle 1,1,\bar{2},0\rangle$ $\langle 0,0,0,1\rangle$ $\frac{1}{3}\langle 1,1,\bar{2},3\rangle$?	$\frac{1}{3}\langle 1,\bar{1},0,0\rangle$ (half disloc.) $\frac{1}{3}\langle 1,0,1,1\rangle$ etc.
Body-centred cubic	$\frac{1}{2}\langle 1,1,1\rangle$ $\langle 1,0,0\rangle$	$\frac{1}{6}\langle 1,1,1\rangle$ (twinning disloc.) $\frac{1}{3}\langle 1,1,1\rangle$,, ,, $\frac{1}{6}\langle 1,1,2\rangle$,, ,, etc.
Diamond	$\frac{1}{2}\langle 1,1,0\rangle$ $\langle 1,0,0\rangle$?	$\frac{1}{6}\langle 1,1,2\rangle$ (half disloc.) $\frac{1}{3}\langle 1,1,1\rangle$ (sessile) etc.
Sodiumchloride	$\langle 1,1,0\rangle$ $\langle 1,0,0\rangle$	
Caesium chloride	$\langle 1,0,0\rangle$ $\langle 1,1,0\rangle$? $\langle 1,1,1\rangle$?	

REFERENCES

AMELINCKX, S., W. BONTINCK, W. DEKEYSER and F. SEITZ (1957); Phil. Mag. **2**, 365.

AMELINCKX, S. (1958); Acta Metall. **6**, 34.

BILBY, B. A. (1950); Proc. phys. Soc. **A63**, 1911.

BILLIG, E. (1956); Proc. phys. Soc. **A235**, 37.

BUEREN, H. G. VAN (1957); Philips Research Reports **12**, 1.

BUEREN, H. G. VAN (1959); Physica (Haag) **25**, 720.

BURGERS, J. M. (1939); Proc. Acad. Sci. Amsterdam **42**, 293 and 278.

BURGERS, J. M. and W. G. BURGERS (1956); in: EIRICH, Rheology, p. 141 (Acad. Press, New York).

COTTRELL, A. H. (1952); Phil. Mag. **43**, 657.

COTTRELL, A. H. (1953); Dislocations in crystals (Oxford, University Press).

COTTRELL, A. H. and B. A. BILBY (1949); Proc. phys. Soc. **A62**, 49.
CRUSSARD, C. (1950); Métaux et corrosion **25**, 203.
DASH, W. C. (1956); J. applied Phys. **27**, 1193.
DIETZE, H.-D. (1952a); Z. Physik **131**, 157.
DIETZE, H.-D. (1952b); Z. Physik **132**, 107.
ESHELBY, J. D. (1949); Phil. Mag. **40**, 507.
ESHELBY, J. D. (1951); Phil. Transact. Roy. Soc. **A244**, 87.
ESHELBY, J. D. (1953); J. applied Phys. **24**, 176.
FOREMAN, A. J. (1955); Acta Metall. **3**, 322.
FOREMAN, A. J., M. A. JASWON and G. K. WOOD (1951); Proc. phys. Soc. **A64**, 156.
FRANK, F. C. (1950); Proc. phys. Soc. **A62**, 131.
FRANK, F. C. (1951); Phil. Mag. **42**, 809.
FRANK, F. C. (1956); Deformation and flow of solids, p. 73 (Springer, Berlin).
FRANK, F. C. and J. F. NICHOLAS (1953); Phil. Mag. **44**, 1213.
FRANK, F. C. and W. T. READ (1950); Phys. Rev. **79**, 722.
FRIEDEL, J. (1956); Les dislocations (Gauthier-Villars, Paris).
FRIEDEL, J. and P. COULOMB (1957); Dislocations and mechanical properties of crystals (Wiley, New York), p. 555.
GILMAN, J. J. and F. JOHNSTON (1959); J. appl. Physics **30**, 2.
HEIDENREICH, A. D. and W. SHOCKLEY (1947); Report Conf. Strength of Solids, Bristol, page 57 (The Phys. Soc., London).
HIRSCH, P. B. (1959); Internal Stress and Fatigue in Metals, p. 140 (Elsevier, Amsterdam).
HOLLOMON, J. (1954); Conf. Defects in Crystalline Solids, Bristol, motion picture.
HORNSTRA, J. (1958); J. Phys. Chem. Solids **5**, 129.
KIMURA, H., D. KUHLMANN-WILSDORF and R. MADDIN (1959); Acta Metall. **7**, 145.
KUHLMANN-WILSDORF, D., R. MADDIN and H. KIMURA (1958); Z. Metallk. **49**, 584.
KUHLMANN-WILSDORF, D. and H. WILSDORF (1954); Report Conf. Defects in Solids, Bristol (London, The Physical Society).
LEIBFRIED, G. (1950); Z. Physik **127**, 344.
LEIBFRIED, G. and H.-D. DIETZE (1951); Z. Physik **131**, 113.
LOMER, W. M. (1951); Phil. Mag. **42**, 1327.
MOTT, N. F. (1952); Phil. Mag. **43**, 1151.
MOTT, N. F. and F. R. N. NABARRO (1948); Rep. Conf. Strength of Solids, Bristol.
NABARRO, F. R. N. (1947); Proc. phys. Soc. **58**, 669.
NABARRO, F. R. N. (1951); Phil. Mag. **42**, 313.
NABARRO, F. R. N. (1952); Advances in Physics **1**, 269.
PEACH, M. (1951); J. applied Phys. **22**, 1359.
PEACH, M. and J. S. KOEHLER (1950); Phys. Rev. **80**, 436.
PEIERLS, R. (1940); Proc. phys. Soc. **52**, 34.
PENNING, P. (1958); Philips Res. Reports **13**, 79.
PENNING, P. (1959); Physica (Haag) **25**, 702.
READ, W. T. (1953); Dislocations in crystals (McGraw Hill, New York).
SEEGER, A. (1954); Z. Naturforschung **9a**, 758, 856 and 870.
SEEGER, A. (1955); Phil. Mag. **46**, 1194.
SEEGER, A. (1956a); Phil. Mag. **1**, 651.
SEEGER, A. (1956b); Encyclopedia of Physics, Vol. VII–1, Crystal Physics I, page 383.
SEEGER, A. and H. BLANK (1954); Z. Naturforschung **9a**, 252.
SEITZ, F. (1950); Phys. Rev. **79**, 529 and 723.
SEITZ, F. (1952); Advances in Physics **1**, 43.
TEGHTSOONIAN, E. and B. CHALMERS (1951); Canad. J. Phys. **29**, 370.
THOMPSON, N. (1953); Proc. phys. Soc. **B66**, 481.

CHAPTER IV

MATHEMATICAL THEORY OF LATTICE IMPERFECTIONS

1. Introduction

In the mathematical theory of elasticity, dislocations and point defects, or rather the elastic singularities that can be regarded as the prototypes of these crystalline imperfections, have been the subject of several classical investigations, dating from 1885 to 1910 (VOLTERRA [1907], LOVE [1927]). After this, interest in the subject dropped till around 1940, when the fundamental importance to crystal physics of these imperfections was realized, crystalline dislocations were "discovered" and the need for a detailed knowledge of their elastic properties arose. Concise accounts of the mathematical theory of imperfections, dealing with the ancient as well as with the modern developments, are scarce and difficult to read. We mention here those of NABARRO [1952] and SEEGER [1956]. ESHELBY [1956] and KRÖNER [1958] in his book have approached the subject from a very general point of view; we come back to this in section 5.

It was therefore deemed useful to include another, simply worded review in this book. The reader not interested in mathematical details can, however, skip this chapter without harm to the understanding of the rest of the work. By way of introduction we shall start with a short survey of the fundamentals of the mathematical theory of elasticity, and follow this by a treatment (largely after LOVE [1927]) of the properties of the simple elastic singularities in continuous media, that correspond to the elementary forms of crystalline point and line defects. After a short digression to singularities of a more general kind we shall finally discuss such properties of imperfections in discrete atomic lattices, as are connected with those of the corresponding elastic singularities in a continuum. We shall not deal with questions of energy of the imperfections (see, however, DE WIT [1958]).

2. Classical Theory of Elasticity

2.1. STRESS, STRAIN AND THEIR EQUATIONS

Consider an infinitesimal cube, oriented along the (orthogonal) coordinate axes, with edges dx, dy and dz (Fig. 1). For an arbitrary state of stress, there usually exists across each face a vector quantity T, with the dimension force per area, that is defined as the limit, for vanishing area, of the force, divided by the area, exerted by the space outside the surface on that inside. T is called the *traction* across the surface;

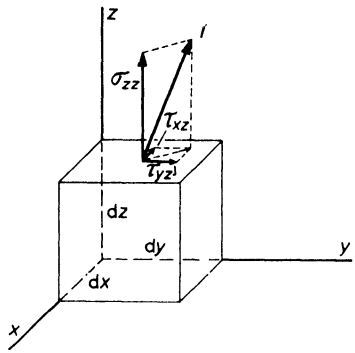


Fig. 1. Illustrating the definitions of traction T and stress components σ_{zz}, τ_{xz} and τ_{yz}.

it has three components T_x, T_y and T_z. In equilibrium, in first approximation T has equal magnitude and opposite sign on opposite faces. Considering the surfaces $dxdy$, the three traction components have been called the *shear stresses* τ_{xz}, τ_{yz} and the *normal stress* σ_{zz} respectively. Analogous definitions apply to the traction components and stress components across the other faces. The state of stress in the cube is fully determined by the nine stress components ("stresses"), σ_{xx}, σ_{yy}, σ_{zz} and τ_{xz}, τ_{yz}, τ_{xy}, τ_{zy}, τ_{yx}, τ_{zx}, which form a tensor of the second rank P, as becomes clear at once when the transformation properties are investigated. From the fundamental condition that in equilibrium the sum of the momenta of all surface tractions must be zero we need not concern ourselves about body forces in the limit of vanishing dimensions), it follows immediately that the tensor is symmetric. Thus only 6 stress components are distinct.

Consider now two neighbouring points P and Q within the cube. After application of an arbitrary stress the body will deform elastically, in such a way that P is displaced to P' and Q to Q' (Fig. 2). Both displacements can be described by a *displacement vector* with components

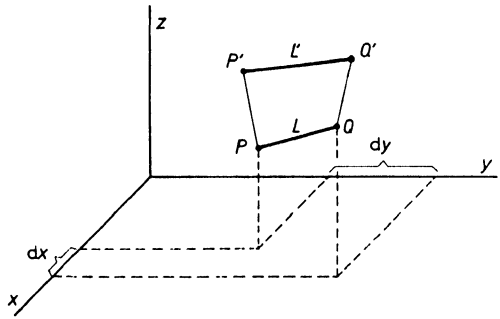


Fig. 2. The line segment PQ goes over into $P'Q'$ after deformation. The components of PP' are the displacements (u, v, w), the components of QQ' are the displacements (u_1, v_1, w_1).

("displacements") u, v, w for P (initial site x, y, z) and components u_1, v_1, w_1 for Q (initial site $x + \mathrm{d}x$, $y + \mathrm{d}y$, $z + \mathrm{d}z$). Then one has, presupposing the differentiability of the displacements:

$$\begin{aligned} u_1 &= u + \frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y + \frac{\partial u}{\partial z}\mathrm{d}z + \ldots\ldots, \\ v_1 &= v + \frac{\partial v}{\partial x}\mathrm{d}x + \frac{\partial v}{\partial y}\mathrm{d}y + \frac{\partial v}{\partial z}\mathrm{d}z + \ldots\ldots, \\ w_1 &= w + \frac{\partial w}{\partial x}\mathrm{d}x + \frac{\partial w}{\partial y}\mathrm{d}y + \frac{\partial w}{\partial z}\mathrm{d}z + \ldots\ldots. \end{aligned} \tag{1}$$

The distance between P and Q was before the deformation:

$$\mathrm{d}l = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2},$$

and is after the deformation:

$$\mathrm{d}l' = \sqrt{(\mathrm{d}x + u_1 - u)^2 + (\mathrm{d}y + v_1 - v)^2 + (\mathrm{d}z + w_1 - w)^2}.$$

In first order approximation we can write

$$dl' = \sqrt{(1 + 2\varepsilon_{xx})dx^2 + (1 + 2\varepsilon_{yy})dy^2 + (1 + 2\varepsilon_{zz})dz^2 + \\ + 2\gamma_{yz}dydz + 2\gamma_{zx}dzdx + 2\gamma_{xy}dxdy}\,. \tag{2}$$

Here we have introduced the *strain* components ("strains"):

$$\varepsilon_{xx} = \frac{\partial u}{\partial x}\,;\quad \varepsilon_{yy} = \frac{\partial v}{\partial y}\,;\quad \varepsilon_{zz} = \frac{\partial w}{\partial z}\,;$$
$$\gamma_{yz} = \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right)\,;\quad \gamma_{zx} = \left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right)\,;\quad \gamma_{xy} = \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)\,. \tag{3}$$

The ε denote the *normal* strains, the γ the *shear* strains. In the literature one often encounters other definitions, with factors 2 introduced or omitted; we have followed the convention used by Love in his book. From the transformation properties it follows again that the symmetric scheme E of the 9 strain components ε_{xx}, ε_{yy}, ε_{zz}, $\frac{1}{2}\gamma_{yz}$, $\frac{1}{2}\gamma_{zx}$, $\frac{1}{2}\gamma_{xy}$ + three additional components $\frac{1}{2}\gamma_{zy} = \frac{1}{2}\gamma_{yz}$, $\frac{1}{2}\gamma_{xz} = \frac{1}{2}\gamma_{zx}$ and $\frac{1}{2}\gamma_{yx} = \frac{1}{2}\gamma_{xy}$, has the characteristics of a tensor of the second rank.

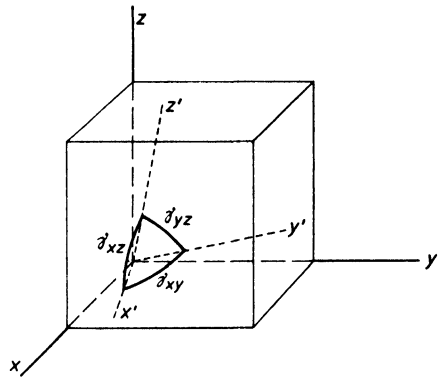


Fig. 3. The significance of the γ_{ij}: they are cosines of the new angles between the axes of the deformation product (dotted lines) of an originally orthogonal system of coordinate axes.

It is easily seen that ε_{xx}, ε_{yy} and ε_{zz} denote in first order approximation the components of the *elongation* of the line element PQ, whereas the γ_{ij} in this same approximation are the three cosines of the angles of mutual inclination of the axes of an, originally orthogonal, coordinate system through P (Fig. 3). Of course, the applicability of the

strain tensor (3) is limited to small strains only, because of this linear approximation (the *variation* of any strain component must everywhere be small compared to this component itself). There is no such restriction on the use of the stress tensor.

By differentiating (3), it is easily found that the various strain components must obey the so-called *compatibility equations*

$$2\frac{\partial^2\varepsilon_{xx}}{\partial y\,\partial z} = -\frac{\partial^2\gamma_{yz}}{\partial x^2} + \frac{\partial^2\gamma_{xz}}{\partial x\,\partial y} + \frac{\partial^2\gamma_{xy}}{\partial x\,\partial z}, \text{ and cycl.};$$
$$\frac{\partial^2\gamma_{yz}}{\partial y\,\partial z} = \frac{\partial^2\varepsilon_{yy}}{\partial z^2} + \frac{\partial^2\varepsilon_{zz}}{\partial y^2}, \text{ and cycl.} \tag{4}$$

For small strains it is found experimentally that *Hooke's law* applies, which in its generalized form can be expressed as follows: each component of the stress at any point of a body is a linear function of the components of the strain in that point. In a general anisotropic solid this linear interdependence is described by the 4th rank tensor of the elastic coefficients (Chapter I). We shall limit ourselves here to the case of an isotropic solid, in which case the *stress-strain-relations* reduce to the well-known form

$$\varepsilon_{xx} = \frac{1}{E}[\sigma_{xx} - \nu(\sigma_{yy} + \sigma_{zz})], \text{ and cycl.};$$
$$\gamma_{yz} = \frac{1}{G}\tau_{yz}, \text{ and cycl.} \tag{5}$$

Of the three quantities E, G and ν occurring in (5) only two are independent, since Young's modulus of elasticity E, the shear modulus or modulus of rigidity G and Poisson's ratio ν are connected by

$$G = \frac{E}{2(1+\nu)}. \tag{6}$$

Often other elastic coefficients are used, especially in more theoretical publications, the most important of which are *Lamé's coefficients* λ and μ, which are defined as

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}; \quad G = \mu; \quad \nu = \frac{\lambda}{2(\lambda + \mu)}. \tag{7}$$

In an isotropic solid with central forces acting between the atoms,

only one independent constant remains, as there $\nu = \frac{1}{4}$ ($\lambda = \mu$). In most materials $\nu \approx 0.3$.

For a body to be in equilibrium, certain conditions between the stresses must be fulfilled. Consider a prism centred at the origin, with edges $2a$, $2b$ and $2c$ along the coordinate axes. In the centre the stress components are σ^0_{xx}, etc. Then on the faces $x = \pm a$ one has:

$$\sigma_{xx} = \sigma^0_{xx} + \frac{\partial \sigma^0_{xx}}{\partial x} a + \frac{\partial \sigma^0_{xx}}{\partial y} b + \frac{\partial \sigma^0_{xx}}{\partial z} c + \ldots .$$

and analogous expressions for the other stress components. Let a volume force f (dimension dyne/cm^3) also act on the body (components in the centre f^0_x, etc., per unit volume). Then on the site x, y, z one has

$$f_x = f^0_x + \frac{\partial f_x}{\partial x} x + \frac{\partial f_x}{\partial y} y + \frac{\partial f_x}{\partial z} z + \ldots ., \text{ etc.}$$

For the body (as a whole) to be in equilibrium, the resultant force must vanish. This force has the components X, Y and Z, where

$$\begin{aligned} X = \iiint \left[f^0_x + \frac{\partial f_x}{\partial x} x + \frac{\partial f_x}{\partial y} y + \frac{\partial f_x}{\partial z} z \right] \mathrm{d}x\mathrm{d}y\mathrm{d}z + \\ + \iint [\sigma_{xx}(x = a) - \sigma_{xx}(x = -a)] \,\mathrm{d}y\mathrm{d}z + \\ + \iint [\tau_{xy}(y = b) - \tau_{xy}(y = -b)] \,\mathrm{d}z\mathrm{d}x + \\ + \iint [\tau_{xz}(z = c) - \tau_{xz}(z = -c)] \,\mathrm{d}x\mathrm{d}y . \end{aligned}$$

Similar formulae hold for Y and Z. By putting $X = Y = Z = 0$ and using the previous formula, one easily derives the following *equilibrium conditions* that must hold at each point of the body:

$$\begin{aligned} \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} + f_x = 0 , \\ \frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} + f_y = 0 , \\ \frac{\partial \tau_{zx}}{\partial x} + \frac{\partial \tau_{zy}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} + f_z = 0 ; \end{aligned} \tag{8}$$

or in tensor notation

$$\mathrm{Div}\, P + \mathbf{f} = 0 . \tag{8a}$$

A second condition for equilibrium, viz. the vanishing of the resultant *moment*, leads to the same equation, which is therefore a necessary and sufficient condition.

The equilibrium conditions can also be written in terms of the strains. Introducing the quantities

$$\textit{dilatation:}\quad e = \operatorname{div}(u, v, w) = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz} \tag{9}$$

and

$$\textit{rotation:}\quad \boldsymbol{\Omega} = \tfrac{1}{2} \operatorname{rot}(u, v, w)\,, \tag{10}$$

with the aid of (5) and (6) one arrives at

$$\frac{1-\nu}{1-2\nu}\frac{\partial e}{\partial x} + \left(\frac{\partial \Omega_z}{\partial y} - \frac{\partial \Omega_y}{\partial z}\right) + \frac{1}{2G} f_x = 0, \text{ and cycl.}, \tag{11}$$

or in vector notation

$$\frac{1-\nu}{1-2\nu} \operatorname{grad} e - \operatorname{rot} \boldsymbol{\Omega} + \frac{1}{2G} f = 0\,. \tag{11a}$$

2.2. DIFFERENTIAL EQUATIONS FOR STRESSES AND STRAINS

The 15 equations (3), (5) and (8) or (11) form a selfconsistent system (together with the appropriate boundary conditions) for the complete solution of the 15 quantities $u, v, w, \varepsilon_{xx}, \ldots, \gamma_{xy}, \ldots, \sigma_{xx}, \ldots, \tau_{xy}, \ldots$. By suitable differentiation and summation the equilibrium equations can be rewritten in terms of the strains as:

$$\begin{aligned} \Delta u + \frac{1}{1-2\nu}\frac{\partial e}{\partial x} &= 0\,; \\ \Delta v + \frac{1}{1-2\nu}\frac{\partial e}{\partial y} &= 0\,; \\ \Delta w + \frac{1}{1-2\nu}\frac{\partial e}{\partial z} &= 0\,. \end{aligned} \tag{12}$$

We take $f = 0$ (no body force) in this section.

For the stresses we thus obtain from the compatibility conditions

the *Beltrami equations:*

$$\Delta\sigma_{xx} + \frac{1}{1+\nu}\frac{\partial^2 s}{\partial x^2} = 0\,,$$

$$\Delta\sigma_{yy} + \frac{1}{1+\nu}\frac{\partial^2 s}{\partial y^2} = 0\,, \tag{13a}$$

$$\Delta\sigma_{zz} + \frac{1}{1+\nu}\frac{\partial^2 s}{\sigma z^2} = 0\,;$$

which by addition give

$$\Delta s = 0\,, \tag{13b}$$

where

$$s = \sigma_{xx} + \sigma_{yy} + \sigma_{zz}\,.$$

Further:

$$\Delta\tau_{yz} + \frac{1}{1+\nu}\frac{\partial^2 s}{\partial y\,\partial z} = 0\,,$$

$$\Delta\tau_{zx} + \frac{1}{1+\nu}\frac{\partial^2 s}{\partial z\,\partial x} = 0\,, \tag{13c}$$

$$\Delta\tau_{xy} + \frac{1}{1+\nu}\frac{\partial^2 s}{\partial x\,\partial y} = 0\,.$$

As boundary conditions, usually the surface displacements (u, v, w) or the surface tractions T are prescribed in each point of the surface of the body. If the unit vector n denotes the outside normal to the surface, the *boundary conditions* read

$$P \cdot n = T\,, \tag{14}$$

where P denotes the stress tensor at the surface point. Written in component language:

$$\sigma_{xx}\cos\alpha + \tau_{xy}\cos\beta + \tau_{xz}\cos\gamma = T_x\,,$$

$$\tau_{yx}\cos\alpha + \sigma_{yy}\cos\beta + \tau_{yz}\cos\gamma = T_y\,, \tag{14a}$$

$$\tau_{zx}\cos\alpha + \tau_{zy}\cos\beta + \sigma_{zz}\cos\gamma = T_z\,;$$

where $\cos\alpha$, $\cos\beta$ and $\cos\gamma$ are the direction cosines of the normal.

Very often these surface tractions are simply zero (stress-free surface).

2.3. TWO-DIMENSIONAL PROBLEMS

When the state of deformation depends on only two of the three coordinates, say on x and y, the equilibrium conditions (8), in the absence of body forces, become simply

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} = 0 ,$$

$$\frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} = 0 ,$$

$$\frac{\partial \tau_{zx}}{\partial x} + \frac{\partial \tau_{zy}}{\partial y} = 0 .$$

We can then introduce two *stress functions* F and χ in such a way that

$$\sigma_{xx} = \frac{\partial^2 F}{\partial y^2} , \quad \sigma_{yy} = \frac{\partial^2 F}{\partial x^2} , \quad \tau_{xy} = -\frac{\partial^2 F}{\partial x \partial y} ;$$
$$\tau_{xz} = \frac{\partial \chi}{\partial y} , \quad \tau_{yz} = -\frac{\partial \chi}{\partial x} . \tag{15}$$

It can be shown that *all* solutions of the equilibrium equations are covered by this symbolism.

If moreover $\tau_{xz} = \tau_{yz} = 0$, only the function F needs to be retained. In this case one is interested in a problem in *plane strain*. The Beltrami equations (13a) become

$$\Delta' \sigma_{xx} + \frac{1}{1+\nu} \frac{\partial^2 s'}{\partial x^2} + \frac{1}{1+\nu} \frac{\partial^2 \sigma_{zz}}{\partial x^2} = 0 ,$$

$$\Delta' \sigma_{yy} + \frac{1}{1+\nu} \frac{\partial^2 s'}{\partial y^2} + \frac{1}{1+\nu} \frac{\partial^2 \sigma_{zz}}{\partial y^2} = 0 ,$$

$$\Delta' \sigma_{zz} = 0 ;$$

where Δ' stands for

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

and $s' = \sigma_{xx} + \sigma_{yy}$.

By addition of the first two equations one has, together with the third equation:

$$\Delta'(\sigma_{xx} + \sigma_{yy}) = 0$$

and together with (15):

$$\Delta'\Delta'F = 0\,. \tag{16}$$

This is the Airy bipotential equation for plane strain; F is the Airy stress function. It does not contain direct information about σ_{zz}; however, it is easily found from (16) that under the conditions pertaining to plane strain, automatically

$$\sigma_{zz} = \nu(\sigma_{xx} + \sigma_{yy}), \qquad \text{thus} \qquad \varepsilon_{zz} = 0\,.$$

In this case the boundary conditions (14a) can also be written directly in terms of F. Consider a cylindrical body with the z-axis as axis. Let l be the length of the circumference of an (x, y) cross-section as measured from an arbitrary origin on the boundary. Then

$$\cos\alpha = \frac{\partial y}{\partial l}\,, \qquad \cos\beta = -\frac{\partial x}{\partial l}$$

and we have with (15) as conditions for the Airy function F_b at the boundary:

$$\frac{\partial^2 F_b}{\partial x\,\partial y}\frac{\partial y}{\partial l} + \frac{\partial^2 F_b}{\partial x^2}\frac{\partial x}{\partial l} = \frac{\partial}{\partial l}\left(\frac{\partial F_b}{\partial x}\right) = -T_y\,,$$

$$\frac{\partial^2 F_b}{\partial y^2}\frac{\partial y}{\partial l} + \frac{\partial^2 F_b}{\partial x\,\partial y}\frac{\partial x}{\partial l} = \frac{\partial}{\partial l}\left(\frac{\partial F_b}{\partial y}\right) = T_x\,.$$

Hence

$$\frac{\partial F_b}{\partial x} = -\int_0^l T_y\,dl + c_1; \qquad \frac{\partial F_b}{\partial y} = \int_0^l T_x\,dl + c_2\,. \tag{17}$$

The two-dimensional problem complementary to plane strain occurs if the *only* existing stress components are τ_{xz} and τ_{yz}. This situation is sometimes referred to as *anti-plane strain*. Now the stress function χ of (15) describes the problem. It follows from (15) that

$$\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} = 0\,. \tag{18}$$

The Beltrami equations (13c) reduce to

$$\Delta' \tau_{xz} = 0 \; ; \qquad \Delta' \tau_{yz} = 0 \, . \tag{19}$$

By differentiating (18) with respect to x and to y respectively, and subtracting from the equations (19), one finds

$$\frac{\partial \tau_{xz}}{\partial y} - \frac{\partial \tau_{yz}}{\partial x} = c \, , \tag{20}$$

where c is a constant that depends neither on x nor on y. Together with (18) and (15) one obtains finally for anti-plane strain the potential equation

$$\Delta' \chi = - c \, , \tag{21}$$

with c depending on the nature of the problem. The boundary condition for χ is easily found from the request that the resulting shear stress at the boundary must everywhere be tangential to it. As

$$\mathrm{d}\chi = - \tau_{xz} \mathrm{d}y + \tau_{yz} \mathrm{d}x \, ,$$

one has $\mathrm{d}\chi_b = 0$, or $\chi_b = \text{const}$. As only the derivatives of χ are of interest to us, we can put *this* constant equal to zero:

$$\chi_b = 0 \, . \tag{22}$$

This is *not* allowed in multiply connected bodies, where χ can be put equal to zero at *one* of the boundary surfaces only. Otherwise (22) would imply a condition for the derivatives of χ also.

3. Singularities in Three-dimensional Strain

In practice, singularities in the strain distribution are usually discussed by considering them due to the action of *singular forces*. We use this term to indicate forces that are neither volume nor surface forces, as they act only in *isolated* points, lines or areas. In these regions the compatibility equations do no longer apply and divergencies occur. We shall discuss here only the case of a force operating in one point, and allied problems. We start from the equilibrium equations (which retain their validity everywhere), written in the form (11a). We, quite generally, express a body force f (per unit volume) as the sum of a scalar and a vector potential:

$$f = - \operatorname{grad} \Phi + \operatorname{rot} \Psi \, , \tag{23}$$

and can then put (11a) in the form:

$$\frac{1-\nu}{1-2\nu}\operatorname{grad} e - \operatorname{rot}\boldsymbol{\Omega} = \frac{1}{2G}\operatorname{grad}\Phi - \frac{1}{2G}\operatorname{rot}\boldsymbol{\Psi}\,. \tag{24}$$

A particular set of solutions is obtained by making the straightforward assumptions:

$$\begin{aligned}\frac{1-\nu}{1-2\nu}e &= \frac{1}{2G}\Phi\,,\\ \boldsymbol{\Omega} &= \frac{1}{2G}\boldsymbol{\Psi}\,.\end{aligned} \tag{25}$$

It is known from potential theory that equation (23) can be inversed to

$$\Phi(r) = -\frac{1}{4\pi}\iiint f(r')\cdot\operatorname{grad}\frac{1}{|r-r'|}\,d\tau'\,,$$

$$\boldsymbol{\Psi}(r) = \frac{1}{4\pi}\iiint f(r')\times\operatorname{grad}\frac{1}{|r-r'|}\,d\tau'\,;$$

integrated over the part of space where f is defined. Here r and r' stand for the radius vector. When f is defined only in a small element of volume $d\tau'$ around the origin, and this volume approaches zero in such a way that

$$\lim_{d\tau'\to 0}\iiint f\,d\tau = f_0$$

(dimension of a force) remains finite, then:

$$\Phi(r) = \frac{1}{4\pi}\frac{f_0\cdot r}{r^3}\,,$$

$$\boldsymbol{\Psi}(r) = -\frac{1}{4\pi}\frac{f_0\times r}{r^3}\,.$$

thus

$$\begin{aligned}e &= \frac{1-2\nu}{1-\nu}\frac{1}{8\pi G}\frac{f_0\cdot r}{r^3}\,;\\ \boldsymbol{\Omega} &= \frac{1-2\nu}{1-\nu}\frac{1}{8\pi G}\frac{f_0\times r}{r^3}\,.\end{aligned} \tag{26}$$

The displacement vector itself is then easily found to be

$$(u, v, w) = \frac{1}{16\pi G}\left\{\frac{3-4\nu}{1-\nu}\frac{\mathbf{f}_0}{r} + \frac{1}{1-\nu}\frac{(\mathbf{f}_0 \cdot \mathbf{r})\mathbf{r}}{r^3}\right\} + \text{const. vector}. \quad (27)$$

This formula describes the displacements in a body where a force $\mathbf{f}_0$ operates *in one point only* (viz. in the origin).

We shall now consider some special situations that can be described with the aid of formula (27).

a) *Force $\mathbf{f}_0$ along x-axis, operating in origin.*

The displacements are

$$u_{0x} = \frac{1}{16\pi G}\frac{1}{1-\nu} f_0 \left(\frac{x^2}{r^3} + (3-4\nu)\frac{1}{r}\right),$$

$$v_{0x} = \frac{1}{16\pi G}\frac{1}{1-\nu} f_0 \frac{xy}{r^3}, \quad (28)$$

$$w_{0x} = \frac{1}{16\pi G}\frac{1}{1-\nu} f_0 \frac{xz}{r^3}.$$

b) *Two equal, oppositely directed forces $\mathbf{f}_0 = \mathbf{g}_0/h$ along the x-axis, operating in the points (0, 0, 0) and (h, 0, 0), in the limiting case that $h \to 0$, thus $h\mathbf{f}_0 \to \mathbf{g}_0$. (Double force without moment, of strength $\mathbf{g}_0$.)*

One has

$$u_{dx} = \lim_{h\to 0} h \cdot \frac{\delta u_{0x}}{h} = \frac{1}{16\pi G}\frac{1}{1-\nu} g_0 \left\{(4\nu - 1)\frac{x}{r^3} - \frac{3x^3}{r^5}\right\},$$

$$v_{dx} = \lim_{h\to 0} h \cdot \frac{\delta v_{0x}}{h} = \frac{1}{16\pi G}\frac{1}{1-\nu} g_0 \left(\frac{y}{r^3} - \frac{3x^2 y}{r^5}\right), \quad (29)$$

$$w_{dx} = \lim_{h\to 0} h \cdot \frac{\delta w_{0x}}{h} = \frac{1}{16\pi G}\frac{1}{1-\nu} g_0 \left(\frac{z}{r^3} - \frac{3x^2 z}{r^5}\right).$$

c) *Three double forces without moment, each of strength $\mathbf{g}_0$, operating in the origin along the three coordinate axes. (Centre of compression or dilatation.)*

One has simply $u_c = u_{dx} + u_{dy} + u_{dz}$, etc., where u_{dy} denotes the x-component of the displacement due to a double force without moment along the y-axis, etc. By straightforward addition one finds:

$$u_c = \frac{g_0}{8\pi G}\frac{1-2\nu}{1-\nu}\frac{x}{r^3},$$

$$v_c = \frac{g_0}{8\pi G}\frac{1-2\nu}{1-\nu}\frac{y}{r^3}, \tag{30a}$$

$$w_c = \frac{g_0}{8\pi G}\frac{1-2\nu}{1-\nu}\frac{z}{r^3},$$

or

$$(u, v, w) = -\frac{g}{8\pi G}\frac{1-2\nu}{1-\nu}\operatorname{grad}\left(\frac{1}{r}\right). \tag{30b}$$

These expressions could also have been obtained by noting, firstly, that:

$$u_c \to h\left(\frac{\partial u_{0x}}{\partial x} + \frac{\partial u_{0y}}{\partial y} + \frac{\partial u_{0z}}{\partial z}\right), \text{ and cycl.},$$

and secondly (from (27) or (28)), that

$$u_{0z} = w_{0x}; \qquad u_{0y} = v_{0x}; \qquad v_{0z} = w_{0y}.$$

Thus:

$$u_c \to h\left(\frac{\partial u_{0x}}{\partial x} + \frac{\partial v_{0x}}{\partial y} + \frac{\partial w_{0x}}{\partial z}\right) = h\,(\mathrm{dil})_x, \text{ etc.} \tag{31}$$

where $(\mathrm{dil})_{x,y,z}$ stands for the dilatations associated with forces along the (x, y, z) axes respectively, operating in the origin. These dilatations are, according to (26), given by

$$(\mathrm{dil})_{x,y,z} = \frac{1}{8\pi G}\frac{1-2\nu}{1-\nu} f_0 \cdot \frac{(x, y, z)}{r^3},$$

from which formula (30) follows immediately. In polar coordinates the displacement components become

$$u_{cr} = \frac{g_0}{8\pi G}\frac{1-2\nu}{1-\nu}\frac{1}{r^2},$$

$$u_{c\varphi} = u_{c\theta} = 0. \tag{32}$$

Thus only a radial displacement exists, which is moreover independent of φ and θ. On a sphere around the centre the surface traction is easily found to be radially directed.

Stress and strain around the centre decrease inversely proportionally to the third power of the distance:

$$\varepsilon_{rr} = -2\varepsilon_{\varphi\varphi} = -2\varepsilon_{\theta\theta} = \frac{g_0}{8\pi G}\frac{1-2\nu}{1-\nu}\frac{1}{r^3};$$
$$\sigma_{rr} = -2\sigma_{\varphi\varphi} = -2\sigma_{\theta\theta} = \frac{g_0}{\pi}\frac{1-2\nu}{1-\nu}\frac{1}{r^3}. \tag{33}$$

All other components vanish.

Obviously, this state of strain is the elastic prototype of the foreign atom or of the vacancy in a crystal, and in first approximation also of the interstitial, although the latter usually has a strongly anisotropic stress field around it (KRÖNER [1958], section 31). In a free body of finite dimensions, (30) does not apply, as ESHELBY [1954] has pointed out. The condition of zero surface tractions necessitates the addition of an image-stress system to (33). The change of *volume* of an infinite solid owing to the displacement equations (30) is

$$\Delta V = \frac{g_0}{2G}\frac{1-2\nu}{1-\nu}. \tag{34}$$

The radius r_0 of a hypothetical hole around the defect has clearly expanded by the amount $(g_0/8\pi G) \cdot (1-2\nu)/(1-\nu) \cdot (1/r_0^2)$.

This volume change also describes that of the solid as a whole, since the dilatation, being the divergence of grad $(1/r)$, is everywhere zero.

The image stresses, now, *increase* this volume change by the constant factor

$$\gamma = 3\frac{1-\nu}{1+\nu}. \tag{35}$$

In crystals the centre can for instance be formed by replacing an atom with radius r_0 by a foreign atom with a radius $(1+\varepsilon)$ times as large. Then

$$\Delta V = \tfrac{4}{3}\pi\varepsilon r_0^3.$$

As follows from (34) corrected by the factor γ, we can in this case replace the strength factor g_0 by

$$g_0 = \frac{8\pi}{9} G \frac{1+\nu}{1-2\nu}\varepsilon r_0^3.$$

In many cases the distortion is unsymmetrical, and the elastic situation could be better described by the superposition of three mutually perpendicular unequal double forces without moment.

The origin in the situations *a*, *b* and *c* is a singularity, in so far as the displacements and also the stresses and strains become infinite there. These regions must thus be excluded from detailed consideration. That is, the singularity must be placed in a hole. At a sufficient distance from the hole, stresses, strains and displacements are single-valued, continuous and differentiable functions, and can be manipulated without any difficulty.

This is not so when the analogous singularities in two-dimensional strain are considered. This we shall do now.

4. Singularities in Two-dimensional Strain

In principle, in the two-dimensional case we could proceed as in the foregoing section, and again discuss first the case of a two-dimensional force operating in a point (resp. a three-dimensional force operating in a line). The vector formalism, however, does not lead us to some of the special features we want to discuss here, and we shall follow Love in adopting a more elegant and powerful method for two-dimensional problems. Let us first consider problems of *plane strain*.

4.1. PLANE STRAIN

In the absence of body forces the equilibrium conditions (11) read

$$\begin{aligned} \frac{1-\nu}{1-2\nu}\frac{\partial e}{\partial x} + \frac{\partial \Omega}{\partial y} &= 0 \ ; \\ \frac{1-\nu}{1-2\nu}\frac{\partial e}{\partial y} - \frac{\partial \Omega}{\partial x} &= 0 \ ; \end{aligned} \tag{37}$$

as only one component of $\boldsymbol{\Omega}$, viz. $\Omega = \Omega_z$ exists. These equations express the fact that the complex function

$$\Sigma = \frac{1-\nu}{1-2\nu} e + i\Omega$$

is an analytic function of $x + iy$. This includes, of course, the corollaries

$$\Delta' e = 0 ,$$

$$\Delta' \Omega = 0 ,$$

Now, we take for Σ a function with a simple pole at the origin:

$$\Sigma = \frac{A}{x + iy} . \tag{38}$$

Putting $\varrho = \sqrt{x^2 + y^2}$; $\theta = \arctan y/x$, we have:

$$e = \frac{1 - 2\nu}{1 - \nu} A \frac{x}{\varrho^2} ,$$

$$\Omega = - A \frac{y}{\varrho^2} .$$

From this again it results after some computation that the displacement components can be written as

$$\begin{aligned} u &= A \log \varrho + \frac{1}{2(1-\nu)} A \frac{y^2}{\varrho^2} + \frac{\partial p}{\partial x} , \\ v &= \frac{1 - 2\nu}{2(1-\nu)} A\theta - \frac{1}{2(1-\nu)} A \frac{xy}{\varrho^2} + \frac{\partial p}{\partial y} ; \end{aligned} \tag{39}$$

where p is still an arbitrary harmonic function of x and y (thus $\Delta' p = 0$).

The peculiar property of the expression (39) for v is that it is *multivalued*. On describing a closed contour around the origin, the displacement in the y-direction increases by the cyclic factor

$$(1 - 2\nu)/(1 - \nu) \cdot \pi A .$$

Besides, the component u of the displacement is logarithmically divergent in the origin. This difficulty does not much trouble us, as we only need to place the origin in a small cavity to evade it. In order to obtain a physically fully sound solution we might, however, try to remove the multivaluedness also, by suitably adjusting the function p. Then we must have

$$\frac{\partial p}{\partial y} = - \frac{1 - 2\nu}{2(1-\nu)} A\theta ,$$

and by reason of the harmonicity of ϕ,

$$\frac{\partial \phi}{\partial x} = \frac{1-2\nu}{2(1-\nu)} A \log \varrho .$$

We then find for the "corrected" components of the displacement:

$$\begin{aligned} u' &= \frac{3-4\nu}{2(1-\nu)} A \log \varrho + \frac{1}{2(1-\nu)} A \frac{y^2}{\varrho^2}, \\ v' &= -\frac{1}{2(1-\nu)} A \frac{xy}{\varrho^2}. \end{aligned} \tag{40}$$

Computation of the stresses from (40) with the aid of the stress-strain formulae (3) and (5), and subsequent evaluation of the resultant tractions across the surface of the small cavity around the origin, reveals that this traction is a force of magnitude $2\pi A$ acting in the direction of the negative x-axis. We thus have here the two-dimensional analogue of problem (*a*) of section 3, a force operating at a point. From it we can, of course, again construct a centre of compression or dilatation.

Another multivalued displacement function is obtained by taking for Σ a function with a *logarithmic* singularity in the origin:

$$\Sigma = B \log (x + iy) . \tag{41}$$

Then one obtains after some computation:

$$\begin{aligned} u &= B(x \log \varrho - x) - \frac{3-2\nu}{2(1-\nu)} By\theta + \frac{\partial \phi}{\partial x}, \\ v &= \frac{1-2\nu}{2(1-\nu)} B(x\theta - y) - \frac{\nu}{1-\nu} By \log \varrho + \frac{\partial \phi}{\partial y} . \end{aligned} \tag{42}$$

These components of displacement are both multivalued and now it is *not* possible to make them both single-valued by adjusting the function ϕ.

Poles of higher order in Σ and other kinds of singularities we shall not discuss here; they lead to stresses and strains that diverge strongly at infinity.

We have removed the multivaluedness of (39) by suitably adapting the auxiliary function ϕ. Is it really necessary to do this in order to

obtain a solution pertaining to a physical situation? In view of the necessity always to exclude a small cavity around the origin from our considerations, it is to *multiply connected* bodies that two-dimensional deformation functions with the described kind of singularities apply. For instance, they can apply to a straight cylinder with the z-axis as axis and with a central bore. In such multiply connected bodies, however, the condition of single-valuedness of the components of displacement is not stringent. Consider a hollow cylinder (Fig. 4) from which a

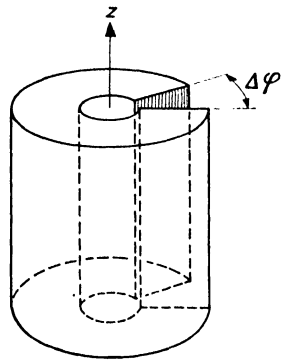


Fig. 4. Illustration of the logarithmic singularity (42).

thin sector of material, bounded by two axial planes inclined at an angle $\Delta\varphi$, has been removed, and which afterwards has again been closed. This cylinder is in a state of *initial stress* or selfstress, actually characterized by a multivalued tangential component of the displacement vector. The residue is $\varrho\Delta\varphi$ where ϱ is the distance from the axis. This is physically a quite acceptable situation. Everywhere the strains are single-valued, and the stresses fully determined. In fact, one can show that the singularity responsible for this state of stress is the same as that of (42), with

$$B = \Delta\varphi/2\pi\,.$$

The pole-singularity leading to (39) can also be interpreted as belonging to a certain state of deformation of a hollow cylinder. Putting p in (39) equal to zero (as it is now no longer necessary to remove the multivaluedness), it follows that the displacement v in the y-direction increases, with the residue $(1 - 2\nu)/(1 - \nu) \cdot A\pi$, each time a complete

contour around the cylinder axis is described. The residue, which we shall call b in the following, is now evidently independent of ϱ. The strain situation can again be realized by removing from the hollow cylinder of Fig. 4 a slice, now not bounded by *axial* planes, but by two *parallel* planes perpendicular to the y-axis and spaced a distance b, and then joining the faces of the cut together again (Fig. 5).

The deformations, illustrated in Figs. 4 and 5, have been called *distorsioni* by Volterra and *dislocations* by Love. The second type,

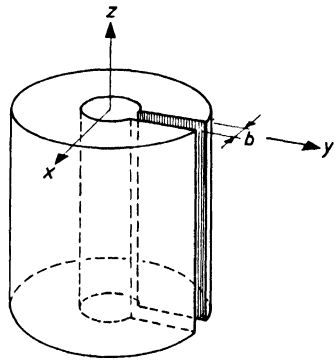


Fig. 5. Illustration of the pole singularity (38).
The designations x and y must be interchanged.

characterized by a *constant translational strength b*, was later called the *edge dislocation*. Although its displacement vector is logarithmically divergent at infinity, the strains, as derivatives of the displacement, remain regular. The same edge dislocation can also be obtained by cutting the cylinder wall not parallel to the x-axis, but by making some other radial fissure and displacing the sides of this cut by the same amount b in the same (y) direction, removing or adding material, and joining the faces together again. A particular case arises when the cut is parallel to the y-axis; then the cut surfaces need only to be sheared, and *no* material needs to be removed or added. An edge dislocation can then be characterized either by a radial shift of the cylinder parts along a cut made parallel to its "strength vector", or by the insertion of a slice of new material into a radial fissure perpendicular to the strength vector. Obviously, the orientation of the cut is immaterial.

This is not true for the dislocation arising from the logarithmic

singularity (41), the *rotational* dislocation as it is sometimes called. From the discussion of Fig. 4 it follows that here the orientation of the cut determines the values of the residue in the x and y coordinates; furthermore these residues are proportional to the radial distance from the axis of the cylinder, and the *strain* components (not only the displacements) diverge logarithmically at infinity. The latter two properties make this dislocation of little interest to crystal physics (see, however, Ch. XIX). The edge dislocation, on the contrary, is of great practical importance, even because of its peculiar simple properties.

In a cylinder of finite dimensions, the surface tractions due to the stresses around the edge dislocation, as these can be derived from (39) by putting $p = 0$, do not vanish. On the end planes σ_{zz} produces a normal pressure or tension $\{2\nu/(1 - 2\nu)\} \cdot (Gbx/\varrho^2)$, and on the cylindrical surfaces normal pressures of magnitude $\{2/(1 - 2\nu)\}. (Gbx/\varrho^2)$ occur. In order to obtain the stresses belonging to a dislocation in a free, *finite*, cylindrical body, we need to superpose on the displacement field (39) another, single-valued displacement, that is accompanied by stresses which, on the surface, cancel the tractions mentioned. It depends on the shape of the body what these additional displacements must be, and it is largely a matter of guesswork how to obtain correct expressions. Therefore, in the early literature different authors proposed different expressions for the actual stress function of an edge dislocation (TAYLOR [1934], KOEHLER [1941], J. M. BURGERS [1939]). In the following, a general approach to the problem will be discussed (NABARRO [1952]).

With the help of the knowledge we already possess about the typical behaviour of the stresses around an edge dislocation, we can try to find a general expression for an Airy stress function from which these stresses in a finite medium can be derived under the condition that the *resultant* traction on the whole surface is zero. The typical behaviour of the stresses can be expressed in the symbolic formula $\sigma = Af(\theta)/\varrho$, where $f(\theta)$ is a dimensionless function of the azimuth only, and A is a numerical factor of dimensions dyne/cm. The stress components being second derivations of the stress function, it is seen that the latter must have the form:

$$F = (Bx + Cy) \log (\varrho/c) + (Dx + Ey)\theta \,, \tag{43}$$

where c is an arbitrary constant length and the constants B, C, D and

E have again the dimension dyne/cm. From the boundary conditions (17) it follows by inversion, that the *resultant* traction on any closed boundary around the origin is given by (in the notation of (17)):

$$T_{ry} = -\oint \frac{\mathrm{d}}{\mathrm{d}l}\left(\frac{\partial F_b}{\partial x}\right)\mathrm{d}l \quad \text{and} \quad T_{rx} = \oint \frac{\mathrm{d}}{\mathrm{d}l}\left(\frac{\partial F_b}{\partial y}\right)\mathrm{d}l\,,$$

where F_b denotes the value of F of (43) at the boundary. To these circuit integrals only those terms of F will contribute that have multivalued first derivatives; that is, only the second term of F contributes to the resultant traction. As the latter must vanish everywhere on the boundary, both D and E must be zero. By a simple rotation of the coordinate axes in the (x, y) plane we can further let one of the constants B and C disappear, say C, and we are left with the stress function

$$F = Bx \log (\varrho/c),$$

describing the stresses of an edge dislocation along the z-axis.

By differentiating we find that v becomes indeed again a multivalued function, with residue $b = 2\pi(1-\nu)B/G$. Expressed in the strength b of the dislocation, we can write

$$F = \frac{Gbx}{2\pi(1-\nu)} \log (\varrho/b)\,, \tag{44}$$

where we have put $c = b$, as this choice for the arbitrary constant is now indicated. Equation (44) is actually the same as the stress function derived by J. M. BURGERS [1939] in a different way (viz. by investigating the analogy with a hydrodynamical vortex line). From (44) can be derived the following expressions for the components of stress and displacements of an edge dislocation in an infinite medium:

$$\begin{aligned}
\sigma_{xx} &= -\frac{Gb}{2\pi(1-\nu)} \frac{x(x^2-y^2)}{\varrho^4}\,;\\
\sigma_{yy} &= \frac{Gb}{2\pi(1-\nu)} \frac{x(x^2+3y^2)}{\varrho^4}\,;\\
\sigma_{zz} &= \frac{2\nu Gb}{\pi(1-\nu)} \frac{xy^2}{\varrho^4}\,;
\end{aligned} \tag{45}$$

$$\tau_{xy} = \frac{Gb}{2\pi(1-\nu)} \frac{y(x^2-y^2)}{\varrho^4};$$

$$\tau_{xz} = \tau_{yz} = 0.$$

And

$$u = \frac{1-2\nu}{4\pi(1-\nu)} b \log(\varrho/b) - \frac{1}{4\pi(1-\nu)} b \frac{x^2}{\varrho^2},$$

$$v = -\frac{1}{4\pi(1-\nu)} b \frac{xy}{\varrho^2} + \frac{b}{2\pi} \arctan\left(\frac{y}{x}\right). \tag{46}$$

4.2. ANTI-PLANE STRAIN

We have yet to inquire whether similar elastic singularities exist also in anti-plane strain. This case is much simpler in principle, as the only existing stress components are τ_{xz} and τ_{yz}. We first look for a simple complex analytic function of the stresses, and find such a function to be

$$\Theta = t_x + it_y, \tag{47}$$

where

$$t_x = \tau_{yz},$$

and

$$t_y = \tau_{xz} - cy,$$

with c an arbitrary constant so far. Then according to (18):

$$\frac{\partial t_x}{\partial x} = \frac{\partial t_y}{\partial y}; \qquad \frac{\partial t_x}{\partial y} = -\frac{\partial t_y}{\partial x}.$$

Writing

$$\Theta = \frac{A}{x+iy},$$

thus assigning to Θ a simple pole in the origin, we have

$$\tau_{xz} = -\frac{Ay}{x^2+y^2} + cy;$$

$$\tau_{yz} = \frac{Ax}{x^2+y^2}. \tag{48}$$

As

$$\tau_{xz} = G\gamma_{xz} = G\frac{\partial w}{\partial x},$$

and

$$\tau_{yz} = G\frac{\partial w}{\partial y},$$

when (48) holds, only then a displacement function w can be found if $c = 0$. In this case we derive

$$w = -\frac{A}{G}\arctan\frac{y}{x}, \tag{49}$$

thus again a multivalued displacement component, with residue $b = -2\pi A/G$. The strain components are

$$\gamma_{xz} = \frac{b}{2\pi}\frac{y}{\varrho^2},$$

$$\gamma_{yz} = -\frac{b}{2\pi}\frac{x}{\varrho^2}. \tag{50}$$

The state of deformation is comparable to that of the edge dislocation because again stress and strain are inversely proportional to the

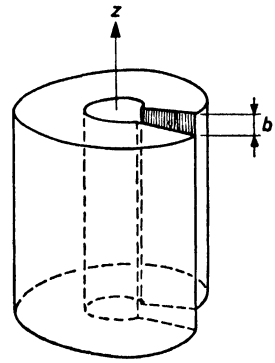


Fig. 6. Illustration of the anti-plane strain singularity (49).

distance. This type of dislocation is the *screw* dislocation. It can be visualized by making a radial cut bounded by two parallel planes in a

cylinder mantle (as for the edge dislocation), but now shear subsequently the two cut surfaces over a relative distance b *in* the direction of the axis before joining them together again (Fig. 6). Again, the azimuthal position of the cut is immaterial.

The logarithmic singularity in plane strain also has its counterpart in anti-plane strain. It corresponds to a multivalued displacement function with variable residue and a logarithmically diverging strain component at infinity.

5. Continuum Theory of Dislocations

5.1. DISLOCATIONS AS SOURCES OF INTERNAL STRESS

Dislocations have in common with so-called thermal strains (which are the result of inhomogeneous thermal expansion) that they can be considered as *sources of initial or internal stress* (ESHELBY [1956]). The strains differ from those caused by an external stress in that the compatibility equations (4) are not everywhere obeyed, because at the sources of internal stress these conditions break down.

To describe the state of deformation of a material, two lines of exposition can in principle be followed, according to the choice of the equilibrium equations (8) or the compatibility equations (4) as a starting point. The introduction of *displacement functions* (u, v, w) which obey the latter equations leads naturally to the adoption of the equilibrium conditions as fundamental to the definition of the stress. This has the advantage that the presence of an external force f can easily be taken into account (cf. equation (8)); however, if the compatibility equations are *not* everywhere fulfilled, this method is less suitable. It is therefore that in the previous sections the complementary approach via the introduction of *stress functions* was used. The stresses, of course, per definition obey the equilibrium conditions, and the compatibility equations, eventually with their singularities included, give us now the displacements.

It has been tried to replace the ad hoc method for finding solutions to the Airy equation, which has been used in section 4, by a more inductive method. For this it was necessary to translate into the mathematical language of the compatibility equations the presence of strain singularities like dislocations. The first approach to this problem has been made by NYE [1953]; BILBY, BULLOUGH and SMITH [1955] and KRÖNER [1955] derived the fundamental equation (54) (see below)

that underlies the modern continuum theory of dislocations, as this has been extensively developed afterwards (see ESHELBY [1956], KRÖNER [1956, 1958]). In this section we shall outline Kröner's method as it was laid down in his book (KRÖNER [1958]); for a description with more emphasis laid on the work of the english authors the reader is referred to ESHELBY's review paper [1956].

Like Kröner, we shall use tensor notation—as we have already done occasionally before, for instance in equation 8a. The intricacies of this notation can be found in many textbooks (see e.g. MORSE and FESHBACH [1953]); we only mention those basic definitions we shall need in this section. Let γ be a system of nine components

$$\gamma \equiv \begin{vmatrix} \gamma_{xx} & \gamma_{xy} & \gamma_{xz} \\ \gamma_{yx} & \gamma_{yy} & \gamma_{yz} \\ \gamma_{zx} & \gamma_{zy} & \gamma_{zz} \end{vmatrix} ;$$

when these obey certain transformation rules they are said to form a tensor (dyadic) of the second rank. *Scalar multiplication* by an arbitrary vector V results in another vector

$$\left.\begin{aligned} W &= \gamma \cdot V = \sum_{ij} \gamma_{ij} V_i a_j \\ W' &= V \cdot \gamma = \sum_{ij} \gamma_{ji} V_i a_j = \gamma^* \cdot V \end{aligned}\right\} i, j = x, y, z ,$$

where $\gamma^*_{ij} = \gamma_{ji}$ denote the components of the conjugate tensor and $a_{i,j,k}$ are the unit vectors in the (Cartesian) system of coordinates. Vectorial multiplication by a vector V results in a tensor, with components:

$$(\gamma \times V)_{kj} = \left[\sum_l (\gamma_{jl} V_{ll}) a_k - \sum_l (\gamma_{kl} V_l) a_j\right].$$

Any tensor field of the second rank can be symbolically represented in terms of a vector F

$$\gamma = a_x \frac{\partial F}{\partial x} + a_y \frac{\partial F}{\partial y} + a_z \frac{\partial F}{\partial z},$$

where

$$\gamma_{ij} = \frac{\partial F_j}{\partial x} .$$

Replacing in the above formulae the arbitrary vector V by the operator ∇ with components

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right),$$

we obtain

$$\nabla \cdot \gamma \equiv \mathrm{Div}\, \gamma = \gamma^* \cdot \nabla .$$

$$(\nabla \times \gamma) \equiv \mathrm{Rot}\, \gamma = -(\gamma^* \times \nabla) .$$

The capitals in the symbols Div and Rot denote that these operations differ from the well-known div and rot operations working on vectors.

By converting the differential relations to integral forms in the usual way, we obtain the extended theorems of Gauss:

$$\iint \gamma \cdot \mathrm{d}A = \iiint \mathrm{Div}\, \gamma \cdot \mathrm{d}v ,$$

and of Stokes:

$$\oint \gamma \cdot \mathrm{d}s = \iint \mathrm{Rot}\, \gamma \cdot \mathrm{d}A ,$$

where $\mathrm{d}s$, $\mathrm{d}A$ and $\mathrm{d}v$ denote line, area and volume element respectively.

5.2. DESCRIPTION OF IMPERFECTIONS

Kröner introduced a *distortion tensor* β describing the state of deformation of a body under internal stress. It is defined as follows: first the body is imagined to be divided into small separate volume elements, held in their deformed state by (imaginary) surface tractions. Then these tractions are removed and the elements relax by an amount which is opposite to the original *distortion*. The total distortion of a solid is made up out of an elastic part β_{E} and a plastic part β_{P}, due to dislocations:

$$\beta = \beta_{\mathrm{E}} + \beta_{\mathrm{P}} ; \tag{51}$$

apart from the strains the quantities β also contain (lattice) rotations. The compatibility equations (4) express the condition that the body

in its deformed state still hangs together everywhere. In tensor notation they can simply be rewritten as

$$\text{Rot } \beta = 0 . \tag{52}$$

Writing

$$\text{Rot } \beta_{\text{E}} = \alpha , \tag{53}$$

we have

$$\alpha = - \text{Rot } \beta_{\text{P}} . \tag{54}$$

The tensor α thus defined is connected with the *distribution* of dislocations in the solid; it is zero if no dislocations are present (no plastic component of the distortion). By definition

$$\text{Div } \alpha = 0 \tag{55}$$

(because Div Rot $\equiv 0$).

It can be shown that α describes the *planar density of dislocations* according to

$$\alpha_{ij} = \lim \Delta b_j / \Delta F_i , \tag{56}$$

where ΔF_i denotes an infinitesimal planar element of area perpendicular to the i-direction, and Δb_j is the component in the j-direction of the resultant Burgers vector of all dislocations passing through this element. In other words, α_{ij} denotes the density of dislocations lying along direction i and with resultant Burgers vector along direction j. The proof is simple. One has, according to the extension of Stokes' theorem, for the total displacement due to dislocations:

$$\boldsymbol{b} = - \int \text{Rot } \beta_{\text{P}} \cdot \mathrm{d}\boldsymbol{F} ,$$

where $\boldsymbol{b}$ stands for the vector sum of the Burgers vectors of all dislocations passing through an arbitrary plane F (dislocation flux). By inserting (54) one obtains

$$\boldsymbol{b} = \int \alpha \cdot \mathrm{d}\boldsymbol{F} ,$$

which by inversion yields (56). Obviously equation (55) expresses the fact that dislocation lines cannot end in the interior.

Formula (54) constitutes the fundamental equation of the continuum theory of dislocations; it was first derived (in the integral form) by NYE [1953]. The close similarity to the mathematical description of a dislocated lattice with the aid of a Burgers circuit will be clear.

The above general treatment is of especial value for the description of *dislocation distributions*.

Kröner proceeded by defining an *incompatibility tensor* as the *symmetric* part of the tensor

$$\alpha \times \nabla = - \operatorname{Rot} \alpha^* . \tag{57}$$

This new tensor is denoted by η, thus

$$\eta = - \operatorname{Sym} (\operatorname{Rot} \alpha^*); \tag{58}$$

it is fully determined by the dislocation density. The compatibility equations can now be written as:

$$\eta = 0;$$

when they are fulfilled everywhere, there are no sources of internal stress. In this way η can be regarded as the *source function* of the dislocation stresses.

Together with the equilibrium equations for the stresses written in the form (8a):

$$\operatorname{Div} \sigma = 0 \tag{59}$$

(no external forces), the "*extended Beltrami equations*" can now be derived which replace the classical Beltrami equations (13) when internal stresses are present. They are:

$$\begin{aligned} \Delta\sigma_{ii} &= 2G\eta_{ii} ; \\ \Delta\tau_{ij} + \frac{1}{1+\nu} \frac{\partial^2 s}{\partial x_i \partial x_j} &= 2G\eta_{ij} . \end{aligned} \tag{60}$$

An identical solution of (59) is obtained by putting

$$\sigma = \nabla \times \chi \times \nabla \equiv \operatorname{Ink} \chi , \tag{61}$$

because one can show that $\operatorname{Div} \operatorname{Ink} \equiv 0$.

The operation Ink is defined by equation (61). The tensor of stress functions χ can then be shown to obey a set of *bipotential equations* that replace the Airy equation (16):

$$\begin{aligned} \Delta\Delta\chi_{ij} &= 2G\eta_{ij} ; \\ \Delta\Delta\chi_{ii} &= \frac{1+2\nu}{1+\nu} \cdot 2G\eta_{ii} . \end{aligned} \tag{62}$$

The classical formalism based on the Beltrami equations can thus be taken over without much change, and leads to logically defined stress functions, in which η, thus α, determines the solution.

The strength of any formalism lies in its applications. In Kröner's book a number of applications is extensively discussed. We mention only the computation of the stresses of an arbitrary dislocation and of the interaction between two dislocations in terms of simple line integrals of the functions χ or η. Other applications are discussed by ESHELBY [1956].

A very important group of applications concerns *point defects*. These can also be considered as elastic singularities as we have seen in section 3, and they can be characterized by the force they exert on their surroundings. As in section 3, three *double forces* in mutually orthogonal directions form components of the symmetric tensor P of the *stress dipole* associated with the defect. The dilatation centre discussed in that section was characterized by $P_{11} = P_{22} = P_{33}$; $P_{ij} = 0$ $(i \neq j)$. However, the generalization to unequal components is easily made. For the P_{ij} simple rules hold. A field of elastic deformation (strain tensor ε) exerts a force with components

$$K_k = \sum_{ij} (\mathrm{grad}_k\, \varepsilon_{ij})\, P_{ij}\,, \tag{63}$$

on the point defect, and by integration one finds for the potential energy of the latter in that strain field

$$U = -\sum_{ij} P_{ij}\, \varepsilon_{ij}\,. \tag{64}$$

The formal analogy to the force exerted by an electric field on an electric dipole is clear. If the strain field is caused by a dislocation, (63) takes account of the interaction between the latter and the point defect. An extension to the case where the defect causes a tetragonal distortion ($P_{11} = P_{22} \neq P_{33}$) is easily made. The classical example is formed by the interaction between a carbon atom and a dislocation in α-iron. This was first studied by COTTRELL and BILBY [1949], who assumed purely hydrostatic stresses around the foreign atom ($P_{11} = P_{22} = P_{33}$); the interaction with a screw dislocation was therefore found zero. Many authors have pointed out that this assumption is too restricted and the tetragonality of the deformation around the C-atom should be considered. The above formalism is then of much use More will be said about this in Chapter IX.

6. Classification of Dislocations in Elastic Continua

6.1. SOMIGLIANA AND VOLTERRA DISLOCATIONS

From the deductive approach to the line singularities in elastic continua that was followed in section 4, it is difficult to conclude in how far the treatment is complete. A more general discussion of the physically possible singularities is thus needed.

Consider a doubly connected body (a body with a toroidal hole in it), in which a cut is made that ends on the hole. Both cut surfaces are elastically deformed and displaced, material is added or removed if necessary, and the surfaces are joined together again. In this way a dislocation of the most general kind, a so-called *Somigliana dislocation* with the hole as its axis, is produced. Of course, the components of stress and strain need now not be continuous across the cut, or single valued in the material. Only the traction normal to the cut surface must be continuous everywhere, in order to ensure equilibrium. A Somigliana dislocation is characterized by a displacement vector $\boldsymbol{U}$ of the cut surfaces that is an arbitrary function of the position on these surfaces.

When the condition is posed that *all* stress components *and* their first derivatives must be *continuous* across the cut and *single-valued* in the material (the displacements itself may yet be multivalued), only *rigid displacements* of the cut surfaces are allowed, as WEINGARTEN [1901] has shown. In this case one speaks of *Volterra dislocations* (VOLTERRA [1907]). All the dislocations discussed in section 4 belong to this class. The most general displacement vector $\boldsymbol{U}$ of the cut surfaces has now the form:

$$\boldsymbol{U} = \boldsymbol{b} + \boldsymbol{d} \times \boldsymbol{r}, \tag{65}$$

where $\boldsymbol{b}$ and $\boldsymbol{d}$ are *constant* vectors, and $\boldsymbol{r}$ is the position vector on the cut surface.

For the edge dislocation of section 4.1, the vector $\boldsymbol{b}$ is perpendicular to the axis of the dislocation (that is, to the toroidal hole, considered as straight), whereas $\boldsymbol{d} = 0$. The screw dislocation of section 4.2 corresponds to $\boldsymbol{b}$ parallel to the axis and $\boldsymbol{d} = 0$. We shall further denote $\boldsymbol{b}$ (formerly called strength of the dislocation) by the term *Burgers vector*. The rotational dislocation (42) of section 4.1 corresponds to the case where $\boldsymbol{b} = 0$ and $\boldsymbol{d}$ is a constant vector directed along the axis.

From (65) it follows that there are six elementary types of Volterra

dislocations, three translational and three rotational, corresponding to the six components of the two vectors $\boldsymbol{b}$ and $\boldsymbol{d}$. These are depicted in Fig. 7. Two of the translational dislocations (*b* and *c*) are edge dislocations of the same type, be it with mutually perpendicular Burgers vectors. They are fully equivalent singularities. The third (*d*) is the screw dislocation. The rotational dislocation (42) is illustrated in Fig. 7*g*.

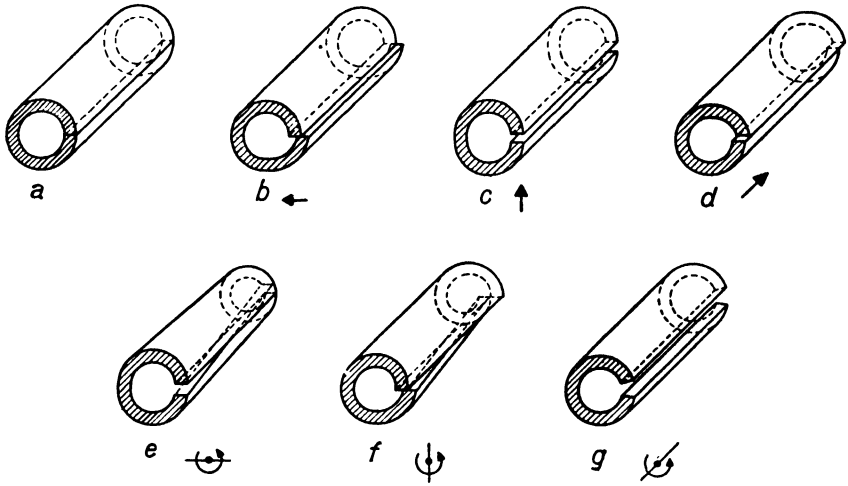


Fig. 7. The six types of Volterra dislocations. *a*) Undistorted cylinder wall, already cut through at one side. *b*) and *c*) edge dislocations, *d*) screw dislocation. Together these form the three translational dislocations. *e*), *f*) and *g*) are the three rotational dislocations, which are not characterized by a constant Burgers vector (after NABARRO [1952]).

Any system of arbitrary dislocations (such as e.g. the loops of a Frank-Read source, section III-6) can be described mathematically as one Somigliana dislocation with stepwise varying strength. This possibility is sometimes exploited to deal with specific dislocation problems, such as that of the *core* of a Volterra dislocation, where stresses and strains tend to infinity. One then replaces the Volterra dislocation of Burgers vector $\boldsymbol{b}$ by a Somigliana dislocation with a displacement vector that is equal to $\boldsymbol{b}$ over most of the cut surfaces, but that tapers off to zero near the cut boundary, that is, the dislocation core (ESHELBY [1949]). A similar treatment is applied when dealing with crystalline dislocations. In the previous section we have already outlined a more modern approach to problems of this kind.

Centres of compression and dilatation can also be regarded as

Somigliana dislocations; then the cut surface is a sphere, and U is normal to its surface, in such a way that as r tends to zero, U tends to infinity proportionally to $1/r$.

In the following we shall, when not indicated otherwise, use the word dislocation to signify only Volterra dislocations of the three (better two) translational kinds, or a mixture of these two. In other words, we shall only be interested in dislocations with a constant Burgers vector.

A full description of the properties of the elastic dislocations would carry us too far (see also MANN [1949]). We shall discuss in sections 7 and 8 two aspects of practical interest, viz. the concept of a force on a dislocation, and the relation between dislocations in elastic continua and those in crystalline solids.

7. Force on a Dislocation

The concept of *force* on a dislocation is defined mathematically in the following manner (NABARRO [1952]). Let a segment of dislocation move the infinitesimal distance $\mathrm{d}\boldsymbol{d}$ under the action of an arbitrary stress tensor. Then the energy E of the complete system as formed by the medium containing the dislocation *and* the sources of the applied stress, changes by, say, an amount $\mathrm{d}E$. The force $\boldsymbol{F}$ on the dislocation segment is now understood as the vector quantity

$$\boldsymbol{F} = \mathrm{grad}_{\boldsymbol{d}}\, E. \tag{66}$$

Starting from this definition, we can easily express the force by which a short element of dislocation $\mathrm{d}\boldsymbol{l}$ with Burgers vector $\boldsymbol{b}$ is acted

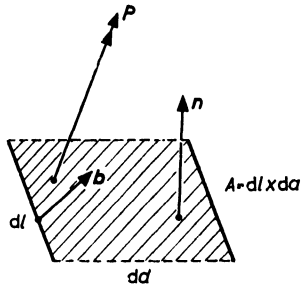


Fig. 8. Illustrating the argument of section 7.

upon by a local stress P (Fig. 8). Let again the segment move the distance $\mathrm{d}\boldsymbol{d}$. The area covered is than $\mathrm{d}\boldsymbol{l} \times \mathrm{d}\boldsymbol{d}$, and the traction

across it is $P \cdot (\mathrm{d}l \times \mathrm{d}d)$. The two sides of this element of area have been relatively displaced by the amount b, and the work done by the stress is therefore

$$\mathrm{d}E = b \cdot \{P \cdot (\mathrm{d}l \times \mathrm{d}d)\} = (b \cdot P) \cdot (\mathrm{d}l \times \mathrm{d}d) = (b \cdot P \times \mathrm{d}l) \cdot \mathrm{d}d\,.$$

Then the force is

$$F = \mathrm{grad}_d E = (b \cdot P \times \mathrm{d}l). \tag{67}$$

Of greatest practical interest is the component of force, f, that acts *in* the glide plane of the dislocation, that is the plane defined by the dislocation segment itself and the Burgers vector. The unit vector normal to this plane is

$$n = \frac{\mathrm{d}l \times b}{|\,\mathrm{d}l \times b\,|}\,.$$

Then:

$$f = F - n(n \cdot F) = -\,n \times (n \times F)\,,$$

and

$$n \times F = n \times \{(b \cdot P) \times \mathrm{d}l\} = (b \cdot P)\,(n \cdot \mathrm{d}l) - \mathrm{d}l\,\{(b \cdot P) \cdot n\} =$$

$$= -\,\mathrm{d}l\,(bPn)\,,$$

thus

$$f = (n \times \mathrm{d}l)\,(bPn)\,. \tag{68}$$

$(b\,Pn)$ is the shear stress component acting in the direction of the Burgers vector. The force per unit length of a dislocation in the glide plane is thus (of course) directed perpendicularly to the axis of the dislocation, and is equal in magnitude to *b times the resolved shear stress along the Burgers vector.* This holds for a stationary dislocation.

8. Defects in Crystalline Solids

8.1. POINT DEFECTS

The approximation of a point defect in an actual crystal by a centre of compression or of dilatation is necessarily very rough, as no account is taken of the discrete atomic structure of the solid. Owing to the rapid variation $(1/r^3)$ with distance of the strains, the part of the crystal within one or two atomic distances from the missing or superfluous atom will be strained most (*core* of the defect), and the continuous approximation is certainly not warranted. Many authors (ESHELBY

[1958]) have tried to replace the elastic theory by using the so-called *atomic* approximation, in which the energy of the crystal containing the defect is minimized in terms of the displacements of the individual atoms surrounding it. Owing to the paucity of knowledge about the interatomic forces, the results obtained by various investigators differ considerably, in the magnitude of the deformation derived as well as in its anisotropy characteristics. In view of this fact, any deepgoing discussion of the problem does not fall within the purview of this book; we only mention the work of KANZAKI [1957] and that of HALL [1957]. The latter includes a valuable comparison between the atomic and the elastic approximation, although the atomic interaction potential chosen, the Lennard-Jones potential, does not apply to the crystals we are mostly interested in, namely metals and ionic crystals. One of Hall's results, namely that the deformation around a vacancy is highly anisotropic even in a cubic crystal, has, however, a much wider significance.

Hall divides the atom positions around the defect into shells, at distances $d\sqrt{k}$ from the defect, where d is the distance to the nearest neighbour atom and k runs from 1 to any whole number. Assuming *central* forces to act and all displacements to be purely radial, he now minimizes these radial displacements of each shell with respect to the interaction energy. Using a certain Lennard-Jones type interaction potential, he finds for the displacement of the first 8 shells around a vacancy in a particular face-centred cubic crystal:

$$\begin{aligned} \delta_1 &= -\ 8.39 \times 10^{-3} & \delta_5 &= +\ 0.33 \times 10^{-3} \\ \delta_2 &= +\ 2.54 & \delta_6 &= +\ 0.02 \\ \delta_3 &= -\ 0.31 & \delta_7 &= -\ 0.02 \\ \delta_4 &= -\ 0.83 & \delta_8 &= 0 \end{aligned}$$

where a minus sign denotes *inward* motion of the shell. The first shell (nearest neighbours) thus moves inward, the second shell outward, the third and fourth inward, etc. As the first shell is made up of atoms situated in ⟨110⟩ directions as seen from the vacancy, the second shell by atoms situated in cube directions and the third shell by atoms along the body diagonal, and higher shells do not contribute much to the displacement, the result of the computation is that the *net* displacements are *out*ward along the cube axes and *in*ward along the other directions. There exists, therefore, a very considerable anisotropy in the

displacement (Fig. 9*a*). The same is true for an interstitial atom (of the same species as the other atoms) in the f.c.c. lattice, and for some other cases illustrated in Fig. 9. Hall points out that the anisotropy will be especially strong in those cases where the two opposite displacements of the first and second shell aid each other (as in Figs. 9*b* and 9*c*).

The displacement around the interstitial atom is about 6 times greater (in the first shell) than around a vacancy in the face-centred cubic lattice. This is not surprising of course. The energy stored in the distortion of the lattice is therefore only considerable for the interstitial, and can be neglected in the vacancy case as compared to the actual energy of formation (see section II–2.1).

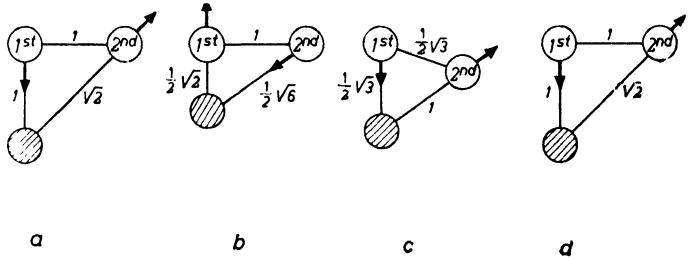


Fig. 9. Movement of neighbours and next-neighbours respectively, on the introduction of *a*) a vacancy in an f.c.c. lattice, *b*) an interstitial in a tetrahedral position of an f.c.c. lattice, *c*) a vacancy in a b.c.c. lattice; *d*) a vacancy in an h.c.p. lattice. In the situations *b*) and *c*) the displacements aid each other, and the expected anisotropy is greatest (after HALL [1957]).

At distances larger than those of the second-nearest neighbours, the continuous elasticity-theoretical approximation describes very well the average radial displacements of the shells. As the appropriate shear modulus G in (36) one should take $\frac{1}{5}(3c_{11} + 2c_{12} + 4c_{44})$ in a cubic crystal. Of course, the conventional theory does not lead to anisotropy, and anisotropic elasticity theory should be applied. In the first instance this can be done by introducing the stress dipole P_{ij} of section 5.

A problem of great practical importance to which the continuum theory of elasticity is eminently applicable, has already been discussed in section 5. It is that of the *interaction* between point defects mutually and between line and point defects. One aspect of this bears directly on the anisotropy mentioned: when many point defects are present (locally), they may with advantage arrange themselves regularly over

the lattice, so that the anisotropic strains of the one compensate to a large extent those of the other, and the crystal structure as such is but little affected. This is a short-range interaction, of course. Elastic long-range interaction is not very important, owing to the rapidly decreasing stresses; its sign depends on the nature of the interacting defects (ESHELBY [1958]). A treatment of short-range interaction has been given by ZENER [1948], following the lines of conventional elasticity theory, and has recently been elaborated upon by FISHER [1958]. The problem is of special interest for some internal friction effects; we refer to the original papers for further details. As pointed out by several authors, other contributions to the long-range interaction, arising from chemical and electrical causes, are often much more important than the elastic interaction as such. These depend, of course, above all on the type of solid considered, and a general discussion here is not feasible.

8.2. DISLOCATIONS

In the theory of dislocations of an elastic continuum, the axis of the dislocation is placed in a hollow tube, in order to make the solid multiply connected and to evade divergencies. In actual crystals containing dislocations, there is no question of the removal of any material. The stresses and strains must be computed up to distances from the mathematical axis that are of the same order as the Burgers vector (viz. one atomic distance). There the strains and stresses have the order of magnitude $1/2\pi$ and $G/2\pi$ respectively, which is far beyond the limits where Hooke's law of linear elasticity applies. Although the theory of elasticity is much better applicable to dislocations than to vacancies, owing to the longer range of the stress field, for the core also here the atomic approximation deserves preference.

Another peculiar feature of a crystalline solid is, that the forces (shear stress) opposing a relative displacement of two atom rows do not increase monotonically with the displacement but are *periodic*, with the atomic distance as period. It has been tried repeatedly and with appreciable success to calculate the effect of this periodicity on the stresses and strains around a crystal dislocation. The starting point for such a calculation is usually the approximation of PEIERLS [1940]. In this approximation the material on both sides of the slip plane is considered as an elastic continuum, and the opposing shear stress is

taken to be a periodic function of the local relative displacement of the two crystal halves (section III-5.1).

Consider an edge dislocation along the y-axis, placed symmetrically with regard to the atom sites as in Fig. 10. On the glide plane $z = 0$, the Burgers vector is directed along the x-axis and is equal in magnitude to the atomic spacing b in the x-direction. Let the atomic spacing

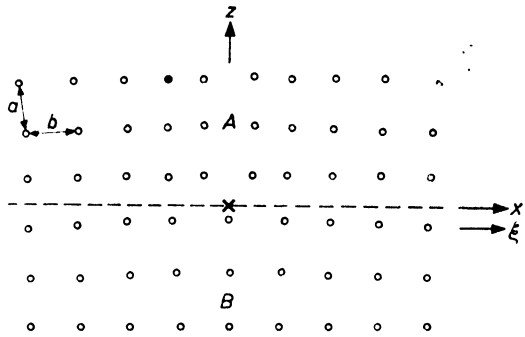


Fig. 10. Illustrating the derivation of the Peierls-Nabarro equations.

in the z-direction be a. The material above the glide plane is denoted by A, that below the glide plane by B. Two forces act on the atoms of A just above the glide plane. The elastic tensions in A tend to *in*crease the diminished local atomic spacing along x, the attraction from B tries to *de*crease it still further. The first effect can be described as follows. One can regard the dislocation (ESHELBY [1949]) as a continuous distribution along the x-axis, of infinitesimal edge dislocations of unit Burgers vector, with density $b'(\xi)\mathrm{d}\xi$ (thus as a Somigliana dislocation of varying strength $b'(\xi)$). Per interval $\mathrm{d}\xi$ these dislocations produce a shear stress at the point $(x,0)$ (compare (45)):

$$\mathrm{d}\tau^{A}_{xz} = \frac{G}{2\pi(1-\nu)} \frac{b'\mathrm{d}\xi}{\xi - x}.$$

Let the displacements of an atom of A be (u, w), and those of an atom of B $(\bar{u}, \bar{w})$. By suitably choosing the origins of these displacements ($\frac{1}{2}b$ distant), we can make $\bar{u} = -u$, $\bar{w} = w$, thus the relative displacement is $\Delta u = 2u$, and one has $b'(\xi) = \mathrm{d}(\Delta u)/\mathrm{d}\xi = 2\mathrm{d}u/\mathrm{d}\xi$. Thus for the total shear stress at the point $(x,0)$ one finds by integrating

$$\tau^{A}_{xz} = \frac{G}{\pi(1-\nu)} \int_{-\infty}^{+\infty} \frac{\mathrm{d}u(\xi)}{\mathrm{d}\xi} \frac{\mathrm{d}\xi}{\xi - x}. \tag{69}$$

The attractive force due to the material B is in the Peierls approximation simply expressed by a periodic shear stress

$$\tau_{xz}^{B} = \frac{Gb}{2\pi a} \sin \frac{2\pi \Delta u(x)}{b}, \tag{70}$$

where the factor before the sinus follows from the fact that Hooke's law must be obeyed for small relative displacements. This law reads:

$$\tau_{xz} = G\Delta u/a ,$$

which is indeed the limiting form of (70). Equating (69) and (70), one obtains the *integral equation of Peierls* for the displacements $u(x)$ in the glide plane, due to an edge dislocation:

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}u}{\mathrm{d}\xi} \frac{\mathrm{d}\xi}{x - \xi} = (1 - \nu) \frac{b}{2a} \sin \frac{4\pi u(x)}{b} . \tag{71}$$

Its solution corresponding to the problem of Fig. 10 is

$$u = -\frac{b}{2\pi} \operatorname{arc\,tan} \frac{x}{\zeta} , \tag{72}$$

where

$$\zeta = \frac{a}{2(1 - \nu)} \tag{73}$$

is the so-called *width* of the dislocation. For $\nu = \frac{1}{3}$ it is only $\frac{3}{4}a$ and thus very small, which makes the application of the Peierls approximation actually a matter of doubt. DIETZE [1952a], using a modified and extended Peierls' model for face-centred cubic metals, found ζ to vary from $0.50a$ for an edge dislocation along ⟨110⟩, to $1.17a$ for a 30°-half-dislocation along ⟨110⟩. The simple form of expression (73) for ζ follows immediately from the adoption of the sinusoidal relation (70). FOREMAN, JASWON and WOOD [1951] used other, more flat-topped periodic functions for τ_{xz} (but left the model as such unaltered); in this way they obtained a much larger width (together with a more complex formula (73)).

From the known value of u the displacement field at an arbitrary site can be computed (LEIBFRIED and LÜCKE [1949]), and from it the complete stress field due to a crystal dislocation. This can be considered as due principally to an "elastic dislocation", which however must then be situated at $z = -\frac{1}{2}\zeta$ for the material A, and at $z = +\frac{1}{2}\zeta$ for B.

NABARRO [1952] has extended Peierls' treatment, taking the discrete

atomic structure of the parts A and B partly into account. It then follows naturally that the strain energy of a dislocation depends on its site relative to the lattice planes. Let the position of the axis be given by the distance αb from the nearest symmetry position. Then the displacements of the various atom rows labelled by the index n of A and B are, according to (72), given by

$$u_n = -\frac{b}{2\pi} \arctan \frac{(\alpha + \frac{1}{2} + n)b}{\zeta}, \qquad n = 0, \pm 1, \pm 2 . \quad (74)$$

Now, the energy of misfit of any row is given by

$$E_n = -\int b\tau \mathrm{d}u ,$$

and according to (70) this is equal to

$$E_n = \int_0^u \frac{Gb^2}{2\pi a} \sin \frac{4\pi u'}{b} \mathrm{d}u' = \frac{Gb^3}{8\pi^2 a} \left[1 - \cos \frac{4\pi u}{b}\right] . \quad (75)$$

Replacing u of (75) by u_n of (74), and carrying out the summation over all n by aid of a Fourier transform method (see e.g. COTTRELL [1953]), we obtain for the strain energy per unit length of the dislocation in first approximation:

$$E = \frac{Gb^3}{2\pi(1-\nu)a} \mathrm{e}^{-4\pi\zeta/b} \cos 4\pi\alpha . \quad (76)$$

which is indeed a periodic function of the dislocation position. By taking $-(\partial E/\partial\alpha)$ we find the so-called *Peierls force*, opposing the glide motion of the dislocation in the crystal lattice:

$$F_p = \frac{2Gb^2}{(1-\nu)a} \mathrm{e}^{-4\pi\zeta/b} \sin 4\pi\alpha . \quad (77)$$

DIETZE [1952a] has given more detailed formulae for the Peierls force pertaining to the various types of dislocation in the f.c.c. lattice. All these formulae depend very sensitively upon ζ which can not be very accurately computed (unless one takes into account all possible variations of the form of the shear stress function (70) and also the influence of temperature, DIETZE [1952b]). Therefore, the Peierls force cannot be predicted within a factor of 10^4.

KUHLMANN-WILSDORF [1960] has criticized the validity of the estimate of F_p found by differentiating the core-energy of a dislocation; she pointed out that, whereas this energy is already very inaccurately

known, small variations in it are still less reliable, numerically. By regarding instead stress and strain on the slip plane and neighbouring planes, she obtained a much higher frictional force than before. A substantial reduction owing to the uncertainty in the position of the dislocation axis because of the *thermal vibrations* of the atoms must be expected, however. This reduction is the smaller, the more rigid the atomic bonding forces are, as in the valence crystals, where the Peierls force is indeed large (see Part III of this book).

REFERENCES

BILBY, B. A., R. BULLOUGH and E. SMITH (1955); Proc. Roy. Soc. **A231**, 263.
BURGERS, J. M. (1939); Proc. Acad. Sci. Amsterdam **42**, 293 and 278.
COTTRELL, A. H. (1953); Dislocations and Flow of Crystals (The Clarendon Press, Oxford).
COTTRELL, A. H. and B. A. BILBY (1949); Proc. phys. Soc. **A62**, 49.
DE WIT, R. (1958); Thesis Illinois University.
DIETZE, H. D. (1952a); Z. Physik **131**, 156.
DIETZE, H. D. (1952b); Z. Physik **132**, 107.
ESHELBY, J. D. (1949); Phil. Mag. **40**, 903.
ESHELBY, J. D. (1954); J. appl. Phys. **25**, 255.
ESHELBY, J. D. (1956); Solid State Physics **3**, 79.
ESHELBY, J. D. (1958); Ann. der Physik **1**, 116.
FISHER, J. C. (1958); Acta Metall. **6**, 13.
FOREMAN, A. J., M. A. JASWON and J. K. WOOD (1951); Proc. phys. Soc. **A64**, 156.
HALL, G. L. (1957); J. Phys. Chem. Solids **3**, 210.
KANZAKI, H. (1957); J. Phys. Chem. of Solids **2**, 24.
KOEHLER, J. S. (1941); Phys. Rev. **60**, 397.
KRÖNER, E. (1955); Z. f. angew. Physik **7**, 249; Z. Physik **141**, 386 and **142**, 463.
KRÖNER, E. (1956); Z. Naturforschung **11a**, 969.
KRÖNER, E. (1958); Kontinuumstheorie der Versetzungen und Eigenspannungen (Continuum theory of dislocations and initial stress) (Springer, Berlin).
KUHLMANN-WILSDORF, D. (1960); Phys. Rev. **120**, 773.
LEIBFRIED, G. and K. LÜCKE (1949); Z. Physik **126**, 450.
LEIBFRIED, G. and H. D. DIETZE (1951); Z. Physik **131**, 113.
LOVE, A. E. H. (1927); The Mathematical Theory of Elasticity (Cambridge University Press).
MANN, E. H. (1949); Proc. Roy. Soc. **A199**, 376.
MORSE, P. M. and H. FESHBACH (1953); Methods of Theoretical Physics, p. 44 (McGraw-Hill, New York).
NABARRO, F. R. N. (1952); Advances in Physics **1**, 269.
NYE, J. F. (1953); Acta Metall. **1**, 153.
PEIERLS, R. E. (1940); Proc. phys. Soc. **52**, 23.
SEEGER, A. (1956); Encyclopedia of Physics, Vol. VII-1, Crystal Physics I, p. 383 (Springer, Berlin).
TAYLOR, G. I. (1934); Proc. Roy. Soc. **A145**, 362.
VOLTERRA, V. (1907); Ann. Ecole Norm. Sup. (3), **24**, 400.
WEINGARTEN, G. (1901); R. C. Acad. Lincei (5), **10**, 57.
ZENER, C. (1948); Phys. Rev. **74**, 639.

PART II

IMPERFECTIONS IN METALS

CHAPTER V

SPECIAL PROPERTIES OF LATTICE IMPERFECTIONS IN METALLIC CRYSTALS

The theory of dislocations was developed originally to understand the mechanical properties of metals. Thus it is not surprising that no more need to be said in this chapter about the elastic properties of dislocations. The considerations of Chapters III and IV apply particularly well to metals, with their non-directed, non-polar bonding orces. It is in respect of the *physical* properties, notably the electrical resistivity, that the description of the various lattice imperfections still requires elaboration.

1. Electrical Resistance of Lattice Imperfections in Metals

1.1. VACANCIES AND INTERSTITIALS

In the alkali metals, and to a lesser extent in the noble and the alkaline earth metals also, the conduction electrons can be regarded as approximately free, that is, they can be described reasonably well by wave functions $e^{ik\cdot r}$, at least as far as their scattering by defects is concerned. Point defects, in turn, can very well be regarded as positive or negative point charges, surrounded by a screened electric field. Many authors have computed the electronic scattering cross-section of a vacancy in the noble metals on this basis. The results do not differ very much: the influence on the resistivity of copper was found both by JONGENBURGER [1953] and by ABELES [1953] to be 1.3 $\mu\Omega$ cm per atomic percent vacancies. BLATT [1955] and SEEGER and STEHLE [1956] have extended the computation considerably and found 1.5 and 1.7 $\mu\Omega$ respectively. These differences illustrate the inherent uncertainty of the theoretical scattering potential. The results pertain

preponderantly to the effect of *charge*; only very recently the influence of the elastic *strains* around a defect has been considered carefully. Especially around interstitials in close-packed metals these strains contribute appreciably. JONGENBURGER [1955] estimated their effect to exceed twice that of the charge; later, more accurate computations by the Stuttgart group showed that this estimate was too large. It appears at this moment again as if the strain term is rather small: a Frenkel pair contributes approximately 2.9 $\mu\Omega$ cm per atomic percent to the resistivity of copper, in agreement with observational results obtained on irradiated metal. We shall return to this matter in Chapters XV and XVI. Except for nickel, no computations of reliability exist for other metals than the noble metals. A review of all relevant data up to 1956 has been given by BLATT [1958].

1.2. DISLOCATIONS

Dislocations scatter the conduction electrons primarily through their elastic strains. Using the Born-approximation and simplified expressions for the stresses and strains, HUNTER and NABARRO [1953] were the first to offer a really serious computational result for the two tensor components $\Delta\varrho_1$ (in the glide direction) and $\Delta\varrho_2$ (normal to the glide plane; the third component parallel to the dislocation axis is of course zero) of the resistivity of an edge dislocation in copper and in sodium. They found an anisotropy ratio $\Delta\varrho_2/\Delta\varrho_1$ of about three.

Extending their computations to screw dislocations, the resistivity of the latter turned out to be much smaller than of edge dislocations in sodium, but not in copper, where the free electron approximation, even if effective masses are taken into account, does not well apply. However, these results have been criticised by several authors; the validity of the Born-approximation has been questioned as well as the use of linear elasticity theory to compute the strains. Taking these matters into account, SEEGER and BROSS [1960] obtained about ten times higher a resistivity of edge dislocations in copper than Hunter and Nabarro did; the anisotropy ratio remained 3. The exact value of the resistivity depends on the so-called "cut-off" radius chosen, outside which the dislocation strains are considered to vanish—since the problem yields a divergent result otherwise—. Taking this radius as about 10^{-4} cm, Seeger and Bross found for the mean resistivity of an edge dislocation in copper 5×10^{-14} $\mu\Omega$ cm/cm^2 (resistance of 1 cm dislocation in 1 cm^{-3} metal). The resistivity of a screw dislocation is

much less, viz. 0.2×10^{-14} $\mu\Omega$ cm/cm^2 (SEEGER and STEHLE [1956]), since the strains are smaller and volume changes are absent.

A complication arises since in f.c.c. metals the dislocations are usually split up into partials, and the contribution to the resistivity by the stacking fault must also be computed. It was shown by KLEMENS [1953] that the conduction electrons may be rather effectively *reflected* at the plane of the fault, which constitutes a break in the periodicity of the lattice, although no strains are associated with it. Klemens estimated the reflection coefficient as about $\frac{1}{4}$, Seeger and Bross adopted $\frac{1}{3}$; both values result in about doubling the resistance of an extended dislocation as compared to a perfect one (according to the Seeger-Bross value). ZIMAN [1956] has disputed this, he maintained a reflection coefficient of only 10^{-4}, resulting in a completely negligible contribution to the resistivity of the stacking fault, which seems unlikely.

The contribution of a dislocation to the resistivity tensor is of course also a tensor. In the coordinate system formed by the axis, the Burgers vector and the normal to the glide plane of an edge dislocation, this tensor has only two components, the components $\Delta\varrho_1$ and $\Delta\varrho_2$ already mentioned earlier. An arbitrary perfect dislocation can always be regarded as the superposition of an edge and a screw dislocation (of the latter the two resistance tensor components are equal), hence the resistance tensor for a straight mixed dislocation is easily written down (VAN BUEREN [1957]). However, in the case of an extended dislocation, this is no longer so. There, three directions, which need not even be mutually perpendicular, play a part: the axis of the dislocation, the normal to its stacking fault and the Burgers vector. Some amount of anisotropy is now certainly also present and we shall see that this property of dislocation scattering can help to distinguish these defects from point defects.

1.3. MAGNETORESISTIVITY OF DISLOCATIONS

The anisotropy of dislocation scattering should show up in the behaviour of the electrical resistance of a dislocated metal in a magnetic field. There are other properties of dislocations too, which may influence the *magnetoresistance* of a metal in some way or another. For instance, the long-range strainfield of a dislocation will disturb the Brillouin zone structure as well as the shape of the Fermi surface of the conduction electrons over appreciable regions of the crystal. It is not easy to form a quantitative estimate of these effects. Only the

influence of anisotropy has been regarded so far (VAN BUEREN [1955, 1957]); it should result in an increase of the relative magnetoresistance of a metal when dislocations are present, which is of the same order as the relative increase of ordinary resistance). This result was confirmed for the case of wires of pure polycrystalline copper (JONGENBURGER [1953]). However, in gold and aluminium a decrease of the magnetoresistance was observed (JONGENBURGER [1960]). This discrepancy may be due to the closeness of the Fermi surface to the Brillouin zone in the case of gold, and to the still more intricate electronic structure of aluminium, in contrast to the relatively simple metal copper.

However this may be, there is little doubt that, in general, dislocations have an appreciable effect on the magnetoresistance of metals. In this they differ significantly from point defects, which affect the magnetoresistance slightly if at all, as is borne out by experiment. This is also obvious theoretically: such defects presumably do not produce much anisotropic scattering in highly symmetric crystal structures. Interstitials, though, with their large and possibly anisotropic strain field, might be expected to produce some anisotropic scattering. An indication of this is present in Jongenburger's work (see section XV–5).

2. Thermal Conductivity and Lattice Imperfections

Not only electrons, but also lattice waves or *phonons* are scattered by lattice imperfections. This effect was first studied by KLEMENS [1953], who pointed out that the temperature dependence of the phonon scattering would be different according to the dimensionality of the imperfection. Dislocation scattering will be proportional to T^{-2}, whereas the lattice thermal resistance owing to point defects is proportional to T. Wide planar imperfections scatter proportionally to T^{-3}. CARRUTHERS [1959] found the following value for the thermal resistivity owing to dislocations in copper:

$$W_{\mathrm{disl}} = \frac{3 \times 10^{-7}}{T^2} N \text{ (watt)}^{-1} \text{ cm degree},$$

where N is the dislocation density. This is much higher than Klemens's original estimate. The discrepancy is due—as in the case of the elec-

Note added in proof: Criticism has been raised against this computation, and Klemens's value might be more appropriate after all (CARRUTHERS, Rev. Mod. Physics **33**, p. 117 (1961)).

trical resistivity—to the slow decrease of strains around a dislocation.

Of course the detection of this type of scattering in metals is only possible when the electronic contribution to thermal scattering is as far as possible reduced. In practice thermal conductivity measurements can therefore only be used to detect imperfections in alloys or superconductors at very low temperatures. (Chapter XV).

3. X-ray and Electron Scattering by Dislocations

Since the presence of a dislocation in a crystal means distortion of the periodic structure over appreciable distances, it should be observable in the X-ray diffraction pattern. If

$$\mathbf{s} = \frac{\mathbf{s} - \mathbf{s}_0}{\lambda} \tag{1}$$

denotes the vector difference between normals to the wavefronts of the incident and reflected monochromatic X-ray beams, divided by the wavelength λ, then the amplitude of diffraction is given by

$$G = \Phi \sum_{1,2,3} j \exp(2\pi i \mathbf{s} \cdot \mathbf{r}_j)$$

where $\mathbf{r}_j$ denotes the position vector of the atoms. Φ depends on the nature of the solid. In a perfect crystal $\mathbf{r}_j$ is periodic

$$\mathbf{r}_j = j_1\mathbf{a} + j_2\mathbf{b} + j_3\mathbf{c}\,,$$

and G differs from zero only when

$$\mathbf{S} \cdot \mathbf{a} = h\,, \qquad \mathbf{S} \cdot \mathbf{b} = k\,, \qquad \mathbf{S} \cdot \mathbf{c} = l\,;$$

h, k, l being integers. In a dislocated crystal $\mathbf{r}_j$ is no longer strictly periodic, and diffraction amplitudes occur also for other reflection angles. A. J. C. WILSON [1949, 1950] has studied the case of a straight dislocation more in detail; he finds that the sharp diffraction spots of the perfect crystal are broadened into rings. It has not been possible to verify this in actual cases because of the mutual interference caused by other dislocations and second order core effects.

Dislocations make their presence known in a typical manner in the integrated intensities of diffraction lines. In a really perfect crystal reflection should occur at the Bragg angles and no reflection anywhere else. The integrated intensity of the (very sharp) lines is then propor-

tional to their width, and therewith to F, the so-called structure factor of the reflection.

In non-perfect crystals mutual interference of incident and reflected beams is less complete, and broadened lines with wings appear. This leads to integrated intensities that are proportional to F^2. The latter relation is nearly always observed in actual crystals and was explained originally by Darwin as due to the so-called *mosaic structure* of all natural crystals. This structure, namely the occurrence of many small blocks of slightly differently oriented material in the apparently strictly "single" crystal, can be explained on more modern lines as being equivalent to the presence of fairly large numbers of dislocations throughout the crystal.

Stacking faults produce a special anomaly in the diffraction pattern, first described by PATERSON [1952]. Interference phenomena caused by stacking faults have been observed by *electron* diffraction also (WHELAN and HIRSCH [1957]).

When a metal is cold-worked, the background scattering of X-rays is enhanced, because the atoms in the dislocation cores scatter as though they formed a liquid (HALL and WILLIAMSON [1951]). This, together with other influences of artificially introduced lattice imperfections on the physical properties of metals will be discussed in detail in Chapter XVI. A general discussion of X-ray phenomena can be found in Chapter XVII of Barret's book (BARRETT [1952]).

4. Rendering Lattice Imperfections Visible

The detection of imperfections by the scattering of electrons, phonons or electromagnetic waves is only an indirect method. It goes without saying that a diligent search for more immediate evidence for the presence of dislocations and point defects has gone on for years. Dislocations especially have been the subject of such investigations, during recent years with remarkable success. Although the techniques for making dislocations visible are not only applicable to metals, and some of them are not at all applicable to opaque materials, it seems nevertheless appropriate to discuss them at this stage, before proceeding with the largely theoretical description of dislocation processes, the occurrence of many of which have been beautifully confirmed by these experimental methods.

4.1. ETCH PITS AND DISLOCATIONS

A promising way to obtain direct optical evidence of the presence of defects in metals is by thermal, chemical or electrolytical *etching* of polished metal surfaces. Depending on the nature of the etchant and the method used, few or many etch pits, bounded by crystallographic planes, can often be observed on metal surfaces. These usually congregate near those zones in deformed metals which were indicated by glide lines before etching (section VI–1, MACHLIN [1957]). It was therefore generally concluded that such etch pits indicate sites where dislocations meet the surface. This conclusion could afterwards be confirmed semi-quantitatively. Up till now it has not been possible, however, to prove that the correspondence between etch pits and dislocations is perfect, that is, that each etch pit corresponds to a dislocation and each dislocation produces an etch pit. Recent evidence obtained by Lacombe and coworkers, e.g. WYON and MARCHIN [1955] and by SUITS and LOW [1957] seems to indicate that, at least in aluminium and silicon iron, only those dislocations form etch pits that are "decorated" with impurity atoms (especially iron atoms in the former case and carbon atoms in the latter case) or conglomerates of them. For instance, in those regions of an aluminium sample where the impurities had been artificially swept away, for instance, by moving crystal boundaries, no dislocation etch pits appeared at all, even after very deep etching (WYON and LACOMBE [1955]) (Fig. 1).

The semiconductors germanium, silicon etc. lend themselves especially well to etching treatments, and in these materials a one-to-one correspondence between dislocations and etch pits, at least on certain crystallographic faces, can be obtained by routine methods. The pattern is apparently independent of the presence of impurities. The technique of etching germanium and silicon surfaces to show dislocations has been brought to a high degree of perfection by Dash, who was even able to show the motion of jogged dislocations, which is accompanied by defect formation, by this method (DASH [1958]).

Using the etching method, GILMAN and JOHNSTON [1957, 1958] succeeded in observing the motion of individual dislocation lines in lithium fluoride. After a given deformation etch pits were developed along slip lines. A second deformation followed by renewed etching revealed new, smaller etch pits, along the same slip band but somewhat displaced, whereas the old etch pits had developed a flat bottom

(Fig. 2). This indicates the absence of the dislocations, originally associated with the latter etch pits; they have migrated to the positions indicated by the new pits. In this way the motion of individual dislocations could be directly followed and even the velocity of dislocations as function of stress could be determined (GILMAN and JOHNSTON [1958]).

Similar observations have been published by DORENDORFF [1957] on germanium surfaces. There, the displacement of etch pits was studied as a function of the time of annealing after a slight previous deformation (polygonization, see Chapter XI). Typical association processes between nearby dislocations were repeatedly observed.

By repeated etching and gradually dissolving away successive surface layers, the course of dislocation lines in the inside of a crystal has been followed; moreover, some aspects of the interaction of dislocations with the surface could be studied in this way. For a discussion of the many observations in this field the reader is referred to: Dislocations and Mechanical Properties of Crystals (Wiley, New York, 1957).

CABRERA [1956] has studied theoretically the formation of etch pits at dislocation emergence sites in solids. The problem bears a close resemblance to the phenomena of preferred evaporation and of crystal growth by a dislocation mechanism (FRANK [1952], see Chapter XXIII). The effect of dislocations on the etching process is explained by the decrease, near a dislocation, of the energy required to remove an atom and or an electron from the surface of the solid near the dislocation. The rate of "evaporation" of atoms from the surface (in contact with the vapor, the etchant liquid or other medium) is given by

$$v = a\nu \exp(-\Delta G/kT) . \tag{1}$$

Here a is the linear dimension of the site from which the atoms evaporate. In the case of an emerging dislocation, this site is a step on the surface, equal in height to the component of the Burgers vector normal to that surface. At the perfect surface, the step height a is the atomic spacing and of the same order of magnitude; ν is a frequency factor, and ΔG is the energy of activation of the evaporation process. The velocity must exceed a certain value, say 10^{-5} cm/sec, for visible etch pits to develop individually and within a reasonable time. Near to a dislocation, ν and ΔG differ from their values ν_0 and ΔG_0 at an arbitrary perfect site of the crystal surface. Whereas ν_0 is about 10^{13}

sec^{-1} in a typical case, ν is only 10^7 sec^{-1}. On the other hand, ΔG will be certainly smaller near dislocations than elsewhere, owing to the disturbance of the lattice that is already present. Etch pits will develop at dislocations only if

$$v > v_0$$

the index 0 again indicating the evaporation rate of the perfect surface. Cabrera obtained the following formula for ΔG near a dislocation site

$$\Delta G = \frac{\pi a \gamma^2 V}{12 \Delta \mu_c} \left(\frac{\Delta \mu_c - \Delta \mu}{\Delta \mu_c} \right)^{3/2}, \tag{2}$$

where γ is the surface energy, V the molar volume and $\Delta \mu$ the difference in chemical potential of an atom inside the crystal and in the surrounding medium. The quantity $\Delta \mu_c$ is given by

$$\Delta \mu_c = \frac{\pi \gamma^2 V}{\alpha G b^2}, \tag{3}$$

where α is a constant of order unity and b is the Burgers vector of the dislocation. $\alpha G b^2$ is simply the line tension of the dislocation (section III–2); $\Delta \mu_c$ is thus the surface energy corrected by the ratio between surface and dislocation energies. On the perfect surface, the activation energy for evaporation is given by

$$\Delta G_0 = \pi a \gamma^2 V / \Delta \mu . \tag{4}$$

From these two formulae one can easily derive a certain critical value that must be exceeded by $\Delta \mu$, in order that a dislocation of given Burgers vector b may give rise to the development of an etch pit. One has in first approximation the condition:

$$\Delta \mu > \beta \frac{\gamma^2 a}{G} (a/b)^2 , \tag{5}$$

where β is a numerical factor of the order of a few tenth.

Thus, etch pits are most easily formed at dislocations with large Burgers vectors, that is, *composite* dislocations, (see Chapter XXIII). In fact, for $b \approx 100a$ and $\gamma a^2 \approx kT$, $\Delta \mu \geqslant 10^{-18}$ erg, which means that concentration of atoms in the liquid must be less than 10^{-4} times that corresponding to saturation. This is easily realized. However, near ordinary dislocations with elementary Burgers vectors ($b \approx a$), the

situation is much less favourable. It is a matter of doubt whether at ordinary dislocations in metals both conditions, that of a sufficiently rapid evaporation rate and that expressed by (5), can actually be fulfilled at the same time. Indeed, Cabrera finds that under fairly general circumstances etch pits can not be generated if $b < \frac{1}{2}a$. It is possible to decrease this limit slightly by taking into account not only the strain energy of the dislocation but also the core energy. However, Cabrera advocates as his opinion that really efficient etch pit formation will only occur when the surface energy is artificially decreased near the dislocation, owing to the presence of another, foreign atom or molecule at the emergence site of the dislocation. This would agree with the result of the experiments cited, which also seem to indicate the necessity of the presence of impurities in dislocations, for the formation of etch pits on *metallic* surfaces to take place. Cabrera's computations show that careful consideration of the theoretical possibility of etch pit formation on pure dislocations is necessary for each type of crystal separately. Recently, a new approach to the problem of chemical attack on imperfections has been made by FRANK [1958]; to discuss this here would carry us too far.

4.2. OTHER METHODS FOR DETECTING INDIVIDUAL DISLOCATIONS

Apart from by etching, followed by microscopical examination, dislocations have been recently made visible by a number of other methods. A very detailed description of some of these methods and their results can be found in the book Dislocations and Mechanical Properties of Crystals, already mentioned before, and we shall therefore be brief.

Decoration of dislocations by letting suitably chosen "impurities" precipitate on dislocation sites, has been applied with success in a number of cases. The opaque precipitates arranged along a dislocation line make the latter detectable under the microscope in optically or infrared transparent crystals. Thus dislocations could be made visible in silicon crystals (DASH [1956], Fig. 3), in silver halides (HEDGES and MITCHELL [1953]) and in alkali halides (AMELINCKX [1956], Fig. 4). Narrow *steps* on a *cleavage surface* of NaCl can be decorated similarly (BASSETT *et al.* [1959], Fig. 5).

Electron-microscopical investigation of thin metal foils has enabled HIRSCH and co-workers [1957] to view dislocations in motion. Contrast is here obtained by reason of the lattice strains around the dislocations,



Fig. 1. Absence of dislocation etch pits in zones swept over by grain boundaries in deformed aluminium that had been submitted to a grain growth treatment (after WYON and LACOMBE [1955], 120×).

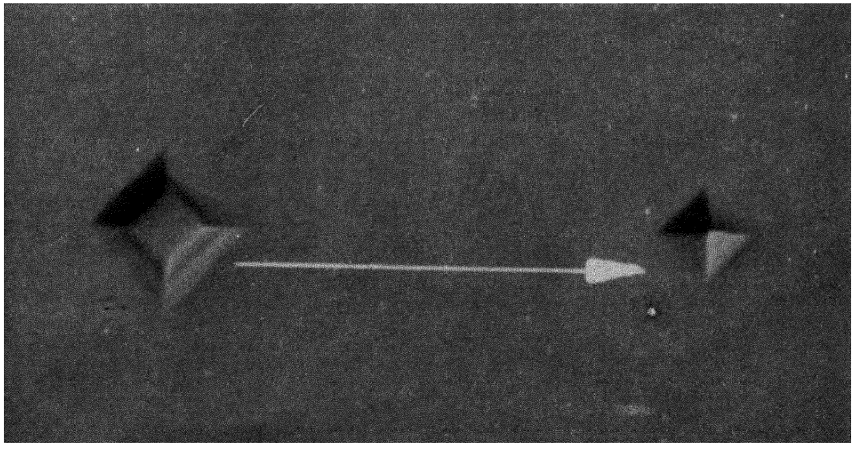


Fig. 2. Etch pits on deformed lithium fluoride. Pit originating from previous deformation (left) developed a flat bottom, an indication that it has lost its dislocation. This has formed a new, smaller etch pit some distance away. (After GILMAN and JOHNSTON [1957], ca. 1000×).

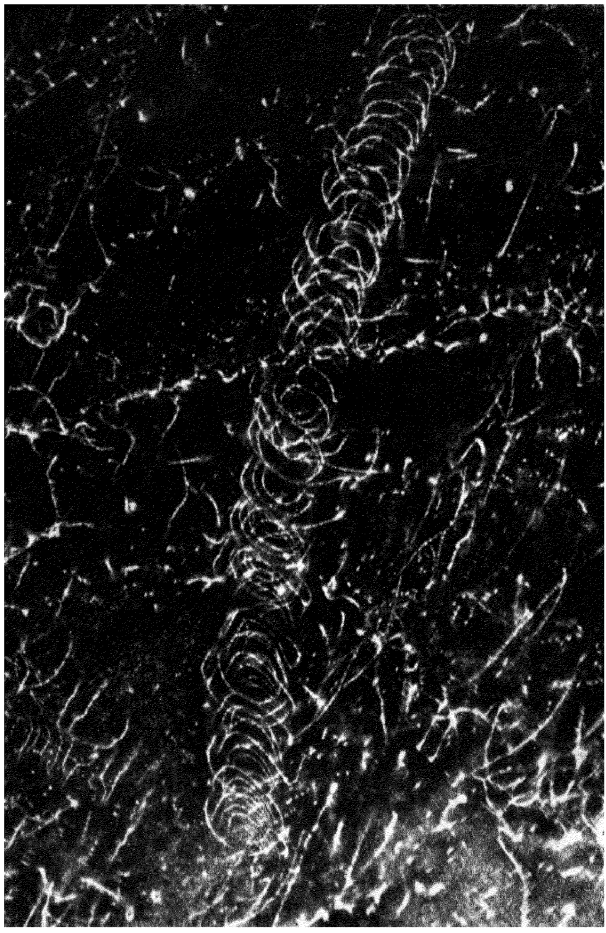


Fig. 3 Dislocation loops along crack front in KCl crystal (after AMELINCKX [1956]).



Fig. 4. Surface sources in cleaved KCl ($+AgNO_3$) crystal. Decoration by irradiation with X-rays followed by annealing (after AMELINCKX [1956]).

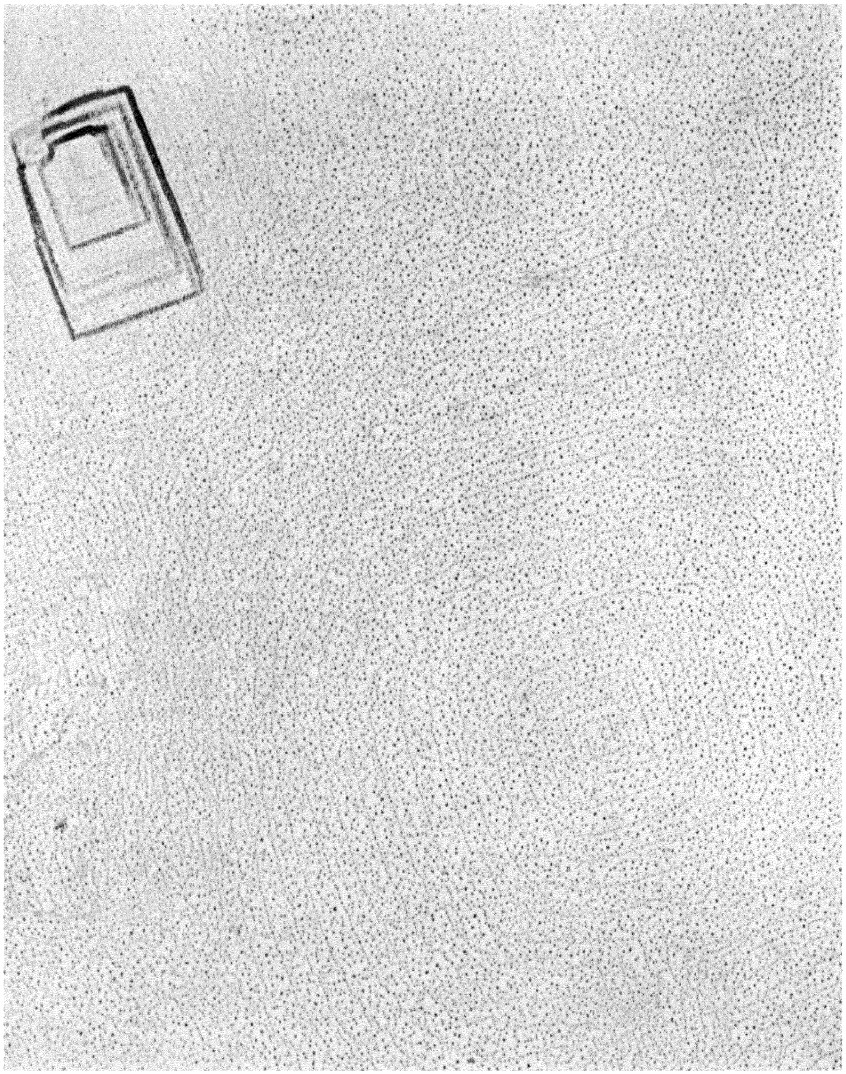


Fig. 5. Shallow steps on a cleaved surface of a rocksalt crystal. They are associated with the screw dislocation in the left hand part of the figure, their height is one atomic distance. These steps have been made visible with a technique originally due to BASSETT *et al.* [1959]: gold vapour is evaporated onto the surface and particles collect along the steps on the latter. A carbon replica is made to which the gold particles become attached. The large spiral below at the right is presumably due to a screw dislocation with a large Burgers vector. Photograph by Haanstra, unpubl., 55 000 ×.

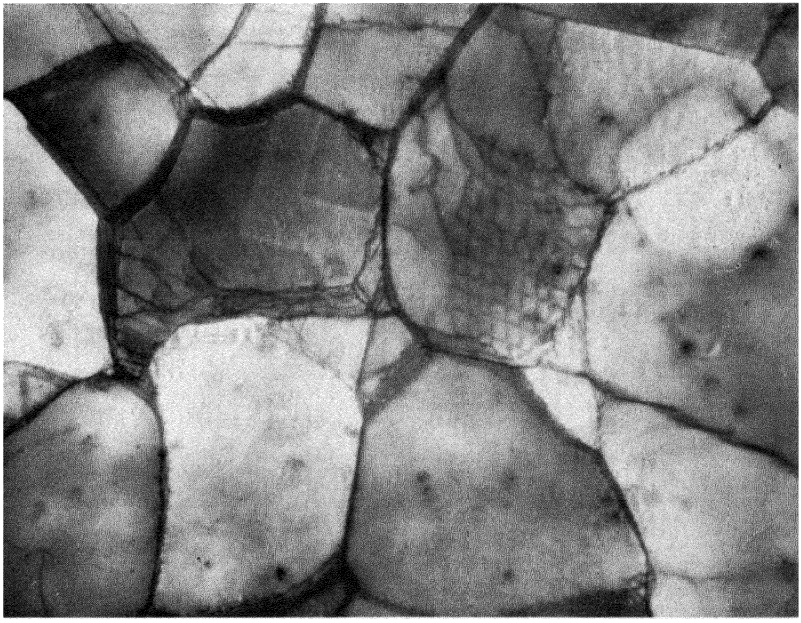


Fig. 6. Well-developed dislocation structure in heavily deformed aluminium. The orientation of the normal to the foil is parallel to [100]; the cross-grid consists of dislocations approximately parallel to [110] direction, and corresponds presumably to a twist boundary on (100) made up of a cross-grid of screws. Note the cross-slip trace in the neighbouring subgrain. (M. J. WHELAN, [1957], Ph. D. Dissertation, University of Cambridge.)

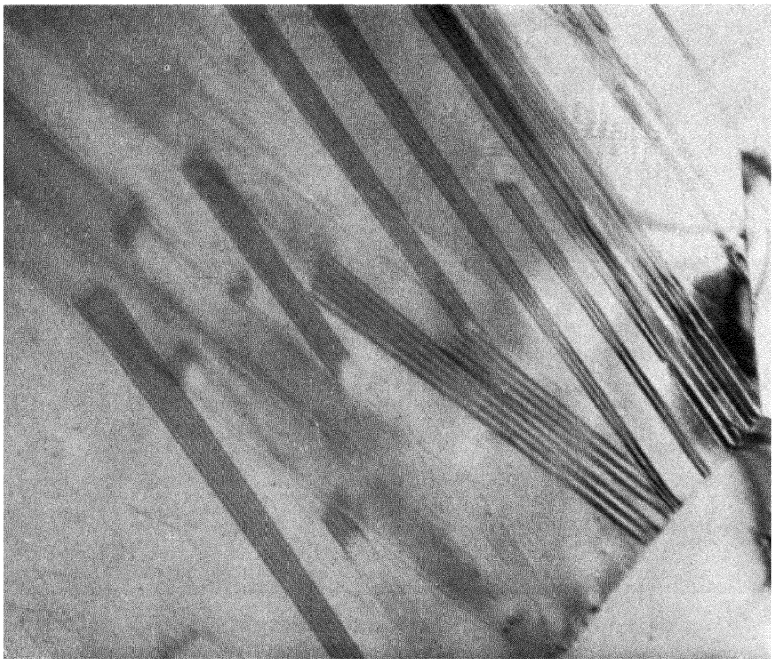


Fig. 7. Wide stacking faults terminated by partial dislocations in stainless steel. The faults traverse the foil from top to bottom. The striations are interference fringes which result from the interference of the electron beams diffracted from the parts of the crystal on the two sides of the fault. (M. J. WHELAN, P. B. HIRSCH, R. W. HORNE and W. BOLLMANN [1957], Proc. Roy. Soc. **A 240**, 524.)

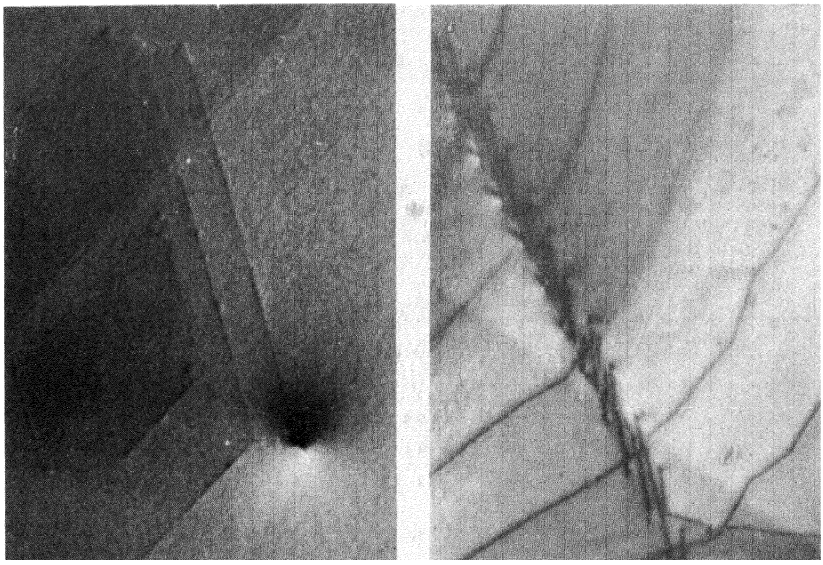


Fig. 8. *a*) Dislocation with growthstep attached to it, viewed in transmission through a thin (0.1μ) slice of ferroxdure, a ceramic magnetic material of composition. $Ba_{12}Fe_6O_{19}$. *b*) Lowangle tiltboundary in the same material. After unpublished work of HAANSTRA and KOOY in the Philips Laboratories in Eindhoven; enlarged about 40000.

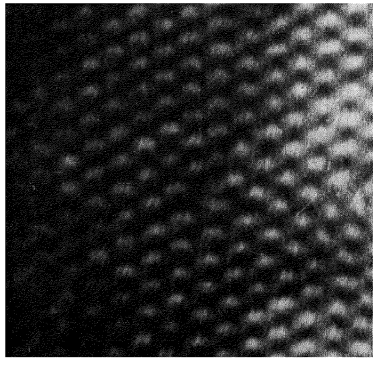


Fig. 9. Moiré pattern from (111) palladium layer deposited onto gold film. A dislocation with Burgers vector $\frac{1}{2}a\langle 110\rangle$ gives rise to a terminating half line of "atoms" (G. A. BASSETT, J. W. MENTER and D. W. PASHLEY [1958], Proc. Royal Soc. **A246**, 345).

which reduce the transmittance of the foil for electrons (Figs. 6, 7). The method is very powerful: the splitting up of dislocations into partials and the interaction between them can easily be studied (the results invariably confirming the already existing theory beautifully). In quenched aluminium the presence of small dislocation loops formed by the collapse of vacancy aggregates could be demonstrated by this method (HIRSCH *et al.* [1958]); see Chapters III and XIV.

Recently, the transmission method has been applied to a large variety of non-metallic crystals, e.g. talc (AlMg-silicate) and ferroxdure, a ceramic, sintered compound of iron and bariumoxyde, which crystallizes in a complicated, hexagonal structure. Fig. 8 gives some examples of dislocations in the latter material.

A third method that looks promising makes use of the moiré-effect. Two very thin layers of crystal, either in parallel orientation but with slightly different lattice parameters, or else with identical parameters but slightly rotated with respect towards each other, are viewed simultaneously in transmission. A lattice imperfection in one of the layers produces a typical disturbance of the moiré "interference" pattern, in which the characteristic structure of the imperfection is reproduced at an enlarged measure (i.e. in the ratio $2d/\Delta d$ where Δd is the difference in lattice parameter d). The method has been applied with success to layers of gold and palladium (Fig. 9) by MENTER *et al.* [1958] and by HASHIMOTO and UYEDA [1957] to layers of cupric sulphide.

Finally the field-ion microscope deserves to be mentioned. This was developed by Müller and coworkers and serves to study the surface layers of tiny single crystals.

A review of electron-microscopical techniques for the detection of dislocations has been recently prepared by MENTER [1958].

REFERENCES

ABELES, F. (1953); C. R. Acad. Sci. Paris, **237**, 796.
AMELINCKX, S. (1956); Phil. Mag. **1**, 269.
BARRETT, C. S. (1952); Structure of Metals (McGraw-Hill, New York), Chapter XVII.
BLATT, F. J. (1955), Phys. Rev. **99**, 1708; (1956), Phys. Rev. **103**, 1905.
BLATT, F. J. (1958); Solid State Physics **5**, 199.
BUEREN, H. G. VAN (1955), in: The Deformation and Flow of Solids (Springer, Berlin).
BUEREN, H. G. VAN (1957); Philips Res. Reports **12**, 1.
CABRERA, N. (1956); J. Chémie Physique **53**, 675.

CARRUTHERS, P. (1959); Phys. Rev. **114,** 995.
DASH, W. C. (1956); J. appl. Physics **27,** 1153.
DASH, W. C. (1958); J. appl. Physics **29,** 705.
DORENDORFF, H. (1957); Z. angew. Physik **9,** 513.
FRANK, F. C. (1952); Advances in Physics **1,** 91.
FRANK, F. C. (1958); unpublished.
GILMAN, J. J. (1956); Trans. AIME **206,** 998.
GILMAN, J. J. and W. P. JOHNSTON (1957), in: Dislocations and Mechanical Properties of Crystals, p. 116 (Wiley, New York).
GILMAN, J. J. and W. P. JOHNSTON (1958); J. appl. Phys. **30,** 129.
HALL, W. H. and G. K. WILLIAMSON (1951); Proc. phys. Soc. **B64,** 937.
HASHIMOTO, H. and R. UYEDA (1957); Acta Crystall. **10,** 143.
HEDGES, J. M. and J. W. MITCHELL (1953); Phil. Mag. **44,** 223 and 357.
HIRSCH, P. B., R. W. HORNE and M. J. WHELAN (1957), in: Dislocations and Mechanical Properties of Crystals (Wiley, New York).
HIRSCH, P. B., J. SILCOX, R. E. SMALLMAN and K. H. WESTMACOTT (1958); Phil. Mag. **3,** 897.
HUNTER, S. C. and F. R. N. NABARRO (1953); Proc. Roy. Soc. **A220,** 542.
JONGENBURGER, P. (1953); Appl. Sci. Res. **B3,** 237.
JONGENBURGER, P. (1955), Nature **175,** 544; (1960), to be published.
KLEMENS, P. G. (1953); Australian J. Physics **6,** 122.
KLEMENS, P. G. (1958); Solid State Physics **6,** 1.
MACHLIN, E. S. (1957); Dislocations and Mechanical Properties of Crystals (Wiley, New York).
MENTER, J. W. (1958); Advances in Physics **7,** 298.
MENTER, J. W., G. A. BASSETT and D. W. PASHLEY (1958); Proc. Roy. Soc. **A246,** 345.
OVERHAUSER, A. W. (1953); Phys. Rev. **90,** 393.
OVERHAUSER, A. W. and R. L. GORMAN (1956); Phys. Rev. **102,** 676.
PASHLEY, D. W., J. W. MENTER and G. A. BASSETT (1957); Nature **179,** 752.
PATERSON, M. S. (1952); J. appl. Physics **23,** 805.
ROSENBERG, H. M. and J. LOMER (1958); Phil. Mag. **4,** 960.
SEEGER, A. (1956); Canadian J. Physics **34,** 1219.
SEEGER, A. (1958); Encyclopedia of Physics, Part VII–2, Crystal Physics II, p. 1 (Springer, Berlin).
SEEGER, A. and H. STEHLE (1956); Z. Physik **146,** 242.
SEEGER, A. and H. BROSS (1960); Z. Naturf. **15a,** 663.
SUITS, J. C. and J. R. LOW (1957); Acta Metall. **5,** 285.
WHELAN, M. J. and P. B. HIRSCH (1957); Phil. Mag. **2,** 1303.
WILSON, A. J. C. (1949); Research (London) **2,** 541.
WILSON, A. J. C. (1950); Research (London) **3,** 387.
WYON, G. and P. LACOMBE (1955); Rep. Conf. Defects in Solids, Bristol, p. 187.
WYON, G. and J. M. MARCHIN (1955); Phil. Mag. **46,** 1119.
ZIMAN, J. M. (1956); Canadian J. Physics **34,** discussion to Seeger's paper p. 1234.
ZIMAN, J. M. (1958); Suppl. al Nuovo Cimento VII, Ser. X, No. 2, 353.

CHAPTER VI

PLASTIC DEFORMATION OF METALS CONSIDERED AS MOVEMENT OF DISLOCATIONS

1. Visual Characteristics of Plastic Deformation

This chapter contains, together with Chapters VII to IX, a description of the most important aspects of the plastic deformation of metals. Whereas the latter three chapters contain quantitative considerations, the first is purely descriptive and may be used to obtain a review of the subject.

In the discussion of the fundamental properties of the dislocations, it has already been pointed out that these imperfections can be considered as boundaries between *slipped* and *unslipped* zones of a crystal. The phenomenon of *slip* or *glide* in metals is known as the most important mode of deformation. It is defined as the shear motion of two parts of a crystal across a common lattice plane under the action of a deforming stress. The visual result on the surface of the crystal is the occurrence of a set or sets of parallel *slip lines* or *bands*, already visible under slight magnification (Fig. 1). This was first explained by Carpenter and Elam in 1920; refer. ELAM [1935]. SCHMID and BOAS [1935] considered a cylindrical sample, elongated by a tensile stress applied along the long axis of the crystal. The maximum shear stress component occurs on a plane at 45° to the axis.

Lamellae of crystal containing this component shear along each other, giving an extended *and* reoriented slipped crystal. Invariably, glide occurs only in certain directions that are determined by the orientation of the shear stress(es) with regard to the crystallographic axes. For a very detailed description of the geometry of slip, the reader is referred to a book by KOCHENDÖRFER [1941].

The common *glide plane* of course contains the glide or slip direction; together they form the *glide* or *slip system*. In polycrystals or arbitrarily oriented single crystals, generally more than one slip system

is active (multiple slip); however, all systems belong to one or two crystalline orientations only, which may be different in different metals. The slip systems of the most important metals are listed in Table 1 (mostly after BARRETT [1952]); often a low-temperature and a high-temperature orientation can be distinguished. The latter occurs in addition to the former at elevated temperatures and sometimes even replaces it. Generally in cubic metals the glide planes are the most densely packed planes, the glide directions close packed directions in these planes.

The crystallography of the deformation by glide has been extensively discussed by CHEN and MADDIN [1956], to which paper we refer for more particulars.

The geometry of slip can be deduced from a microscopical investigation of the surface, and from X-ray studies. The appearance of the surface of a previously polished deformed metal crystal shows a number of typical structures, of which Figs. 1 to 3 give examples. Most conspicuous already under slight magnification are the slip lines mentioned above. According to the orientation of the surface with respect to the glide direction they are (in face-centred cubic crystals) long and straight (surface almost parallel to slip vector) or branched and interconnected (if the surface makes an angle with the slip vector). Of course no slip lines appear when the surface is strictly parallel to the glide planes, because then no steps are formed on it. Superposed on the slip line pattern so-called *deformation* or *kink bands* often appear, fairly wide zones approximately perpendicular to the glide lines, in which the latter show a slight curvature (Fig. 2). Apart from the branching already mentioned, which is apparent in Fig. 1 and is often denoted by the term cross-slip, sometimes so-called *bands of secondary slip* or *striae* are noted, in which, in zones approximately parallel to the primary glide lines, a large number of short slip lines belonging to another slip system are observable (Fig. 3).

The detailed slip line pattern of deformed *face-centred cubic crystals* depends strongly on amount of strain, temperature, orientation, etc. SEEGER [1958] recently outlined the various details in his review paper on the properties of this important class of metals. We refer to this paper for particulars and name but a few characteristic observations.

In aluminium the slipline pattern depends significantly on the temperature of straining. Single crystals suitably oriented for single slip show at −180 °C or below long straight, finely divided separate

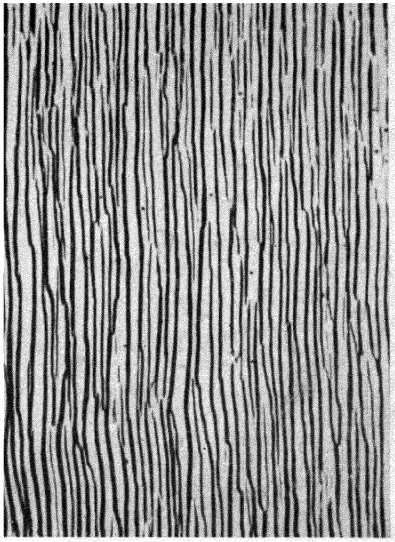


Fig 1. Slip lines on the surface of a 20% deformed aluminium crystal (190×, after CAHN [1951]).

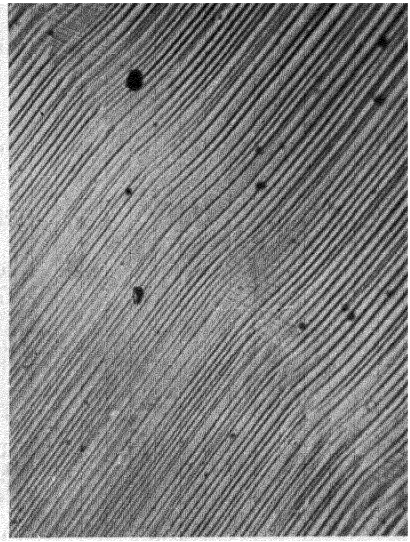


Fig. 2 [illegible] band on 20% deformed aluminium (after CAHN [1951], 250×).

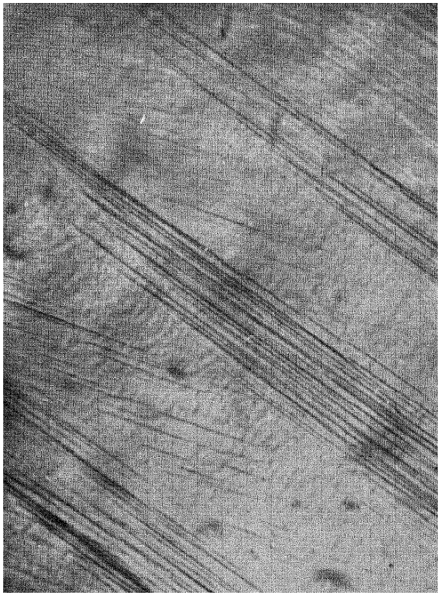


Fig. 3. Bands of secondary slip on deformed aluminium crystal (after JAOUL [1957]; 150×).

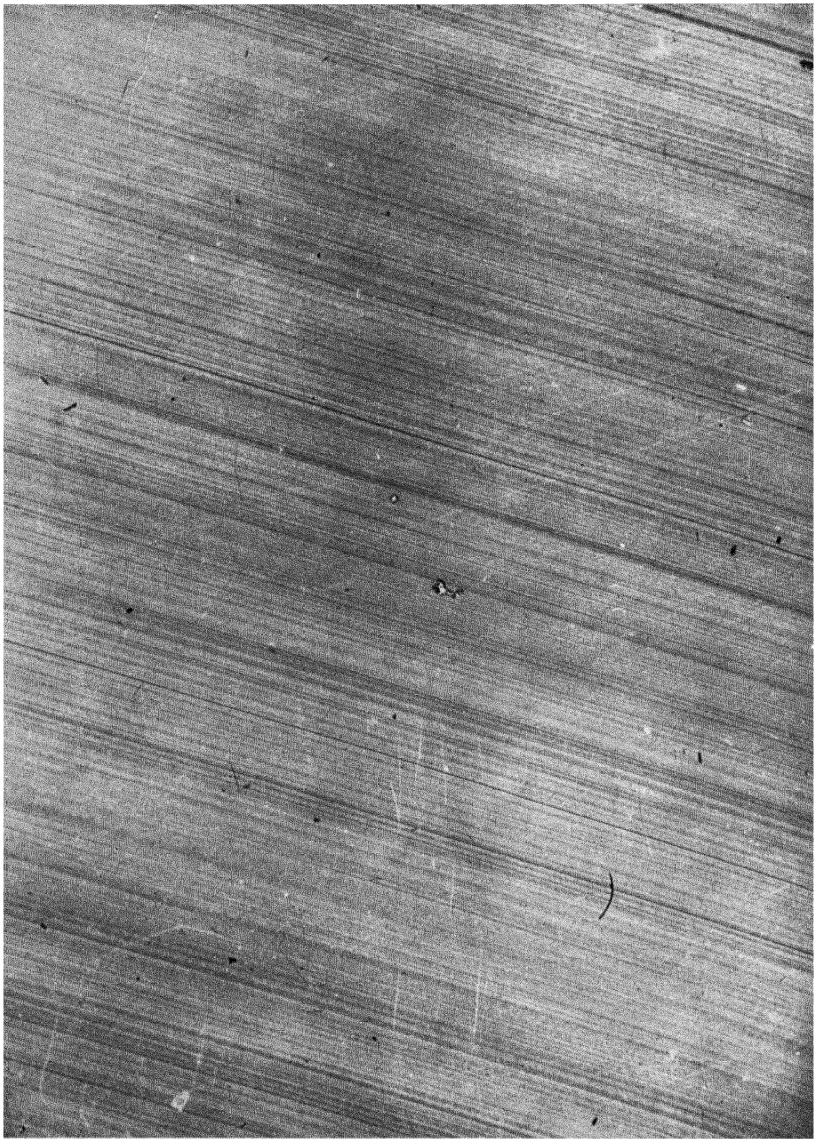


Fig. 4*a*. Electromicrograph (12 500 ×) of aluminium crystal deformed 20% at 90°K. Only fine slip (elementary slip) is observable.

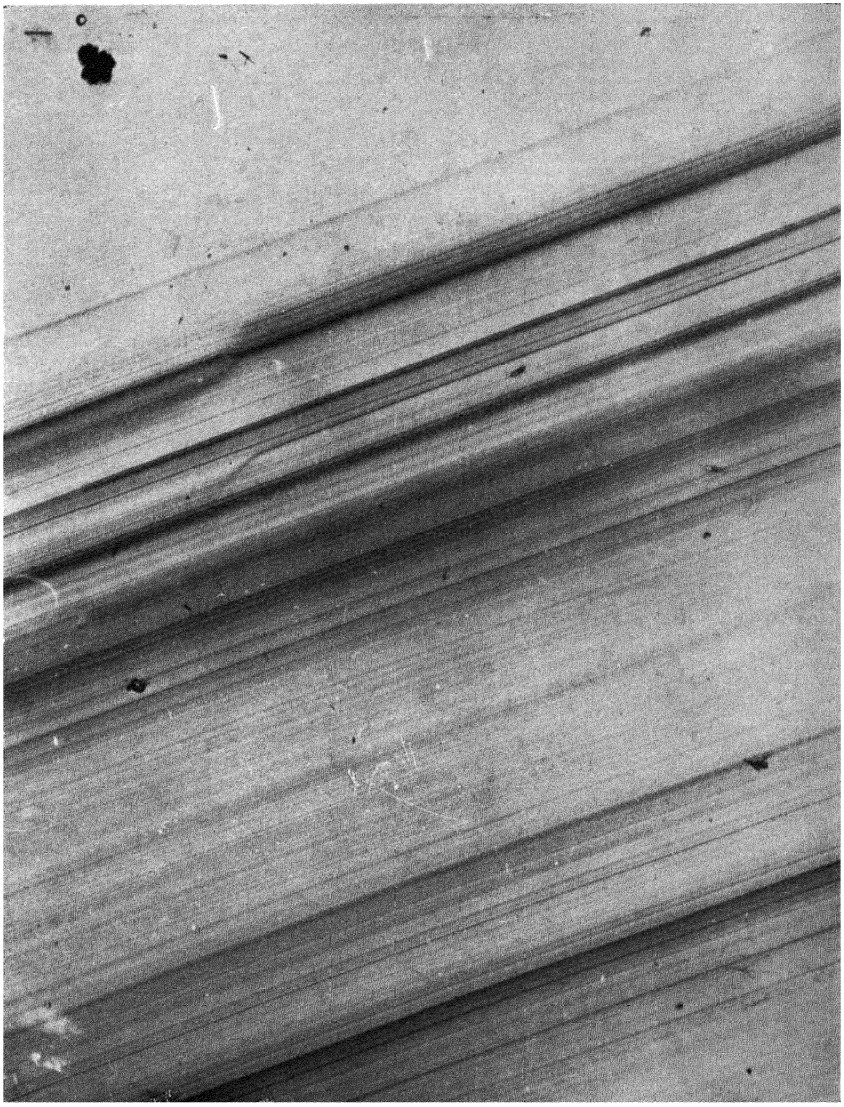
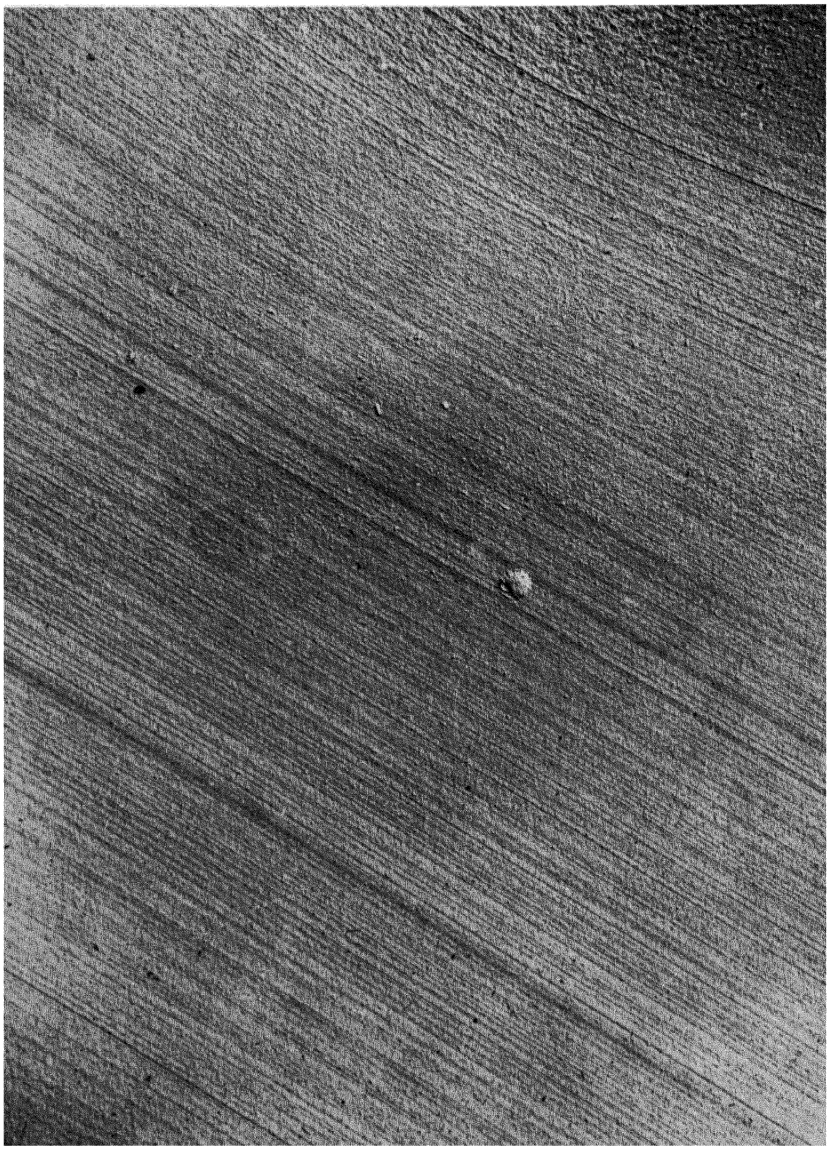


Fig. 4*b*. Aluminium crystal deformed 30% at room temperature (12500 ×). Most of the fine slip lines have clustered into slip bands



Figs 5–8. Slip line patterns on copper single crystal deformed by various amounts at room temperature (electron micrographs).
Fig. 5: $\varepsilon = 8\%$ ("stage I"): fine slip (18000×).

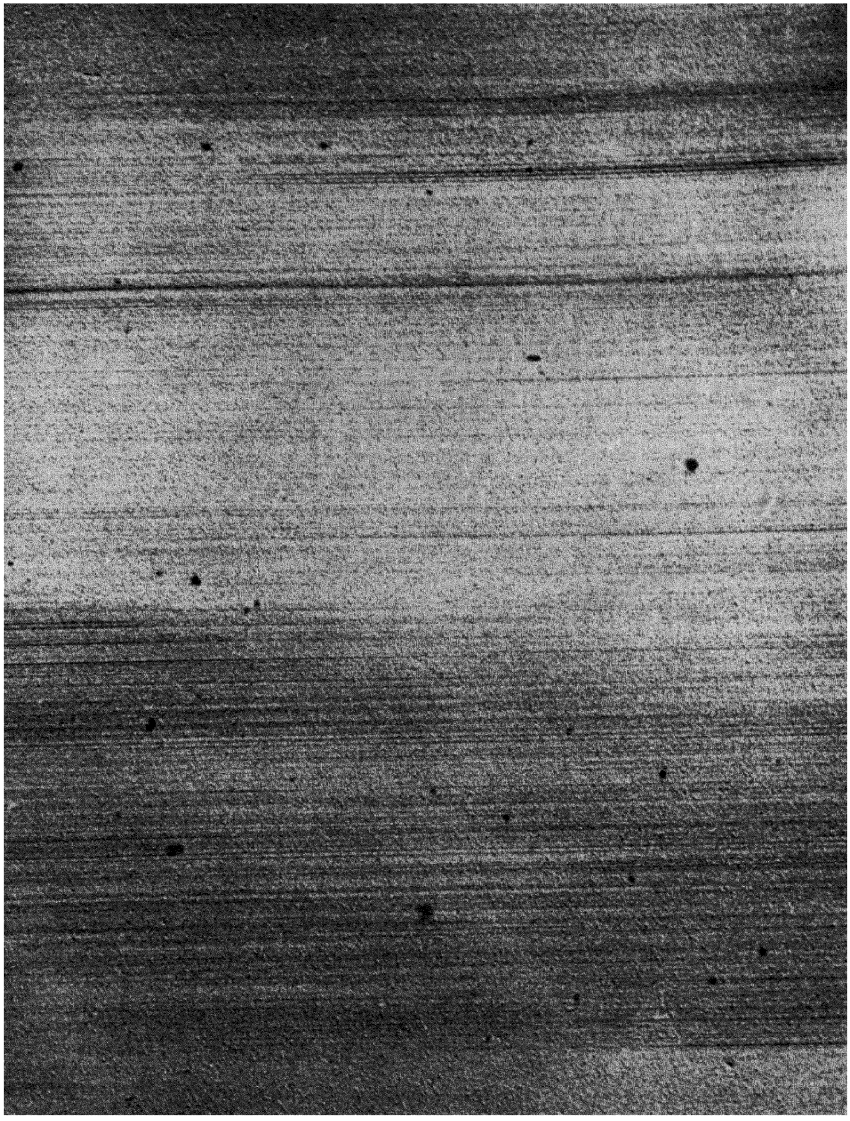


Fig. 6: $\varepsilon = 25\%$ ("stage II"): structural fine slip (13000×).

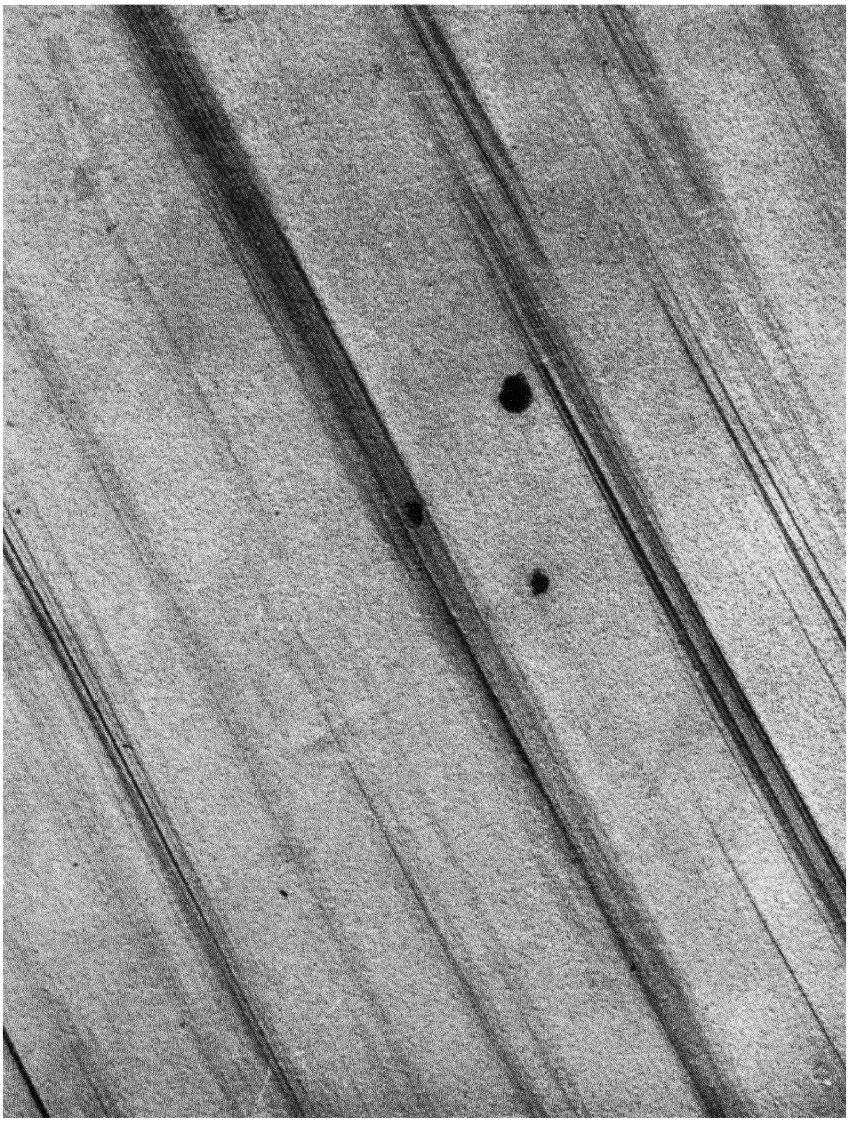


Fig. 7: $\varepsilon = 45\%$ ("stage III"): slip bands with cross slip (12500×).



Fig. 8: $\varepsilon = 65\%$ ("stage III"): fragmented slip bands (7000 ×).

I am much indebted to Dr S. Mader for letting me use his previously unpublished photomicrographs reproduced in Figs. 4 to 8.

TABLE 1

Slip systems in metals and alloys (mostly after Barrett)

Structure	*Metal or alloy*	*Low temperature*		*High temperature*	
		Slip plane	*Slip direction*	*Slip plane*	*Slip direction*
Face-centred cubic	Al	(111), (010)	[$\bar{1}$01]	(010)	[$\bar{1}$01]
	Cu	,,	,,	(111)	,,
	Ag	,,	,,	,,	,,
	Au	,,	,,	,,	,,
	Ni	,,	,,	,,	,,
	Cu–Au	,,	,,	,,	,,
	α Cu–Zn	,,	,,	,,	,,
	α Cu–Al	,,	,,	,,	,,
	Al–Cu	,,	,,	,,	,,
	Al–Zn	,,	,,	,,	,,
	Au–Ag	,,	,,	,,	,,
Body-centred cubic	α–Fe	(110), (211), (321)	[$\bar{1}$11]	(110)	[$\bar{1}$11]
	Mo	(211)	,,	(110)	,,
	W	(211)	,,	(110), (321)	,,
	K	(321)	,,	(321)	,,
	Na	(211)	,,	(110), (321)	,,
	β Cu–Zn	(110), (211)?	,,		
	α Fe–Si	(110)	,,		
Close-packed hexagonal	Mg	(0001) (10$\bar{1}$0)?	[2$\bar{1}\bar{1}$0]	(10$\bar{1}$1)	[2$\bar{1}\bar{1}$0]
	Cd	(0001)	,,	,,	,,
	Zn	(0001)	,,	,,	,,
		(11$\bar{2}$2)	[11$\bar{2}\bar{3}$]	,,	,,
	Be	(0001)	[2$\bar{1}\bar{1}$0]	(10$\bar{1}$0)	,,
	Ti	(10$\bar{1}$0), (0001)	,,	(10$\bar{1}$1)	,,
	Zn–Sn	(0001)	,,	,,	,,
	Zn–Cd	(0001)	,,	,,	,,
Rhombo-hedral	Bi	(111)	[10$\bar{1}$]		
	Hg	(100) and others			
Ortho-rhombic	α–U	(001) and other	[100] systems	(001) and other	[100] systems
Tetragonal	β–Sn	(110) (100) (10$\bar{1}$) (121)	[001] [001], [011] [101] [101]	(110)	[111]
Diamond	C, Ge, Si	(111)	[$\bar{1}$01]	(111)	[$\bar{1}$01]

slip lines which at higher temperatures tend to cluster into more inhomogeneously distributed slip "bands" (Fig. 4). After increasing amounts of deformation, the number of bands increases, until a certain minimum spacing between them is reached; from then on the number of lines per band starts to increase (BROWN [1952]). Similar changes have been found in other face-centred cubic metals.

LANGE and LÜCKE [1953], STAUBWASSER [1954], MÜLLER and LEIBFRIED [1955] (see also LEIBFRIED and HAASEN [1954]), WILSDORF and KUHLMANN [1953] and most recently MADER and SEEGER [1959], have made detailed optical and electron microscope studies and have described the appearance of slip lines on extended single crystals of aluminium, resp. of nickel and copper, as a function of temperature, rate and amount of strain and crystallographic orientation of the direction of elongation. The optical results for aluminium can be summarized as follows (for slight deformation). Elongation at room temperature in [110]-direction: evenly spaced slip lines begin to appear after 5% elongation, on the $(11\bar{1})$ [101] system only. Their separation decreases with increasing strain. Kinkbands not very regularly spaced, also appear approximately perpendicular to the slip lines. Moreover, "cross-slip" from one slip line to another is occasionally present. Apart from these local disturbances, the slip line pattern remains fairly regular.

After elongation in the [100]-orientation the slipbands appear in clusters separated by distances of 0.1 to 1 mm or larger. Slip lines on all four octahedral systems occur, predominantly traces of $(11\bar{1})$ and (111) are observed, however. The clusters at first contain (3.5% strain) about 10 lines. On increasing deformation they broaden and become denser by addition of more lines. Lines parallel to $(1\bar{1}\bar{1})$ occur increasingly in the still empty regions of the surface. They combine with the former by the mechanism of cross-slip, which occurs frequently in this orientation, also on a smaller scale between slip lines of the same system. Kinkbands are found but rarely.

In orientations on the connecting line from [100] to [111] in the stereographic triangle, typical transition phenomena appear, such as the occurrence of bands of secondary slip in the empty zones between the clusters of slip lines on the principal system(s). Crystals elongated in a [111]-orientation differ, in turn, appreciably from the others in their surface appearance. Again the slip lines occur in clusters, but their boundaries are not parallel to the glide lines, as in the former cases.

The clusters are a few tenths of a millimeter wide, they contain very many narrowly spaced lines of the $(11\bar{1})$ [101] system. Kinkbands are again relatively rare in this orientation: kinkbands and striae seem to be mutually exclusive.

With increasing deformation the density of slip lines increases; this has been studied quantitatively by MADER [1957].

In body-centred cubic metals the slip direction is always a body diagonal; however, the slip planes are less well defined. This results in so-called *pencil glide*, and is due to the simultaneous presence of several distinct crystallographic planes which are nearly equally densely packed. The slip lines on these metals therefore often show a wavy character, in contrast with the long straight slip lines on the surfaces of face-centred cubic metals and hexagonal metals. The latter invariably show basal plane slip, in some cases (if the axial ratio $c/a > \sqrt{3}$) together with slip on the prism faces. The slip direction is invariably along a second order diagonal of the basal plane (the close packed direction).

The wavyness of the slip lines in b.c.c. metals has recently been shown to be intimately connected with the occurrence of *cross-slip* (see section VIII–1) (LOW and GUARD [1959], BRANDON and NUTTING [1959]).

The discovery of a fine structure of narrow lines (*"slip band"*), where at low magnification only a single line seems to be present was the most important result of *electron microscopical* investigations of the surface of a deformed crystal. In general a higher resolution leads to the observation of a much more homogeneous deformation pattern than could have been expected from an optical investigation alone. WILSDORF and KUHLMANN-WILSDORF [1953], using an electron-microscope and a very careful replica technique, have discovered the occurrence, in the very beginning of the deformation, of a fairly *homogeneous* background pattern of parallel, closely packed slip lines, about a few tenths of a mm in length, and representing steps on the surface of only 30–100 Å height. This *elementary structure* (see for instance Figs. 4 to 6) occurs in all pure close-packed cubic metals studied, but not in α-brass and other alloys that show a tendency to order. The macroscopic slip lines only represent a modulation of this structure.

Very interesting results have been obtained by studying the dependence of the slip line pattern on the amount of deformation in face-

centred cubic crystals. Figs. 5 to 8 illustrate the change from "fine slip" to "coarse slip" in a copper single crystal. In combination with mechanical measurements a distinction between three *deformation stages* has been carried through. The discussion of this is postponed to Chapter VIII.

The electron microscope has shown the optical slip bands to consist of a number of undeformed glide lamellae, separated by slip steps. With progressing deformation the slip lines gradually become shorter, their length varying from about 1000 micron after one or two percent strain to about 10 microns after 20% strain. The amount of slip per band also increases (from 10 to 500 Å) as does the number of lamellae per band (2 to 100). Finally a state is reached in which strong, heavy, often fragmented slip bands (Fig. 8) occur, and deformation bands also incidentally appear.

Plastic deformation need not take place exclusively by slip, but under circumstances where the slip process is especially unfavourable, it can proceed by *twinning* or by *kinking*. Twinning is a kind of homogeneous shear in which two crystal halves come into a mirror orientation with respect to each other. It is also a dislocation process but differs from slip in a number of aspects. For its occurrence usually a higher shear stress than for slip is needed. Usually twinning takes place in body-centred cubic or less symmetric metals at low temperatures, in hexagonal metals and occasionally in face-centred cubic metals; deformation by compression favours twinning. We refer to Chapter XIII for a further discussion.

Hexagonal crystals compressed along a direction parallel to the basal plane cannot deform in slip, because no suitable slip planes are available (see below). Under certain conditions kinking rather than twinning can achieve the necessary deformation. This process can best be compared to the formation of a giant deformation band, and its description will also be postponed to Chapter XIII.

We shall not discuss here in further detail the different optical and geometrical features of deformation, and refer the reader for this to the paper by SEEGER [1958] in the Encyclopedia of Physics. We will rather turn our attention to the mechanical characteristics of the phenomenon of plastic glide, the most important of which are the low value of the critical resolved shear stress and the shape of stress-strain curve of metals.

2. Critical Shear Stress for Slip in Metals and Stress-Strain Relation

The relationship between stress and strain of a pure metal single crystal is represented very schematically by the curve in Fig. 9. At very low stress levels the relation is linear and reversible, that is, on removing the stress, the strain also disappears. We are in the *elastic* region and Hooke's law applies, at least approximately. On increasing the strain the proportionality with the stress gradually vanishes (non-linear elasticity) and the reversibility is no longer perfect any more: a region of *anelasticity* occurs. We shall return to this in Chapter XVII.

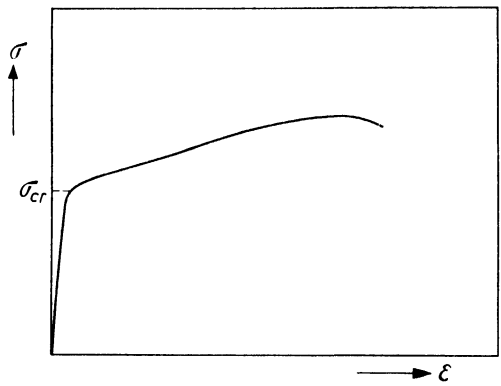


Fig. 9. Schematic stress-strain curve of a pure metal.

Applying a stress in the anelastic region for a *very long time* often results in a (slight) *permanent* deformation, thus in a kind of plastic strain. This is a form of *creep*. On increasing the stress further, ordinary *plasticity* sets in, and large permanent deformations are produced instantaneously, which remain after removal of the stress. The slope of the stress-strain curve rapidly decreases, although it usually remains positive. That is, a certain amount of *work-hardening* occurs. The shape of the curve remains similar up to the stress level where fracture becomes imminent (Chapter XXII), and consequently the slope becomes undefined. The actual shape of the stress-strain curve of a particular metal specimen depends primarily on the nature and crystallographic orientation of the deforming stresses, on the structure of the specimen (mono- or polycrystalline), on the degree of purity, the temperature and

the rate of straining. A sharp transition from elastic to plastic behaviour usually does not exist, although for many purposes a certain *critical stress* for the beginning of plastic flow is often defined (e.g. where the plastic strain exceeds a certain very small value). It depends, of course, on the orientation of the slip systems and the applied stress.

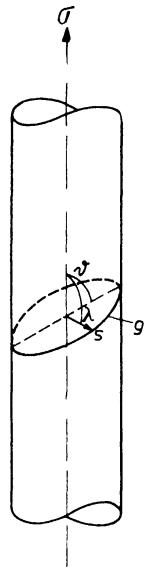


Fig. 10. On the definition of the resolved shear stress: g is the glide plane, s is the slip direction; the tensile stress σ acts along the cylinder axis.

It is possible to define with much less ambiguity a critical *resolved shear stress*, above which ordinary plasticity sets in in the form of slip on a *given* slip system. Although the *applied* critical stress for the beginning of slip may vary, on resolving the stress components it is observed that the critical *shear* stress in the slip plane along the slip direction is approximately the same. Suppose a single crystal bar of material is extended along its axis by a tensile stress σ (Fig. 10). Let the slip plane make an angle ϑ with the axis of the bar, and the slip direction an angle λ. Then the resolved shear stress in the slip direction is

$$\tau = \sigma \sin \vartheta \cos \lambda , \tag{1}$$

and it is only when *this* quantity exceeds a certain critical value, that

slip begins. This "law of constant resolved shear stress" was first enunciated by SCHMID [1935]. No effect at all of the stress component normal to the slip plane was observed. This finds a natural explanation in the crystallographic theory of the glide process. The shear stress needed to shear two perfect crystal parts over one another on a given lattice plane has been computed by FRENKEL [1926]. Assuming (as in Chapter IV) a sinusoidal relation to exist between the shearing force between two atom rows and their relative displacement, this *theoretical* critical shear stress is

$$\tau_{th} = \frac{b}{a}\frac{G}{2\pi}, \tag{2}$$

where b is the period of the shearing force (atomic distance in the shear direction) and a the spacing between the rows. Later refinements of Frenkel's calculation (MACKENZIE [1949] unpublished, see SEEGER [1958]) in which a less simple but more realistic approximation to the force function was used, and possibly stable, intermediate positions of two lattice planes were also considered, resulted in a value about five times smaller. That is, $\tau_{th} \approx G/30$ in copper and other noble metals.

Now, the *observed* critical shear stress in pure well annealed single crystals of face-centred cubic metals is many orders of magnitude

TABLE 2

Critical resolved shear stress for slip in crystals at room temperature (BARRETT [1952])

Metal and structure		*Impurity content* %	*Critical shear stress* g/mm^2
Cu	f.c.c.	0.1	100
Ag		0.01	60
Au		0.01	90
Ni		0.2	580
Mg	h.c.p.	0.05	85
Zn		< 0.001	150
Zn		< 0.04	95
Cd		0.004	80
β–Sn	tetragonal	0.01	130
Bi	rhombo-	≈ 0.1	220
Hg	hedral	≈ 10^{-6}	7
NaCl	sodium	≈ 0.1	200
AgCl	chloride	≈ 0.1	100

smaller, viz. of the order of $10^{-5}-10^{-4}\,G$, as shown in Table 2. (The slip system is always the principal system of Table 1.)

This, together with the appearance of slip lines etc. on the surface of a crystal and the law of the critical resolved shear stress, have led to the theoretical assumption that slip does not occur *simultaneously* over the whole slip plane, but first nucleates somewhere, and subsequently propagates under the action of the stress over the slip plane. This then happens in such a way that at each moment only relatively few atoms near to the boundary between the slipped and unslipped zones of the crystal are in energetically unfavourable positions. All other atoms remain more or less perfectly periodically arranged. This boundary of course represents a linear lattice imperfection, which moreover can wander easily across the slip plane. An imperfection with these properties was first devised by DEHLINGER [1929] and called "Verhakung". A few years later, G. I. TAYLOR [1934], OROWAN [1934] and POLANYI [1934] independently proposed a closely related imperfection, that possesses better mobile properties. This imperfection was the edge dislocation discussed in Part I. J. M. BURGERS [1939] completed the discussion by inventing the complementary imperfection, the screw dislocation, that could perform the same tasks as the edge dislocation. The hypothesis of the existence of these two types of dislocations (and of their intermediate forms) enabled others to explain the various features of crystallographic slip, as we shall discuss in the next section.

3. Dislocation Theory of Glide

The propagation of a dislocation over its glide plane results in the shearing of the two crystal halves on both sides of this plane with respect to each other. The amount of strain produced by a piece of dislocation that has swept across an area A was found to be (section III–5.1):

$$\varepsilon = \frac{Ab}{V}, \tag{3}$$

where V is the volume of the crystal. Fig. 11 illustrates this. Dislocation glide motion occurs only under the action of a stress. According to section III–2 a *force* is exerted on the dislocation by a shear stress τ in the direction of b per unit length of dislocation, which is simply

equal to τb. However, a certain resistance to the glide motion is present; as discussed in Chapter III and IV this is due to a number of causes.

a) The *Peierls force,* arising from the periodicity of the atomic positions on both sides of the slip plane. The magnitude of this fundamental force resisting dislocation motion in actual crystals has been

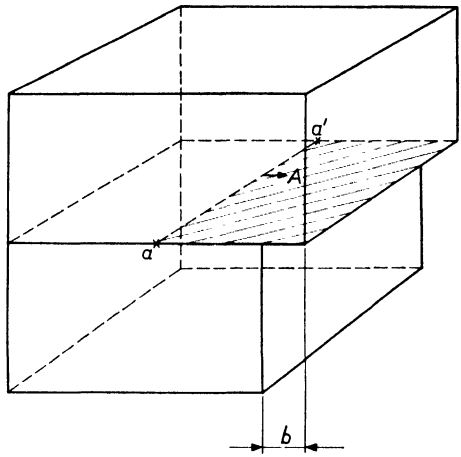


Fig. 11. A dislocation with Burgers vector b, originally at the position of the dotted line, has moved out of a unit cube of material, thereby sweeping over the hatched area A of the glide plane. The fraction $A/1$ of the horizontal cross section of the cube has then suffered the shear displacement b. Averaged over the cube (volume $V = 1$), the average shear strain is $(A/V) \cdot b$.

calculated accurately only for face-centred cubic metals (DIETZE [1952]). Under ordinary circumstances this force is barely comparable to or even considerably smaller than that corresponding to the observed critical shear stress. There are notable exceptions, however. Already mentioned previously was the indication from internal friction experiments that in copper at very low temperatures the Peierls force is no longer negligible at all. In the hexagonal metal zinc (and in some other metals also) $(10\bar{1}1)$ slip is not found at room temperature, but occurs only above 250 °C; this is because the Peierls force on that plane is too large for it to be operative as a glide plane at low temperatures (GILMAN [1956]). The occurrence of new slip systems at elevated temperatures in other metals is presumably due to the same cause.

b) The *elastic interaction* between dislocations causes a frictional

force that can be estimated as follows. Two parallel dislocations at a distance r exert a force upon each other that can be expressed as an equivalent shear stress of approximate magnitude

$$\tau_{\text{int}} \approx \frac{\gamma G b}{r},$$

where γ is a numerical factor of order 0.1.

Denoting the dislocation density by N, one has on the average

$$r = \sqrt{\frac{1}{N}}.$$

Thus the mean internal shear stress level is

$$\tau_{\text{int}} \approx \gamma G b \sqrt{N}.$$

In *undeformed* metals, N lies somewhere between 10^6 and 10^8 cm^{-2} and then $\tau_{\text{int}} \approx 10^{-5} G$. This is again of the order of magnitude of the observed critical shear stress. In deformed metals it becomes much greater, and determines part of the work-hardening.

c) During the very beginning of the plastic deformation of pure metals, possible other (*geometrical*) factors that hinder dislocation motion are relatively unimportant, at least at not too low temperatures. It will thus be clear that the small observed value of the critical shear stress for the beginning of plastic flow in pure metals can be satisfactorily explained in terms of the dislocation mechanism of slip.

d) In impure materials the interaction between impurity atoms and dislocations may well be much more important than any of the interactions discussed above. We return to this in Chapter IX, and limit ourselves here to the case of pure metals. It must be noted, however, that even the purest materials do contain *in principle* enough impurities to "fill up" completely all dislocation lines present in the undeformed crystal. A dislocation density of 10^8 cm^{-2} means that about 10^{16} atomic sites per cm^3 occur in the immediate vicinity of a dislocation line. These can, in principle, all be occupied if impurities are present to a concentration not larger than 10^{-4} atomic per cent. The highest degree of purity that can be obtained by using the recently developed zone-melting techniques (PFANN [1958]) amounts to about the same value. Of course, usually a small fraction of the impurity atoms present actually occur on dislocation sites, and therefore the impurity-dislocation interaction is less intensive than would follow from the above

estimate. Various recent researches seem to indicate that in most metals impurities begin to loose their influence on plastic flow and related phenomena below a minimum concentration of about 10^{-3} atomic per cent.

To explain the large amounts of plastic flow observed in practice under moderate stresses, dislocation *sources* must be present, which become operative at a stress of the order of $10^{-4}\,G$ or less. The Frank-Read source, discussed in section III–6, begins to operate at the shear stress Gb/l, where l is the length of the source. When $l = 10^4\,b$, or, for instance in copper, when $l = 3 \times 10^{-4}$ cm, the above condition is indeed fulfilled. When the Frank-Read sources are segments of the dislocation network present in the undeformed crystal, this length would correspond to a mean dislocation density of $10^7 - 10^8$ cm^{-2}, which agrees well with the actual density in well-annealed metal crystals.

However, it is rather doubtful whether indeed the dislocation network plays such a large role in the generation of new dislocations under stress. In those instances where the network can be observed (by decoration), changes of importance were not observed to occur during the beginning of plastic flow, whereas in several cases the newly generated dislocations were found to come from the surface, or even from sites where nothing particular had been detected before. The question of the dislocation sources, their nature and the magnitude of the critical shear stress stands in the focus of attention during the time of writing of this book (1958), it is still not completely solved. The interior Frank-Read source as such does undoubtedly exist, but also the surface acts as a prolific source of dislocations, and at present the latter seems to be more important than the dislocation network (GILMAN and JOHNSTON [1957]).

It is also a matter of doubt whether the actual observed critical shear stress in metals is determined by the critical stress for *source operation*, or by the resistance to *dislocation motion*. From the results of experiments of PARKER and collaborators [1952] about the mobility of small-angle grain boundaries under a shear stress (see Chapter XIX), one is tempted to prefer the latter alternative. The dislocations making up such a boundary, which were very long compared to the average source length in the metal (zinc), yet needed a stress of the same order as the observed critical shear stress of the crystals in order to become mobile. SEEGER *et al.* [1957] advocated the theory that in face-centred cubic metals the mutual elastic interaction between dis-

locations, especially the interaction between the moving dislocations and the original network, determines the critical shear strength of the metal. We refer for a further discussion to Chapters VII and VIII.

4. Work-hardening

In all metals work-hardening takes place, that is, the flow stress of a deformed metal increases with the amount of strain. The explanation of this all-important phenomenon in terms of the decreasing mobility of dislocations is the fundamental problem of the dislocation theory of crystal plasticity.

Again, one or more of the processes enumerated in the foregoing section must be responsible for the work-hardening. The Peierls force can be excluded, because it is too small to explain the often very high rates of work-hardening (the flow stress increases by factors 100–10000 during the first 100% of strain) and secondly because the Peierls stress must be independent of the prestrain. The increase of work-hardening with increasing strain points to a dependence on the dislocation density. The latter quantity is also observed to increase very rapidly with the amount of strain. It rises from $10^6 - 10^8$ cm^{-2} in undeformed metals to a few times 10^{11} cm^{-2} in heavily worked metals, and saturates at about 10^{12} cm^{-2} (HIRSCH [1957]). This increase can easily be observed by X-ray methods or microscopic examination of the surface (see also Chapter XVI) and is one of the most remarkable effects of the deformation.

The mutual elastic interaction of dislocations obeys the condition that it increases with the dislocation density; in first approximation it is about proportional to the square root of the density. In heavily cold-worked metals, with a dislocation density of $10^4 - 10^5$ times that in undeformed metals, the flow stress should therefore be 100–1000 times higher than the original critical shear stress. Ratios of this order of magnitude are indeed encountered in practice. Therefore, an important part of the work-hardening is undoubtedly due to this elastic interaction. Nevertheless, a detailed study of the work-hardening characteristics of metal single crystals as a function of strain, temperature and orientation, together with careful investigation of the surface markings, shows that more complicated interactions also play a significant part in the general work-hardening phenomenon.

Presumably most important of these socalled "*geometrical*" effects is

ıe mutual encounter of (nearly) parallel dislocations moving on dif-
rent intersecting slip planes, resulting in a combined form of disloca-
ɔn that is immobile (*sessile*). In section IV–5.1 we have already
entioned a possibility for the formation of such a *barrier* in face-
ıntred cubic crystals. A Frank-Read source on a slip plane where a
ırrier occurs can emit only a limited number of dislocation loops,
'en under increasing shearing stress, because the loops cannot move
rther from the source than the barrier permits, and will *pile up*
;ainst it. Long-range elastic stresses are associated with such a
slocation pile up, and it can be shown that at the site of the source
ıese stresses counteract the activating shear stress, rapidly causing
ıe latter to decrease below its critical value. Large stresses also occur
ı other slip planes in the neighbourhood of the pile-up, and these
nit the dislocation mobility on other slip systems (MOTT [1952]).

In the neighbourhood of piled-up groups of dislocations the lattice
curved, as illustrated in Fig. 12 (for the case of pure edge dislocations).
his curvature can easily reach observable magnitudes; for instance,
hen 100 parallel identical dislocations occur on a space of 5000
omic distances, the lattice planes are bent over an angle of a few
:grees. A similar curvature of the slip planes is observed in the
nk bands shown in Fig. 5. This surface phenomenon is indeed directly
lated to the existence of piled-up groups of dislocations (section
II–1).

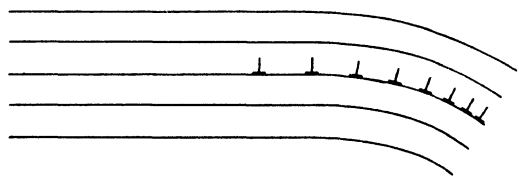


Fig. 12. Curvature of lattice planes near a pile up of edge dislocations. The thin lines denote the extra half planes associated with these dislocations.

A third important factor in the work-hardening of metals is the
utual intersection of *non-parallel* dislocations. This usually results
the formation of jogs which must move non-conservatively together
ith the parent dislocations (section III–5.3), and in the subsequent
rmation of point defects such as vacancies and interstitial atoms.
ɔr both these processes an activation energy is needed that depends
ı the applied stress. In the absence of appreciable thermal activation,

the necessary shear stress to move the jogged dislocation must therefore increase. This process usually becomes of importance only in the later stages of work-hardening, when the dislocation density has increased so much that intersections occur frequently (after, say, 20–50% of strain).

Not only can dislocations on intersecting slip planes produce obstacles for dislocation motion, but impurities and irregularities on a scale larger than a few atoms can also act as effective obstacles for slip. Quite generally such obstacles are very effective slip inhibitors because of their twofold action: restriction of dislocation paths, and closing up of sources.

All inhomogeneities in the crystal structure of metals act as obstacles for dislocation motion. Very efficient inhibitors are also grain boundaries in polycrystalline metals, which can even act as "trapping" centres for moving dislocations. On having arrived at such a boundary, the dislocation is sometimes "dissolved" in it and lost to the further deformation process. Often pile ups form, which, having reached a certain extent, promote the propagation of the slip on the other side of the boundary, that is, into the next grain.

Polycrystalline samples generally present an appreciably higher critical shear stress than single crystals in favourable orientations. This is due to the fact that it is not possible to orient a polycrystalline specimen in such a way that slip occurs in every crystallite only on the system where the resolved shear stress is highest. In order to accomodate the strains of the various crystallites, usually two or more, up to five or even six slip systems must be activated in every grain. This of course necessitates higher applied stresses, and also produces much faster rates of work-hardening.

It will be clear from the foregoing that not only the low critical shear stress, but also the work-hardening of metal crystals can qualitatively be explained by the dislocation hypothesis. In the next two chapters we shall discuss this hypothesis and its effects more quantitatively.

The presence of large numbers of dislocations and point defects in plastically deformed metals does not only show itself in the mechanical properties but also in the physical properties. For instance, the electrical resistivity increases appreciably, especially so in pure metals at low temperatures where the thermal and residual resistivities are small. Also several forms of internal friction of metals become much

more apparent after plastic strain, and the diffusion of foreign atoms through the metal, as well as the rate of self-diffusion, may be appreciably exhanced. Chapters XIV to XVIII will be especially devoted to a quantitative discussion of these phenomena.

REFERENCES

BARRETT, C. S. (1952); Structure of Metals (McGraw-Hill, New York).

BLEWITT, T. H., R. R. COLTMAN and J. K. REDMAN (1955); Report of Conference on Defects in Solids, Bristol, p. 389 (London, The Physical Society).

BRANDON, G. G. and J. NUTTING (1959); Acta Metall. **7**, 101.

BROWN, A. F. (1952); Advances in Physics **1**, 427.

BURGERS, J. M. (1939); Proc. Kon. Ned. Acad. Wet. Amsterdam **42**, 293 and 378.

CAHN, R. W. (1951); J. Metals **79**, 129.

CAHN, R. W. (1954); Advances in Physics **3**, 363.

CHEN, N. K. and R. MADDIN (1956); Progress in Metal Physics **5**, 53.

DEHLINGER, U. (1929); Ann. d. Physik **5**, 2 and 749.

DIETZE, H. D. (1952); Z. Physik **131**, 156.

ELAM, C. F. (1935); Distortion of Metal Crystals (Oxford, Clarendon Press).

FRANK, F. C. (1952); Advances in Physics **1**, 91.

FRENKEL, J. (1926); Z. Physik **37**, 572.

GILMAN, J. J. (1956); Trans. AIME **206**, 998.

GILMAN, J. J. and W. G. JOHNSTON (1957), in: Dislocations and the Mechanical Properties of Crystals, p. 117 (Wiley, New York).

HAASEN, P. and G. LEIBFRIED (1955); Fortschritte der Physik **2**, 79.

HIRSCH, P. B. (1957); Progress in Metal Physics **6**, 236.

JAOUL, B. (1957); J. Mech. Phys. Solids **5**, 95.

KOCHENDÖRFER, A. (1941); Plastische Eigenschaften von Kristallen (Springer, Berlin).

LANGE, H. and K. LÜCKE (1953); Z. Metallk. **44**, 183.

LEIBFRIED, G. and P. HAASEN (1954); Z. Physik **137**, 67.

LOW, J. R. and W. GUARD (1959); Acta Metall. **7**, 171.

MACHLIN, E. S. (1957), in: Dislocations and the Mechanical Properties of Crystals (Wiley, New York).

MADER, S. and A. SEEGER (1959); Acta Metall. **8**, 513.

MÜLLER, H. and G. LEIBFRIED (1955); Z. Physik **142**, 87.

MOTT, N. F. (1952); Phil. Mag. **43**, 1151.

OROWAN, E. (1934); Z. Physik **89**, 614 and 635.

PARKER, E. S. and J. WASHBURN (1952); Trans. AIME **194**, 1076.

PFANN, W. G. (1958); Zone melting (Wiley, New York).

POLANYI, M. (1934); Z. Physik **89**, 660.

SCHMID, E. and W. BOAS (1935); Kristallplastizität (Springer, Berlin).

SEEGER, A. (1958); Encyclopedia of Physics, Part VII–2, p. 1.

SEEGER, A., J. DIEHL, S. MADER and H. REBSTOCK (1957); Phil. Mag. **2**, 323.

STAUBWASSER, W. (1954); Dissertation Göttingen University, unpublished.

SUITS, J. C. and J. R. LOW (1957); Acta Metall. **5**, 285.

TAYLOR, G. I. (1934); Proc. Roy. Soc. **A145**, 362.

WILSDORF, H. and D. KUHLMANN-WILSDORF (1953); Acta Metall. **1**, 394.

WYON, G. and J. M. MARCHIN (1955); Phil. Mag. **46**, 1119.

WYON, G. and P. LACOMBE (1955); Report of Conference on Defects in Solids, Bristol, p. 187 (London, The Physical Society).

CHAPTER VII

PRODUCTION OF IMPERFECTIONS DURING PLASTIC DEFORMATION

1. Sources under Stress in Single Crystals

An undeformed or well-annealed metal crystal is believed to contain a spatial network of dislocations (sections III–6 and VI–3). On applications of a stress, the network elements can, in principle, act as Frank-Read sources, because their end points are more or less fixed. In section VI–3 it has already been pointed out that often not the network elements, but other structures, the nature of which is yet unknown but which are probably connected with the surface, act as sources. Presumably these structures are also dislocation elements or loops, and for them holds the same as for the network elements, viz. that only those elements will emit dislocations for which the resolved shear stress along their Burgers vector exceeds the critical value (III–16):

$$\tau_{cr} = \alpha G b / l \,, \tag{1}$$

where l is the length of the element and α a factor of order unity. Operation as a source occurs if no other restrictions exist which inhibit source action. Primarily a suitable orientation of the Burgers vector of the source element is necessary. In face-centred cubic metals most dislocations have presumably ⟨110⟩ Burgers vectors (see Table III–1), and thus, when a uniaxial stress is (locally) present, in general about 1/6 of the source elements will be favourably oriented, (assuming an arbitrary distribution of source elements). Then, the longest elements will start earliest, subsequently shorter sources will be activated later and new orientations will become active in subsequent stages.

Inhibition of the source action can occur either through the presence of other dislocations and imperfections, or by immobilization of the original source element by, say, impurity atoms.

The critical shear stress for source action is always enhanced by the presence of other dislocations, because these inhibit the free expansion of the Frank-Read source elements. Also the mobility, that is, the flow stress, is decreased. The interaction between impurity atoms and dislocations also often results in an increase of the critical shear stress. In contrast to the former case, however, *once* it has broken loose, the dislocation segment can move under a far smaller stress than was needed to mobilize it. This results in the observation of a so-called upper yield point. We shall postpone the discussion of this phenomenon to Chapter IX.

Thermal vibrations may help a dislocation to pass the obstacles mentioned, or to liberate itself from its anchoring atoms. Appreciable thermal activation will occur only if, in order of magnitude, $U \leqslant 10\,kT$, where U is the height of the obstacle (of atomic dimensions) to be overcome (U may be, for example, the difference between the binding energy of the impurity atom and the line tension of the free dislocation, or the energy to form a jog). Questions of thermal activation will also be treated in Chapter IX. The action of a "pure" Frank-Read source is of course independent of temperature, at least if the Peierls stress can be neglected.

Each loop of radius R emitted by a source exerts a *back stress* on the latter, which can easily be computed (e.g. by the method of Peach and Koehler, section III–2). It is:

$$\tau_b = \gamma Gb/R\,, \tag{2}$$

where

$$\gamma = \frac{1}{2\pi}(1-\nu)/(2-\nu) \approx 0.1$$

in metals. Although the back stress decreases when the loop expands, it nevertheless prevents the continuous operation of the source. The latter will only emit its dislocation loops at separated time intervals: when the resultant stress at the site of the source has again attained a sufficiently large value. For this the previously emitted loop must have covered a distance δ such that

$$\gamma Gb/\delta \ll \tau_{cr} = \alpha Gb/l\,,$$

or

$$\delta \gg \frac{\gamma}{\alpha} l\,. \tag{3}$$

Remembering that $\gamma/\alpha \approx 0.1$, for purposes of computation we can write (3) as

$$\delta \geqslant l . \tag{4}$$

Let the expansion velocity of the dislocation loops be v — about this quantity not much more is known in metals than that it is rather high, of the order of 10^3—10^4 cm/sec; in polar and homopolar crystals the velocity varies with stress and temperature over many factors of 10. Then, owing to the back stress alone, every Frank-Read source possesses an inherent minimum delay period θ which is, in order of magnitude, given by

$$\theta \approx \frac{l}{v} . \tag{5}$$

In metals this time is usually very short, about 10^{-7} sec, but in polar and homopolar solids delay periods of many minutes can under circumstances be observed (see Chapter XXX).

The operation of Frank-Read sources results in an increase of the dislocation density of the material; owing to the intersection-and-moving-jog mechanism discussed in Chapter III also the density of point defects increases. It has been tried by a number of authors to derive quantitative expressions for the increase of the imperfection density with stress or strain. Such expressions must necessarily depend sensitively on the details of the *model* which is adopted for the description of the shape of the moving dislocation line, the mode of formation of point defects, the existence or not of internal barriers, etc. A discussion of the merits of these models cannot yet be given at this stage. Because of the great importance of a quantitative theory for the production of defects during plastic flow, especially when the influence of the latter process on physical properties is considered, we shall, however, mention already in this chapter some of the results obtained so far.

Dislocations

1) The simplest model imaginable is that of a number of independent Frank-Read sources, which emit concentric circular loops. The latter expand until they meet the surface. As long as the influence of the surface can be neglected, each loop contributes proportionally to its area to the strain, whereas of course the dislocation length increases propor-

tionally to the radius. Therefore, *per loop*

$$d \sim \varepsilon^{\frac{1}{2}}, \tag{6}$$

where d stands for the newly generated dislocation density and ε for the plastic strain due to that loop. Since the number of loops increases with the time, and the expansion does not go on indefinitely, this formula does not apply to the material as a whole. Taking also the production of point defects into account and balancing the strain energy gained by expanding against the energy stored in the form of ever more lattice imperfections, VAN BUEREN [1955], [1957] arrived at the following formula for the overall dislocation density D in the early stages of deformation:

$$D = 10^6 (D_0 N)^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} \text{ cm}^{-2} \tag{7}$$

(a number of uninteresting factors and constants have been collected together in the numerical factor 10^6, which applies to copper). D_0 is the density of the *dislocation forest* or *network* already present before the deformation, and N is the volume density of sources. This formula applies to slip on one system only. When multiple slip occurs to such an extent that the newly formed dislocations begin to intersect each other (stage III, see Chapter VIII), instead of (7) the linear relation

$$D = 10^8 N^{\frac{1}{2}} \varepsilon \tag{8}$$

should be used†.

2) HAASEN and LEIBFRIED [1955] and FRIEDEL [1956] considered the presence of fixed barriers in the material, which confine the expansion of the loops to a certain maximum distance L. Also the number of loops per source is then limited because the back stress exerted by all the emitted loops together ultimately stops the source. This model ("fixed barrier model") leads to a linear dependence of the dislocation density on the strain:

$$D = \frac{2}{Lb} \varepsilon . \tag{9}$$

3) SEEGER [1957] finally considered the case that the limiting distance L depends on the density of dislocations. He argued that, as the bar-

† I am indebted to Miss Jenifer Lomer for pointing this out to me.

riers are formed by the mutual interaction of these imperfections, they will certainly become more numerous as the deformation proceeds ("growing barrier model"). Basing himself on electron microscope observations of the length of slip lines in copper (see for instance SEEGER *et al.* [1957] and Chapter VI), Seeger took:

$$L = L_e \text{ is constant for edge dislocations;}$$
$$L = L_s = \frac{\Lambda}{\varepsilon - \varepsilon^*}, \ (\varepsilon > \varepsilon^*) \text{ for screw disl.,} \tag{10}$$

where $\Lambda \approx 3 \times 10^{-4}$ cm (observed value) and ε^* amounts to a few per cent. Considering straight dislocation lines, one then arrives at the formula for the density of dislocations:

$$D = \frac{\varepsilon - \varepsilon^*}{L_e b} + \frac{(\varepsilon - \varepsilon^*)^2}{\Lambda b}. \tag{11}$$

Owing to the decrease of the slip length with increasing deformation, the dislocation density increases faster than proportionately to the strain.

Inserting numerical values applying to copper in all these formulae ($D_0 = 10^8$ cm^{-2}, $N = 10^{10}$ cm^{-3}, $b = 2.5 \times 10^{-8}$ cm) the following (approximate) theoretical dislocation densities are obtained:

TABLE 1

Dislocation density as function of strain

	Energy balance (cm^{-2})	*Fixed barriers* (cm^{-2})	*Growing barriers* (cm^{-2})
$\varepsilon = 1\%$	10^8	10^8	10^8
$\varepsilon = 10\%$	5×10^8	10^9	3×10^9
$\varepsilon = 100\%$	$3 \times 10^{10} - 10^{11}$	10^{10}	10^{11}

As is well-known, after heavy deformation the dislocation density saturates in the neighbourhood of 10^{12} cm^{-2} in metals.

Point defects

Quantitative predictions regarding the production rate of point defects are in an even less satisfactory state. Not only the model description of dislocation generation and motion determines this rate,

but also the mechanism of formation of point defects: the moving-jog mechanism, or a more static process such as the combination of two edge dislocation segments of opposite sign on neighbouring slip planes, resulting in a row of point defects (FRIEDEL [1956], see Chapter III). Variants on the recombination mechanism have been proposed, but these are certainly of less importance and will not be discussed here.

According to the model and the mechanism, we can distinguish between several theoretical formulae. To van Bueren's model the moving-jog mechanism is inherent. For each expanding loop the number of jogs is proportional to the number of intersected dislocations, thus to $D_0 A$, where A is the area covered. The number of vacancies or interstitials left in the wake of these jogs is again proportional to the radius radius of expansion (after averaging). Therefore, *per loop*

$$f \sim \varepsilon^{\frac{1}{2}} , \tag{12}$$

where f stands for the number of point defects formed. Considering again also the formation of new loops and the energy balance, VAN BUEREN [1956] arrived at the following expression for the volume density F of point defects in the early stages of deformation:

$$F = 10^{16} \left(\frac{D_0^3}{N}\right)^{\frac{1}{4}} \varepsilon^{\frac{1}{2}} \text{ cm}^{-3} . \tag{13}$$

When multiple slip occurs, this should be replaced by

$$F = 10^{21} \varepsilon^2 . \tag{13a}$$

In the fixed barrier model, not only the dislocation density but also the concentration of point defects increases proportionately to the strain, as both are proportional to the area covered. Assuming the moving-jog mechanism to apply, one arrives at

$$F = \frac{L D_0}{b^2} \varepsilon , \tag{14}$$

(L being again the ultimate extension of the loops). When recombining dislocation segments should be the most important source of point defects, one has instead:

$$F = \frac{l}{L^2 b^2} \varepsilon , \tag{15}$$

where l is the length of the recombining parts. This is presumably of the order of the source length.

In Seeger's growing barrier model it makes very much difference which mechanism of point defect formation applies. To analyse this complicated case here would carry us too far, the result may yet be mentioned:

moving-jog mechanism → logarithmic increase:

$$F = \frac{\Delta D_0}{b^2} \log (\varepsilon - \varepsilon^*) \; ; \tag{16}$$

recombination mechanism → quadratic increase:

$$F = \frac{l}{2\Lambda L_e b^2} (\varepsilon - \varepsilon^*)^2 \, . \tag{17}$$

In Table 2 numerical data derived from these formulae have been collected (again for the metal copper).

TABLE 2

Rate of formation of point defects

	Energy balance	*Fixed barriers*		*Growing barriers*	
	(cm^{-3})	*Moving jogs* (cm^{-3})	*Recombination* (cm^{-3})	*Moving jogs* (cm^{-3})	*Recombination* (cm^{-3})
$\varepsilon = 1\%$	5×10^{17}	10^{19}	10^{14}	10^{18}	10^{15}
$\varepsilon = 10\%$	10^{19}	10^{20}	10^{15}	10^{19}	10^{17}
$\varepsilon = 100\%$	10^{21}	10^{21}	10^{16}	10^{19}	10^{19}

Note added in proof

A recent theoretical study by SAADA (Thesis Paris 1960), based on a careful consideration of the energetic changes involved when two dislocations, which interact elastically and which may be extended and liable to recombination (see section III. 5), intersect, leads to a general relation between the production rate of defects and the coefficient of work-hardening—at least when a model theory of the latter process as proposed by Friedel (see next Chapter) is adopted. If the coefficient of work-hardening is proportional to ε^n, the production rate should be proportional to ε^{n+1}. For polycrystals, where generally $\sigma \sim \sqrt{\varepsilon}$, this relation yields the $\frac{3}{2}$ power law; in single crystals with linear hardening a quadratic powerlaw for the vacancy concentration is obtained, as was indeed observed by BLEWITT *et al.* [1954] (see Chapter XV).

It will be clear that the recombination mechanism, at least in the form proposed by Friedel, yields much smaller defect concentrations than the moving-jog mechanism. We shall see in Chapter XV that observations show a much greater density to be present. In this light, we tend to favour the moving-jog mechanism.

The numerical differences between the various models based on this mechanism in Tables 1 and 2 are but slight, although the arguments by which the numbers were derived are greatly different. This gives us some reason to trust the numerical data in the two tables, in as far as the moving-jog mechanism is concerned.

Arguments can be, and have been, advanced to show that one should not compute defect density as a function of strain, but of *stress* (that is, flow stress) because it is the latter that is determined by the state of imperfection of the lattice. This argument is true as it stands. However, most measurements so far have been done as a function of the easily measurable variable quantity strain (see Chapter XV), and, secondly, it is not known very clearly in what quantitative manner the flow stress is determined by the state of imperfection.

The study of this is the study of the phenomenon of work-hardening, which is the topic of the following chapter.

REFERENCES

BLEWITT, T. H., R. R. COLTMAN and J. K. REDMAN (1954); Report Conf. on Defects in Solids, Bristol, p. 369 (The Physical Society, London).

BUEREN, H. G. VAN (1955); Acta Metall. **3**, 519; (1957) Philips Research Rep. **12**, 1.

ESHELBY, J. D., F. C. FRANK and F. R. N. NABARRO (1952); Phil. Mag. **42**, 351.

FRIEDEL, J. (1956); Les dislocations (Gauthier-Villars, Paris).

HAASEN P. and G. LEIBFRIED (1955); Fortschritte der Physik **2**, 79.

SEEGER, A. (1954); Z. Naturforschung **9a**, 758, 856 and 870.

SEEGER, A. (1957); Dislocations and the Mechanical Properties of Crystals, p. 243 (Wiley, New York).

SEEGER, A., J. DIEHL, S. MADER and H. REBSTOCK (1957); Phil. Mag. **2** 323

CHAPTER VIII

WORK-HARDENING OF FACE-CENTRED CUBIC CRYSTALS

1. Interaction between Dislocations

Since it is not the purpose of this book to duplicate already existing papers on the theory of plastic deformation of metals, we shall only discuss briefly the various ideas put forward to explain the phenomenon of work-hardening in metals. Attention will be limited to face-centred cubic crystals only, for these have been studied most extensively. Work-hardening is principally due to the interaction between dislocations: as the strain increases, the number of dislocations increases and the interaction becomes more important. It then becomes increasingly difficult to effect further deformation.

One of the most obvious interaction phenomena that can be observed directly under the microscope, is the formation of kinkbands (Fig. VI–2). These have already been discussed in Chapter VI; the bands are nearly perpendicular to the slip system to which they belong. In them, the glide lines are curved like a flattened S. They are visible most clearly on the sides of the deformed crystal both perpendicular to the glide planes and parallel to the glide direction. Their presence is due (MOTT [1952]) to the mutual encounter of large numbers of dislocations of opposite sign on neighbouring slip planes, which cannot pass each other for some reason (Fig .1). From the observed curvature of the glide planes and their distance, the number of *piled-up* dislocations in each plane can be derived. If the distance between glide planes is d and the number of piled-up dislocations is ν per unit length, there are ν/d dislocations per cm^2. These, if all pure edge dislocations, produce a local curvature

$$\frac{1}{R} = \frac{b\nu}{d}. \tag{1}$$

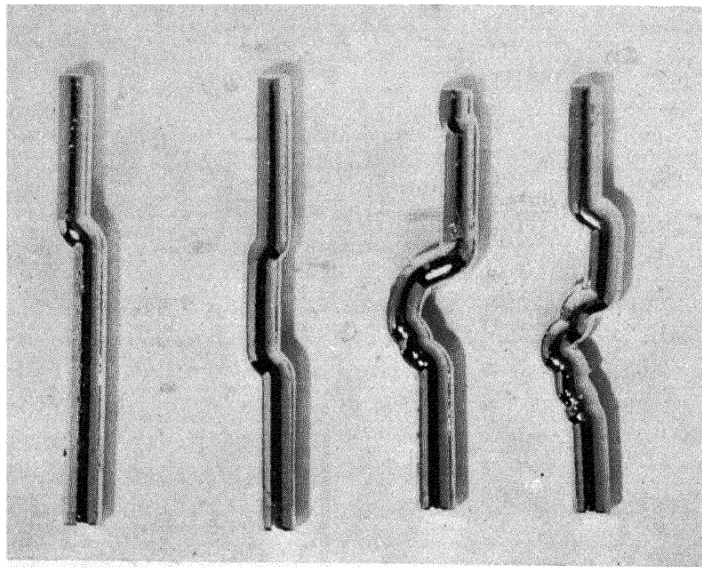


Fig. 2*a*. Kinks in zinc monocrystals (after Gilman).

Using values representative for aluminium after a medium amount of strain, ν is found of the order of a million per cm. Since a kinkband is usually about 10^{-3} cm wide, the number of piled-up dislocations per glide plane is then of the order 1000.

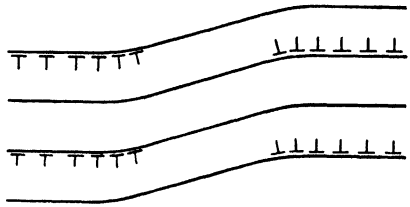


Fig. 1. Dislocation model of a kinkband (viewed in the glide plane perpendicular to the glide direction).

The origin of deformation bands must lie in the presence of barriers of some kind (the most probable of which are dislocations in an intersecting glide plane), which interact by their long range stresses with the original dislocations, causing the latter to be trapped in the forward as well as in the reverse direction (MOTT [1952]). LOMER [1951] and COTTRELL [1952] have pointed out that extended dislocations on intersecting glide planes can produce locks in a purely geometrical manner, at least in close-packed structures, by the formation of so-called Cottrell-Lomer sessile dislocations (section III–5.1). This is in accord with the observation that in single crystals kinkbands only form when multiple slip takes place, and mostly appear simultaneously with the slip lines on the secondary system(s). In hexagonal metals kinkbands of this type do rarely occur (because of the difficulty of glide on other than basal planes), but on compression the related phenomenon of *kinking* is observed (Fig. 2*a*). Such a kink is bounded by walls of edge

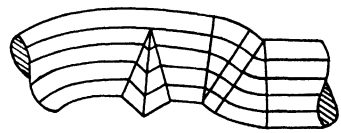


Fig. 2*b*. Structure of kink boundaries.

dislocations resembling boundaries, see Chapter XIX (Fig. 2*b*). The glide planes are not bent continuously everywhere, but change their orientation sharply.

Another very important mechanism whereby the motion of dislo-

cations is impeded, is the mutual intersection of non-parallel dislocations. This has already been discussed in Chapters III and VII. A complication arises when the dislocations are extended, as then intersection is possible only after a *constriction* in the stacking fault ribbon has occurred (Fig. III–10). For the formation of such a constriction a certain additional energy is required, which was computed by SCHOECK and SEEGER [1955] to be considerably larger in copper than in aluminium. The difference is connected with a corresponding difference in the surface tension of the stacking fault: the latter is high in aluminium (a large fraction of Gb^3 = a few eV) and consequently the extended dislocations are narrow (of the order of a few atomic spacings) and a constriction is easily formed. In copper and especially in gold the stacking fault energy amounts to one hundredth or even less of Gb^3, that is a few tenths of an eV, and the extended dislocations are therefore wide, 3 to 20 atomic spacings. To form a constriction in these is difficult. More is said about this in section 5.

In face-centred cubic crystals the formation of a constriction also plays a dominating role in the phenomenon of *cross-slip* of extended dislocations. An extended screw dislocation in a $(\bar{1}11)$ plane can, under circumstances, change over to the intersecting $(1\bar{1}1)$ plane, for instance if on the latter a sufficiently large shear stress component acts along the (invariant) direction of the Burgers vector [110] (Fig. 3). For this change-over to occur, first the dislocation line must become fully constricted over a certain length l. Cross-slip will be therefore a far

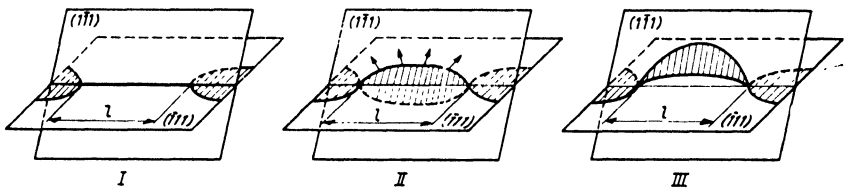


Fig. 3. Cross-slip of extended screw dislocation according to SEEGER [1957]. I, II and III are successive steps of the process. First of all a constriction must be formed over the whole length l.

more probable process in aluminium than in the noble metals. FRIEDEL [1957] has pointed out that it should not be necessary to have the dislocations constricted over the full length of the cross-slipping portion simultaneously: the cross-slip can occur as a *wave* along the dislocation (Fig. 4). The activation energy is now dependent on the stress and can

be computed as follows (WOLF [1960]). First the difference of the energy of the cross-slipped portion of dislocation (Fig. 3-II) and that of the non-slipped part is derived as a function of the stress on the cross-slipplane and of the length l. Then this difference is minimized with respect to this length: the minimum distance from where the "Friedel wave" can start is the length where this minimum is reached, since from then on any enlargement of the cross-slipped portion decreases the energy of the system. The corresponding minimum energy difference is the activation energy for cross-slip; it depends on the stress according to

$$U \approx \frac{0.35Gb^3}{1 + 180\gamma/Gb} \log \{\sigma/\sigma_3(0)\}, \tag{2}$$

where γ is the stacking fault energy and $\sigma_3(0)$ a parameter, viz. the stress where cross-slip should occur without *any* thermal activation; it can also be computed and derived experimentally from the work-hardening curve at very low temperatures (see section 5). In metals where $180\gamma/Gb$ is larger than unity the activation energy depends on the stacking fault energy; these metals are called "high-s.f.-energy" metals. In them, the activation energy is relatively small, since the stacking fault is narrow; aluminium is an example of such a metal.

TABLE 1

Stacking fault energy and width in metals (mostly after BERNER [1960].)

Metal	*Stacking fault energy* γ ergs/cm²	*Stacking fault width (in atomic distances)*	$180\,\gamma/Gb$
copper	160	1.5— 3.5	2.2
gold	10	11 —36	0.22
nickel	300	1.1— 1.9	2.9
stainless steel	20	50	0.18
α-brass	30	20	0.2
aluminium	238	1.0— 1.8	5.5
lead	100?	2?	—

In low-s.f.-energy metals where the ratio $180\gamma/Gb$ is smaller than unity the activation energy is large as there the stacking faults are wide, but the value of γ does no longer determine the energy. Gold is an example of this class of metals. In Table 1 values of the s.f.-energy

and width in various metals are given, together with the ratios mentioned. It is seen that copper forms in this aspect some sort of average case. Its properties do most resemble those of the low-s.f.-energy metals, though.

Cross-slip occurs when dislocations are held up before a barrier. A pile up forms, the dislocations in front of it are already partly compressed by the stress due to the others, and cross-slip will be initiated.

Cross-slip is responsible for the sudden change from one slip plane (octahedral plane) to another, as observed from the course of the slip lines on the faces of moderately deformed copper crystals (Fig. VI–7) and for the occurrence of *fragmented* slip bands in heavier deformed crystals (Fig. VI–8). The latter phenomenon probably reflects the occurrence of *double cross-slip*, a mechanism in which the cross-slipped portion of a dislocation acts as a source of a number of new dislocation loops on a slip plane parallel to that on which the dislocations moved. The fundamental role played by the stacking fault energy in cross-slip has been stressed by SEEGER and his coworkers in a number of papers [1955, 1957].

The differences in stacking fault energy between cubic metals are closely related to corresponding differences in the energy of a coherent twin boundary. This becomes clear when it is realized that in the face-centred cubic lattice a stacking fault can be regarded as a twin which is only a few atom layers thick. From the twin boundary energy, if known, the stacking fault energy can thus be estimated; it is roughly two times the latter. An accurate value for this energy is only obtained after a more elaborate calculation, however.

The differences in stacking fault energy between the various metals and alloys are largely determined by the influence the fault has on the energy of the conduction electrons. The ion-ion interaction is practically the same in the two close-packed configurations, the more so as only distances between next-nearest neighbours are affected by the presence of the fault.

It might therefore be expected that metals with a large electron-to-atom ratio would possess a higher stacking fault energy than those with a small ratio. In accordance herewith is that aluminium as a trivalent metal has a much larger stacking fault energy than the monovalent noble metals and that recently a more or less parabolic dependence on the electron-atom ratio has been established in a series of metals (CAHN, private communication). However, alloying copper with

multivalent elements like zinc, beryllium or germanium to obtain Hume-Rothery phases invariably results in a smaller fault energy. Evidently the change in Brillouin zone structure owing to the inclusion of foreign atoms overcompensates the increase of the electron-atom ratio.

2. Stress-Strain Curve of Face-Centred Cubic Single Crystals

It follows from the work started by Masing and coworkers in Göttingen and extended by DIEHL and coworkers [1955, 1956a,b] and many others (see CLAREBROUGH and HARGREAVES [1959]), that the

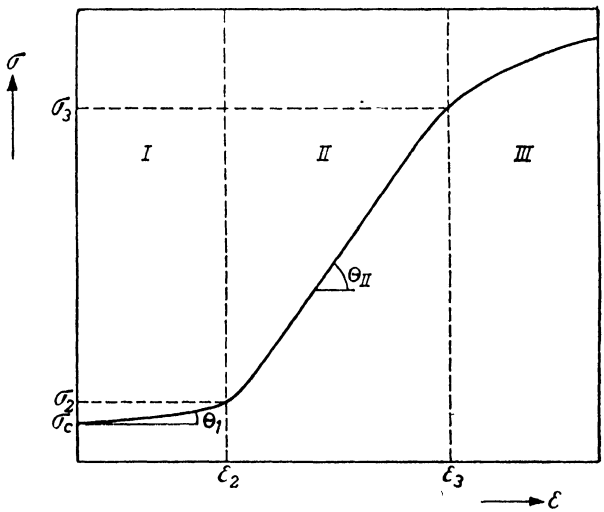


Fig. 4. Schematic stress-strain curve of face-centred cubic metal. I is the easy glide region, II the linear-work-hardening region and III the temperature dependent part.

stress-strain curve of f.c.c. metal crystals can be divided into three parts (Fig. 4). A linear initial part with a relatively small slope is followed by a much steeper, also approximately linear part, which finally makes place for a less steep, but strongly temperature dependent part, ending in fracture.

The first part is often called the *easy glide* region or *region I*, its extension depends strongly on the orientation of the crystal with respect to the axis of the deforming stress. From this and the observed

slip line pattern (long straight parallel slip lines) it can be unambiguously concluded that easy glide is characterized by the activity of only *one* of the twelve possible {111} ⟨110⟩ glide systems (BROWN and HONEYCOMBE [1951], SEEGER [1957]). Impurities present in small amounts tend to extend the easy glide region, presumably by still further inhibiting slip on less favourably oriented systems.

Kinkbands begin to occur about at the level of strain corresponding to the first bend in the stress-strain curve (ε_2). This, together with other micrographical evidence (the occurrence of shorter slip lines, in several directions), and the rapidly increasing rate of work-hardening, points to the occurrence of at least one second glide system in *region II*. The slope θ_{II} of this second part of the stress-strain curve is remarkably independent of temperature, impurity content etc., and SEEGER [1957] has pointed out that this part reflects especially clearly the properties of a dislocated lattice as such. In contrast, the form and slope of the third part of the stress-strain curve depend markedly on the temperature of deformation. According to SEEGER and coworkers [1956] it starts when *cross-slip* of screw dislocations, which is, as we have seen, a temperature-activated process, begins to take place. The number of kinkbands also increases very rapidly. In favour of this assumption we can adduce the simultaneous occurrence of large coarse slip bands, which indeed can easily form by this mechanism. The space in between these bands becomes ultimately filled with fine short slip lines. Local stress concentrations, leading to excessive local slip and fracture, occur in the ultimate parts of the stress-strain curve. This region does not lend itself to a detailed physical description.

In Table 2 the dependence of the principal parameters of the work-hardening curve of face-centred cubic crystals (as indicated in Fig. 4) upon temperature, orientation of a uniaxial loading stress and impurity content, are presented. This table is based largely on the work of Seeger and coworkers in Stuttgart. The variations quoted have been corrected, where necessary, for simultaneous variations of the elastic constants.

Note the relative invariance of $\sigma_2-\sigma_0$; even if the rates of work-hardening in regions I or II are changed, the relative extent in the stress scale of the easy glide region remains the same.

Although the general form of the flow curve is the same for low- and high-stacking fault energy cubic metals, there are differences, especially in the temperature dependent part. In a metal like aluminium with a

high stacking fault energy, cross-slip already starts at a relatively low stress and temperature, much lower than, say, in the noble metals. The second region is even often practically absent in aluminium. Only at very low temperatures does it become apparent (see Fig. 9).

TABLE 2

Variation of parameters of work-hardening curve of f.c.c. metal single crystals

Parameter (see Fig. 5)	*Temperature increases:*	*Orientation of stress goes from centre to edge of stereogr. triangle:*	*Substitutional "impurity" content increases:*
critical shear stress: σ_0	- -	slightly +	strongly +
stress at which region II begins: σ_2	—	slightly +	+
$\sigma_2 - \sigma_0$	0	0	+
stress at which region III begins: σ_3	strongly −	slightly +	+
strain at which region II begins: ε_2	—	strongly −	very strongly + (overshoot)
strain at which region III begins: ε_3	−	−	strongly —
work-hardening rate in region I: θ_I	0	strongly +	strongly —
work-hardening rate in region II: θ_{II}	0	+	0
work-hardening rate in region III: θ_{III}	−	0	?

In the hexagonal metals zinc, magnesium and cadmium the flow curve is usually fairly linear over the entire range (LÜCKE *et al.* [1955], DERUYTTERE and GREENOUGH [1956]). This is especially so at low and at high temperatures. In the intermediate temperature range the curve sometimes shows a fairly sharp bend at a strain of about 1 or 2. Not much dependence on orientation exists, but the rate of work-hardening

varies strongly with the temperature, at room temperature and above. At low temperatures the stress-strain curve becomes temperature independent and can presumably be compared directly with the easy glide part of the flow curve of cubic crystals. The temperature dependence above room temperature probably arises from the simultaneous occurrence of recovery during the plastic deformation, in the form of dislocation *climb* (SEEGER and TRAUBLE [1960]).

The work-hardening curve of body-centred cubic metals is as a whole strongly temperature-dependent, and complicated because of the interference of twinning at low temperatures.

We shall now discuss the theory of work-hardening of face-centred cubic crystals in some detail.

3. Work-hardening in Easy Glide

The first theory of work-hardening is due to TAYLOR [1934]. It explains the phenomenon by the elastic interaction of dislocations on different (parallel) glide planes. Two parallel edge dislocations (Fig. 5) cannot pass each other and become stuck when the driving shear stress τ does not exceed the value $Gb/8\pi(1-\nu)d$, when d is the distance between the slip planes. When at a certain stress level τ the density of dislocations has increased to N, arranged in a square lattice, in which each dislocation has moved an average distance L before it became stuck, the plastic strain is

$$\varepsilon = NbL,$$

whereas, since $d \approx N^{-\frac{1}{2}}$, the shear stress opposing further motion is

$$\tau_w \approx \frac{Gb}{8\pi(1-\nu)} N^{\frac{1}{2}}. \tag{3}$$

Equating this stress to the applied shear stress τ (a dubious procedure in itself), one finds the following parabolic stress-strain relation

$$\tau = \frac{1}{8\pi(1-\nu)} G\sqrt{\frac{\varepsilon b}{L}}. \tag{4}$$

Now, it is a peculiar coincidence that until recent times such a parabolic relation was indeed usually observed (in aluminium). However, in the easy glide region, to which (4) should apply in the first place

(elastic interaction of parallel dislocations only) the stress-strain relation is, as we know, actually linear in close-packed crystals.

MOTT [1952] was the first to point out the all-important role that barriers for dislocation motion should play in work-hardening. He showed that a linear work-hardening law could indeed be obtained on the basis of a model in which *piled-up groups* of dislocations act as sources of internal stress. This idea has been elaborated upon by a number of authors, among whom Friedel and Seeger. The latter author (SEEGER [1957, 1958]) has recently proposed a very detailed theory

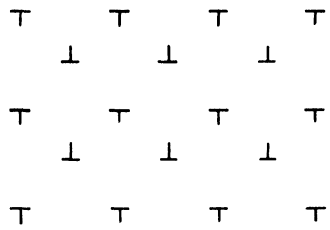


Fig. 5. On Taylor's model of work-hardening (TAYLOR [1934]).

of the work-hardening of face-centred cubic metals, and the following discussion will be largely based on this work. In the easy glide range the dislocations cover large distances on their {111} slip planes, and many of them reach the surface (long straight slip lines, Fig. VI–5). The work-hardening should be practically zero, if not barriers for the dislocation movement would gradually develop. These barriers are the so-called Lomer-Cottrell barriers discussed in Chapter III, that originate from the coalescence of two extended dislocations on intersecting slip planes. If the orientation of the crystal with respect to the stress axis is such that one slip system is greatly favoured over all the others, very few barriers are produced, the work-hardening rate θ_{I} is small and the easy glide region extends unto high strains. If the orientation is more at random, barriers form relatively early and easy glide extends only over a smaller range of strain. Observations show that even when no barriers are present, the work-hardening coefficient θ_{I} remains finite; this fundamental work-hardening is probably connected with the Peierls force.

The occurrence of Lomer-Cottrell barriers thus determines the extent and orientation dependence of region I. The activity of each dislocation source becomes limited by the formation of barriers in its glide plane, lying along one of the three possible ⟨110⟩ directions in

that plane (because they are formed by the interaction with dislocations on other {111} slip planes).

Let the number of loops emitted per source be n, and the mean area covered by them A (this is of the order of the square of the distance between the source and the barriers). If there are ν sources per cm^3 the plastic strain amounts to

$$\varepsilon = \nu n b A \,. \tag{5}$$

To obtain an estimate of the internal stress opposing dislocation generation and motion it should be realized that the dislocations are not distributed at random but occur in bundles of n each. Applying (3) to this situation, one has

$$\tau_w = \gamma G b n \sqrt{\nu A^{\frac{1}{2}}} \tag{6}$$

where γ is again a numerical factor of order 0.1. Hence we find the work-hardening coefficient θ_I in the easy glide range to be:

$$\theta_I = \frac{d\tau_w}{d\varepsilon} = \gamma G (\nu A^{\frac{1}{2}})^{-\frac{1}{2}} \,. \tag{7}$$

It seems reasonable to assume that ν and A are both fairly independent of stress and temperature, as they are largely determined by the geometry of the initial dislocation distribution. Therefore, indeed a constant rate of work-hardening is obtained in this model. Numerical agreement with observed values of θ_I — which are of the order of $10^{-3}\,G$ (see Figs. 8 and 9) — can be reached by putting: $\nu A^{3/2} \approx 10^4$. From the appearance of his electron micrographs MADER [1957] has concluded that ν should not exceed the (very small) value 10^8 cm^{-3} in aluminium; this would imply a mean path of the dislocations in region I of the order of 1 mm. From the observed strain, n, the number of loops per source would then follow as about 10, in accordance with the observed stepheight of only a few tens of Ångström units of the elementary slip lines on slightly deformed crystals (Fig. VI–5). The agreement between this simple model and the observations is thus to be regarded as satisfactory.

4. Work-hardening in Region II

As the deformation proceeds, the number of barriers increases, and ultimately each source will become completely *surrounded* by barriers,

as illustrated in Fig. 6. For this, slip must have taken place on at least three systems, viz. that of the sources considered and on two different {111} planes. The old argument that easy glide ends where multiple slip starts has now got a more defined meaning: easy glide ends when multiple glide has occurred to such an extent that the glide motion of dislocations becomes restricted in *every* direction. It is not surprising that the orientation of the crystal with regard to the axis of the applied stress determines to a large extent the shear stress σ_2 at which region II begins. GARSTONE, HONEYCOMBE and GREETHAM [1956] found that

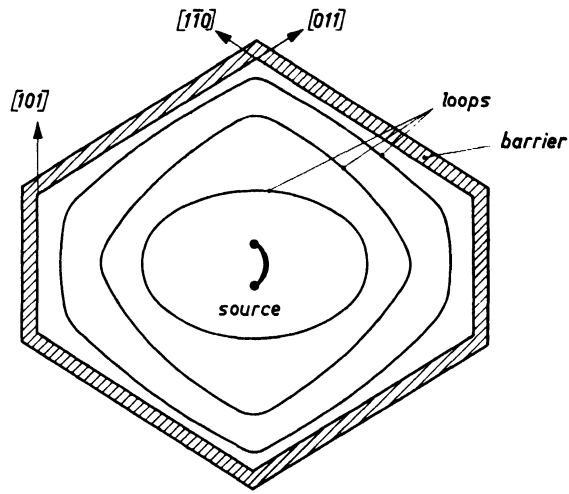


Fig. 6. Frank-Read source surrounded by Cottrell-Lomer barriers. This illustrates the situation in stage II (after SEEGER [1957]).

in pure copper and aluminium single crystals the easy glide range ends when the difference between the applied *resolved* shear stress on the slip system and the critical resolved shear stress reached a value which is *independent* of the orientation of the crystals (see also Table 2). Taking the latter stress as that necessary to activate a Frank-Read source (which is dubious), these authors can indeed explain their observations by assuming that easy glide ends when the stress is large enough to activate sources on the secondary slip systems. They concluded that these pile-ups must then have attained a length of about half the distance between the barrier and the nearest source.

According to DIEHL and REBSTOCK [1956], during the second part of

the work-hardening curve the mean path L of the successively released dislocation loops *decreases*, according to the formula:

$$L = \frac{\Lambda}{\varepsilon - \varepsilon^*} ; \tag{8}$$

in which ε^* is a parameter slightly smaller than ε_2, the strain at which region II begins, and Λ is of the order of 10^{-3} cm. This decrease of the slip length was deduced from measurements of the length of microscopically visible slip lines in copper and nickel, and was interpreted as the result of the continuous formation of *new Cottrell-Lomer barriers*.

From the observed slight dependence on temperature of the flow stress in region II it can be deduced that again the elastic interaction of dislocations determines the work-hardening (see also section 5). If the number of loops per source of diameter L is *constant* n, the density of sources ν',

$$\tau_w = \gamma G b n \sqrt{\nu' L} . \tag{9}$$

If $\mathrm{d}N$ new dislocations move through the distance L, the increment of strain is

$$\mathrm{d}\varepsilon = bL^2 \mathrm{d}N = \frac{b\Lambda^2}{(\varepsilon - \varepsilon^*)^2} \mathrm{d}N \tag{10}$$

(with the aid of (8)). By differentiating equ. (9) with respect to L one can, with (10), express $\mathrm{d}\tau_w$ in terms of $\mathrm{d}\varepsilon$, and one finds for the work-hardening coefficient in stage II:

$$\theta_{II} = \gamma G \sqrt{\frac{bn}{3\Lambda}} . \tag{11}$$

In practice $\theta_{II} \sim 4 \times 10^{-3}\, G$, and $\Lambda \approx 4 \times 10^{-4}$ cm, which gives $n \approx 25$, in accordance with observations (section VI-1).

Observed is a *constant* rate of work-hardening. This can be understood by assuming that the number of piled-up dislocations per source varies but insignificantly during stage II. Here a difference exists between this theory and a theory of work-hardening proposed at about the same time by FRIEDEL [1957]. This author also sees as the principal cause of work-hardening in stage II the obstruction by Lomer-Cottrell barriers. However, Friedel assumes these barriers to be formed all together at the beginning of stage II (when multiple glide starts). After this the number of barriers remains fixed, and their positions too.

Thus the path of the dislocations remains invariant, only the quantity n, the number per pile-up, increases continuously, in contrast to Seeger's theory. This increase being proportional to the stress (according to the simple formula (7)) and the strain being porportional to n, a linear hardening law is obtained again. To explain the small work-hardening coefficient of the order of $10^{-2}\,G$, Friedel must in his theory take into account the "reduction" of the back stress of dislocations on their own source, through the influence of dislocations emitted by sources on the other glide planes. This reduction must amount to not less than 98% to explain the observations.

It is not possible yet to decide between the two theories. Objections to Friedel's theory include the inherent instability under stress reversal of his dislocation distribution, and the observed inverse proportionality of the slip line length with the strain. Against Seeger's theory can be advanced the consideration that the assumption of a constant number of dislocations per source seems at first sight rather improbable, although experiments confirm that the depth of slip lines does not change much during stage II.

5. Third Region of Stress-Strain Curve. Influence of Temperature on Critical Shear Stress and Flow Stress

In the first as well as in the second stage of the stress-strain curve of face-centred cubic crystals, work-hardening is observed to be slightly dependent on the temperature, apart from the trivial changes caused by the temperature variation of the elastic constants (PAXTON and COTTRELL [1954], ADAMS and COTTRELL [1955]). The third part of the curve, however, shows an appreciable temperature effect, the rate of hardening decreasing with increasing temperature. This temperature dependence points to the influence of obstacles that must be passed by the dislocations with the aid of *thermal fluctuations*. Such obstacles are predominantly Cottrell-Lomer barriers that must be passed either by *cross-slip* or by sheer force (break through).

The former process is the more probable; it has a stress-dependent activation energy given by formula (2). The meaning of the parameter $\sigma_3(0)$ now becomes apparent: it is the stress at which region III would commence at the absolute zero of temperature. At finite temperature the transition stress $\sigma_3(T)$ is found as follows. The frequency of elementary cross-slip processes is given by $\nu = \nu_0 \exp(-U/kT)$; hence

the strain rate due to the motion of N dislocations which move over an area A between two cross-slips is

$$\dot{\varepsilon} = NAb\,\nu_0\,\mathrm{e}^{-U/kT}. \tag{12}$$

Reversing this formula and inserting (2) one finds for the strongly *temperature dependent* flow stress in region III, at a given strain rate:

$$\sigma_3 = \sigma_3(0)\,\exp - \left\{\frac{\log\,(\dot{\varepsilon}/NAb\nu_0)}{0.35\,Gb^3/(1+180\gamma/Gb)}\,kT\right\}. \tag{13}$$

This brings us to the discussion of the temperature dependence of the *critical shear stress* and the flow stress in *earlier* stages. SEEGER [1955] has shown that these quantities can be split into a temperature independent part τ_G which arises from the source properties and the long-range elastic interaction, and a temperature dependent part τ_S which should be caused by the necessary mutual intersection and defect formation during the motion of dislocations. As we have seen earlier (section III-5), an activation energy determines the rate of this process also, it is given by

$$U = U_a - \tau w.$$

The activation volume w is for extended dislocations equal to bdl_0. The rate of strain is now formally determined by

$$\dot{\varepsilon} = NAb\nu_0\,\exp\left\{-\frac{U_0 - v(\tau - \tau_G)}{kT}\right\}, \tag{14}$$

where ν_0 is the atomic frequency factor (about 10^{13} sec^{-1} in metals), A is the area covered by a segment of dislocation between two successive activations, N is the density of dislocations and τ_G is, as usual, given by

$$\tau_G = \gamma G b N^{1/2}. \tag{15}$$

Solving (14) for τ, one has

$$\begin{aligned} &\tau = \tau_G && \text{when } T \geqslant T_0 \\ &\tau = \tau_G + \tau_S; \quad \tau_S = \frac{U_0 - kT\log\,(NAb\nu_0/\dot{\varepsilon})}{bdl_0} && \text{when } T < T_0 \end{aligned} \tag{16}$$

where

$$T_0 = \frac{U_0}{k\log\,(NAb\nu_0/\dot{\varepsilon})}. \tag{17}$$

Thus the stress is made up by a practically temperature independent elastic part τ_G and a temperature dependent part τ_S; the latter is of importance only for $T < T_0$ (Fig. 7).

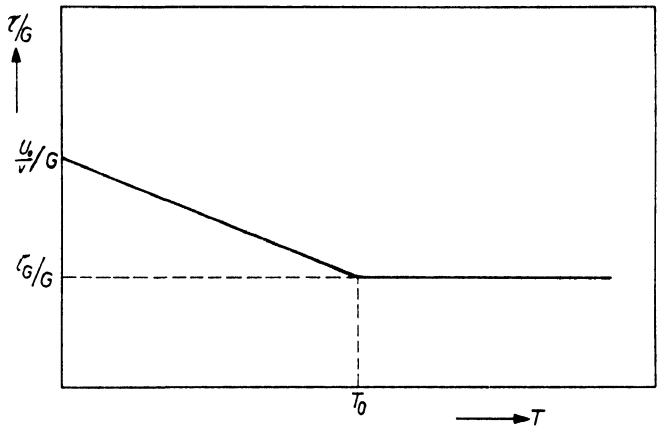


Fig. 7. Schematic dependence of flow stress on temperature for aluminium (high stacking fault energy). T_0 is the transition temperature of formula (17).

When the work-hardening in stage II has increased to such large values that thermally activated processes can begin to compete with unactivated slip, the rate of increase of the work-hardening curve lessens. We have now reached stage III, in which the temperature influences the stress-strain curve to a large extent (Figs. 8 and 9). Seeger coined for this relaxing influence of thermal activation aided by stress in stage III the term *dynamic recovery*. The influence of the applied stress in this recovery is essential: the activation energy U_0 is usually much too large to be overcome by thermal fluctuations alone (in contrast to static recovery, to be discussed in Chapter XI).

This activation energy is in the two cases we are interested in (intersection *and* cross-slip) the energy required to form a *constriction* in an extended dislocation. We have already seen earlier in this chapter that this energy is closely connected to the *stacking fault energy*. From the temperature variation of the stress-strain curve, an estimate can thus be made of the latter energy. SCHOECK and SEEGER [1955] obtained the following values for the constriction energy of an extended dislocation:

Cu, edge dislocation	3.9 eV;
Cu, screw dislocation	0.84 eV;
Al, edge dislocation	0.21 eV;
Al, screw dislocation	0.11 eV.

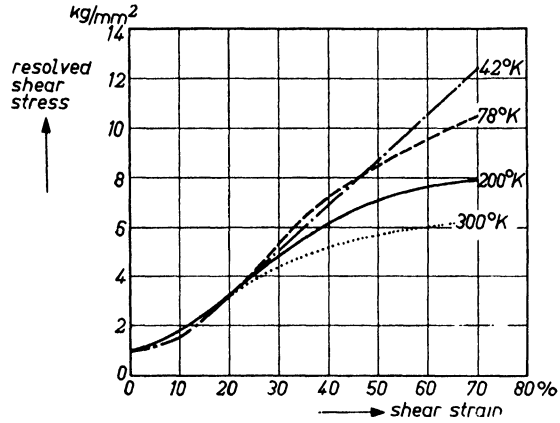


Fig. 8. Temperature variation of stress-strain curve of copper single crystals. Variations are present only in stage III, that occurs after a fairly large amount of strain (after SEEGER [1957]).

New evidence has shown the values for copper to be about 40% smaller; however, even then, for this noble metal and for the other noble metals also, the constriction energy is so large that intersection of dislocations or cross-slip can *never* be originated by thermal fluctuations alone. Every intersecting dislocation, whatever be its nature or orientation, thus acts as an obstacle for slip; the transition temperature T_0 is near to or above room temperature and the flow stress depends linearly on the temperature in stage III (formula (16)). Also, no cross-slip occurs at a measurable rate. This is indeed what is observed in copper and gold (Fig. 8). (Silver forms an exception, probably owing to the presence of an oxyde layer.) More about this will be said in section 6.

In aluminium, the constriction energy is much smaller and the critical temperature below which the flow stress decreases with temperature, is consequently much lower. Temperature dependence of the flow curve is already observable after slight amounts of strain (Fig. 9): the rate of work-hardening is now only independent of temperature at very low strain levels and begins to decrease with temperature

already after a few per cents of strain. That is, region II is practically absent in aluminium, and stage III begins much earlier than in copper. The transition temperature under normal conditions of deformation is about 200 °K in aluminium: this means that aluminium at liquid

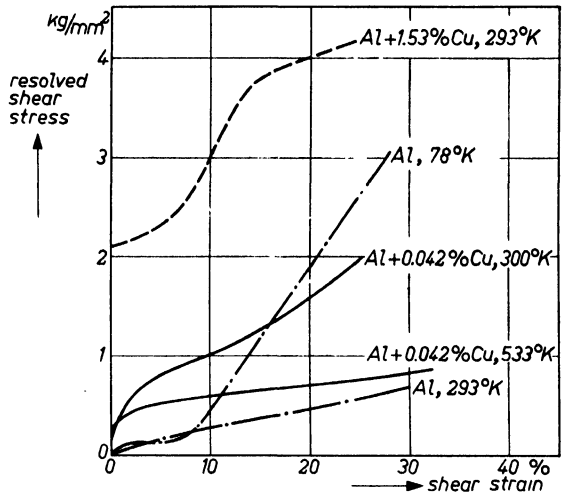


Fig. 9. Temperature variation of stress-strain curve of aluminium and aluminium-copper single crystals. Part II is practically absent above 78 °K.

air temperature does compare *very well to* copper at room temperature (Fig. 9). It should be mentioned that this difference is *not* connected with the difference of the melting points, but with the difference in stacking fault energies. Between copper at room temperature and copper at liquid air not much difference exists. Aluminium at room temperature and above shows the phenomena of cross-slip and fragmented slip bands very clearly. About the latter process and its occurrence in the noble metals also (although not through the intermediate of cross-slip) we must refer the reader to SEEGER's extensive paper in the Encyclopedia of Physics [1956]).

6. Deformation at Different Temperatures: Work-softening

It was observed by ADAMS and COTTRELL [1955] that copper and aluminium crystals when heavily worked at 90 °K, showed a considerable drop of the flow stress on further straining at a higher temperature,

i.e. 473 °K (Fig. 10). Also an apparent yield point (Chapter IX) developed. The phenomenon was called *work-softening* by COTTRELL and STOKES [1955].

These authors explained the effect as due to the release of piled-up groups, formed during the low-temperature deformation, by some thermally activated process only active at the higher temperature.

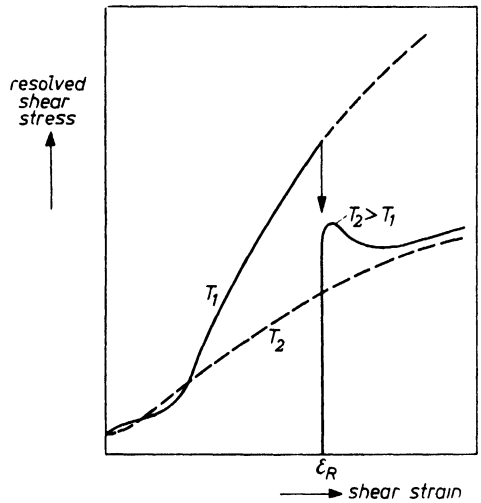


Fig. 10. Work-softening of face-centred cubic metal crystal. At the strain ε_R the temperature is *in*creased from T_1 to T_2. In stage III the flow stress then drops to a lower value, and an apparent yield point is observed (after SEEGER [1957]).

KELLY [1956] found that work-softening in aluminium is accompanied by an increase in the amount of cross-slip, and by an increase in slip line depth, in good agreement with Seeger's theoretical picture, in which the process mentioned *is* simply cross-slip. Adams and Cottrell could, by applying their *temperature cycling* method, study very accurately the *reversible* variations with temperature of the flow stress of *copper* single crystals at the end of stage II and in stage III. Correcting for temperature variation of the shear modulus, they obtained a flow stress against temperature curve consisting of three parts. In the first part, at temperatures (for copper) below about 200 °K, there was but little temperature dependence; in the second region (to about 300 °K) the flow stress fell linearly with T, to become again less temperature-dependent above 300 °K (Fig. 12).

(One should not confuse the *first* of these transition temperatures with the T_0 mentioned in section 5.)

From Seeger's theory discussed in the previous section we know already that a certain temperature dependence of the critical shear strength is to be expected as it is connected with the activation energy necessary to constrict extended dislocations when they intersect with "screw-like" dislocations. In copper this energy is so large that this dependence is simply linear.

SEEGER [1955] has attempted to explain the remarkable temperature dependence in this metal by remarking that owing to the large energy of formation of constrictions, the following three processes must be distinguished: jog formation in edge dislocations, the same in screw

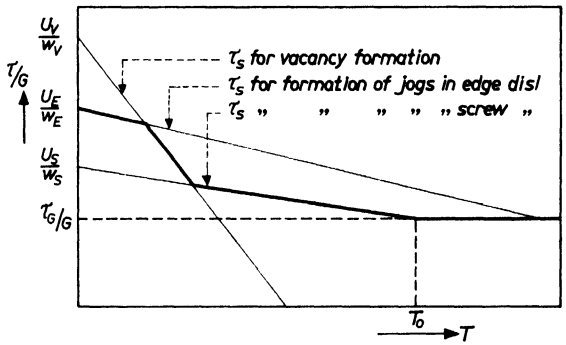


Fig. 11. Flow stress as a function of temperature of a metal with relatively narrow stacking faults. The third transition temperature is T_0 of formula (17).

dislocations, and the formation of point defects in the wake of the latter. As follows from (16), the slope of the τ_s *versus* T curve is determined by the quotient $\log(NAbv_0/\dot{\varepsilon})/w$; thus it is inversely proportional to w, whereas the intercept with the stress axis is given by U_0/w. It can be shown that w is about the same for the two processes of jog formation, but it is much smaller, of the order of one atomic volume, for the third. Consequently, the slope corresponding to defect formation is much larger than that due to the other two processes. Since also the ratio U_0/w is larger for the third process (and smallest for the second, since U_0 is smallest for screw dislocations), one obtains three different intersecting curves for the T-dependence of the flow stress, which combine to one broken curve as shown in Fig. 11. By comparison the experimental flow stress-temperature dependence

(Fig. 12) shows the remarkable similarity between the two shapes. It appears as if between 200 and 260 °K the flow stress in copper is for a large part determined by the formation of defects (see also DIEHL and BERNER [1960]).

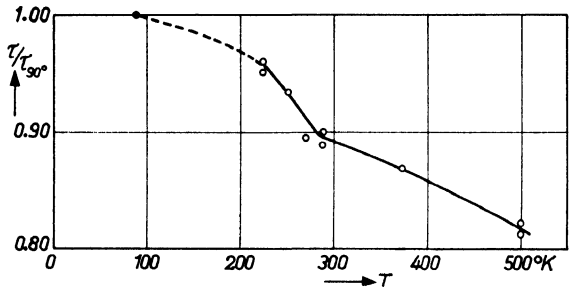


Fig. 12. Temperature variation of the flow stress of copper single crystals as determined by ADAMS and COTTRELL [1955].

Schematically a temperature dependence like that in Fig. 11 will hold for all f.c.c.metals with a low stacking fault energy, but not for aluminium and for the hexagonal metals. To aluminium does the simple theory of section 5 apply.

REFERENCES

ADAMS, M. A. and A. H. COTTRELL (1955); Phil. Mag. **46**, 1187.
ANDRADE, E. N. DA C. and C. HENDERSON (1951); Phil. Trans. Roy. Soc. **A244**, 177.
BERNER, R. (1960); Z. Naturf. **15a**, 689.
BROWN, A. F. and R. W. K. HONEYCOMBE (1951); Phil. Mag. **42**, 1146.
CLAREBROUGH, L. M. and M. E. HARGREAVES (1959); Progress in Metal Physics **8**, 1.
COTTRELL, A. H. (1952); Phil. Mag. **43**, 657.
COTTRELL, A. H. and J. STOKES (1955); Phil. Mag. **46**, 1213.
DERUIJTTERE, A. and P. B. GREENOUGH (1956); J. Inst. Metals **84**, 337.
DIEHL, J., S. MADER and A. SEEGER (1955); Z. Metallk. **46**, 650.
DIEHL, J. (1956a); Z. Metallk. **47**, 331; (1956b); id., 411.
DIEHL, J. and H. REBSTOCK (1956); Z. Naturf. **11a**, 169.
DIEHL, J. (1958); Diskussionstagung zur Metallplastizität (unpublished), Aachen.
DIEHL, J. and R. BERNER (1960); Z. Metallk. **51**, 522.
FRIEDEL, J. (1957), in: Dislocations and the Mechanical Properties of Solids, p. 330 (Wiley, New York).
GARSTONE, J., R. W. K. HONEYCOMBE and G. GREETHAM (1956); Acta Metall. **4**, 485.
KELLY, A. (1956); Phil. Mag. **1**, 835.
LOMER, W. M. (1951); Phil. Mag. **42**, 1327.
LÜCKE, K., G. MASING and K. SCHRÖDER (1955); Z. Metallk. **46**, 792.
MADER, S. (1957); Z. Physik **149**, 73.

MOTT, N. F. (1952); Phil. Mag. **43**, 1151.
PAXTON, H. W. and A. H. COTTRELL (1954); Acta Metall. **2**, 3.
SCHOECK, G. and A. SEEGER (1955); Report Conference on Defects in Solids, Bristol, p. 340 (London, The Physical Society).
SEEGER, A. (1955); Phil. Mag. **46**, 1194.
SEEGER, A. (1956); Encyclopedia of Physics, Part VII–1, p. 323.
SEEGER, A. (1957), in: Dislocations and the Mechanical Properties of Solids, p. 243 (Wiley, New York).
SEEGER, A. (1958); Encyclopedia of Physics, Part VII–2, p. 1.
SEEGER, A. and H. TRAUBLE (1960); Z. Metallk. **51**, 435.
TAYLOR, G. I. (1934); Proc. Roy. Soc. **A145**, 362.
WOLF, H. (1960); Z. Naturf. **15a**, 180.

CHAPTER IX

THE INTERACTION OF DISLOCATIONS WITH IMPURITIES

1. Interaction between Dislocations and Impurity Atoms

1.1. YIELDING

In section VII–1, when dealing with the critical stress needed to activate a Frank-Read source, it was already mentioned that this stress can be enhanced in practical situations by the presence of impurity atoms that block the dislocation segment. This blocking action is due to decrease of the free energy of the system when the impurity atom penetrates into the stress field of the dislocation, as has been explained in section III–4. Not only elastic, but also electrical (Coulomb) interaction and chemical interaction (formation of molecules or even a second phase) may contribute to such an energy decrease. Chemical as well as physical point defects are attracted by dislocations in metals.

The great difference between the influence of vacancies or of interstitial atoms of the matrix metal, and that of foreign atoms, is that the latter remain as such in the lattice whether bound to the dislocation or not, whereas the former can be completely absorbed (at jogs, section III–4), leaving the dislocation practically unchanged too. Therefore, single vacancies and interstitials do not necessarily inhibit the mobility of a dislocation very much whereas individual impurity atoms are very effective in this respect. In the plastic properties of crystals impurities thus play a very important part.

The most spectacular effects of impurities are found in the *initial* stages of plastic deformation. In many body-centred cubic alloys the first part of the stress-strain curve of a well-annealed specimen has the typical shape as depicted in Fig. 1. The metal shows a pronounced *yield point*: below a certain stress the deformation is elastic but above this stress (the upper yield stress) a finite amount of plastic deformation

takes place "suddenly". Partly owing to the "hardness" of the tensile machine the applied stress then drops to the level of the "lower yield stress". Ordinary work-hardening starts only after a certain amount of plastic strain has been achieved.

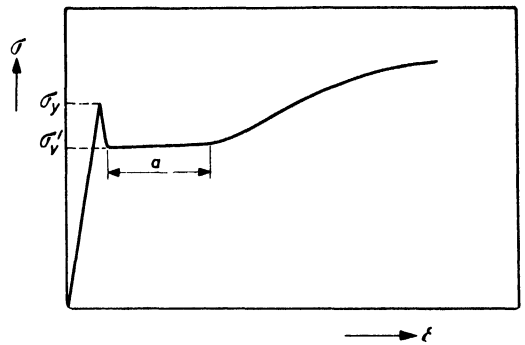


Fig. 1. Schematic flow curve of α-iron containing some carbon in solid solution. σ_y is the upper yield stress, σ'_y the lower yield stress and a denotes the Lüders band region ("Lüders strain").

The extent a of the "constant stress" portion of the curve, after the yielding has occurred, as well as the value σ_y of the yield stress itself, depend markedly on the temperature. PAXTON and BEAR [1955] observed in single crystals of α-iron the following values:

$T = 20$ °C: $a = \frac{1}{4}\%$; $\sigma_y = 5$ kg/mm²;

$T = -70$ °C: $a = 5$–15%; $\sigma_y = 15$ kg/mm².

In this portion of the curve so-called *Lüders bands* appear. These are zones where slip has occurred and slip bands have formed, which start at some place of high stress (e.g. near the grips) and *propagate* with a definite velocity through the test bar (Fig. 2).

These phenomena were explained by COTTRELL [1950] in terms of the immobilization of dislocations by impurity atoms. Cottrell's theory was based on the best-known practical case, that of body-centred iron-carbon alloys, containing carbon in solid solution. Yielding occurs when the stress is sufficient to *unpin* the dislocations. The sudden release leads to a catastrophic process in which barriers such as grain-boundaries can be overcome. The *lower* yield stress should be a measure for the strength of *these barriers*.

1.2. ELASTIC INTERACTION

COTTRELL, and COTTRELL and BILBY [1949], investigated in detail the interaction between solute foreign *interstitial* atoms and *edge* dislocations. They only considered the purely elastic interaction due to the hydrostatic stresses associated with the impurity (regarded as a compression centre, section IV–3) and with the dilatation field around the dislocation. The energy of interaction can be written as

$$U = A \sin \theta / r \,, \tag{1}$$

where r and θ are the cylindrical coordinates of the interstitial impurity with respect to the (straight) dislocation (in the direction of the Burgers vector $\theta = 0$, Fig. 3) and A is a factor of the dimension

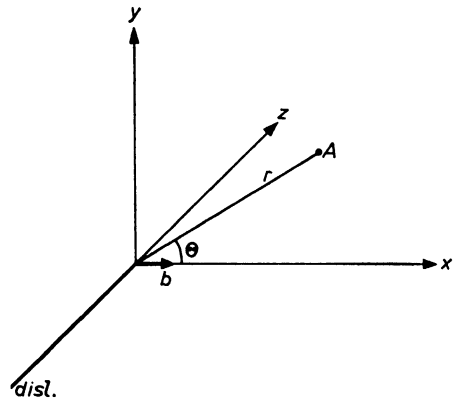


Fig. 3. On the elastic interaction of an impurity atom with an edge dislocation along the z-axis. The x-axis is in the direction of the Burgers vector $\boldsymbol{b}$; A is the impurity atom (in the xy-plane), r and θ denote cylindrical co-ordinates.

dyn · cm², which for this simple case can be written as (compare section III–4):

$$A \approx -\alpha G b \varepsilon \varrho_0^3, \tag{2}$$

where ϱ_0 is the radius of the rigid sphere, which, if introduced into the lattice at the site of the impurity, would produce no dilatation there, and ε measures the relative difference $(\varrho' - \varrho_0)/\varrho_0$ between the actual radius ϱ' of the impurity and the former quantity. Of course, equation (2) is not valid within, say, 1 atomic distance of the dislocation axis; however, we shall neglect this complication here. Neither are the effects

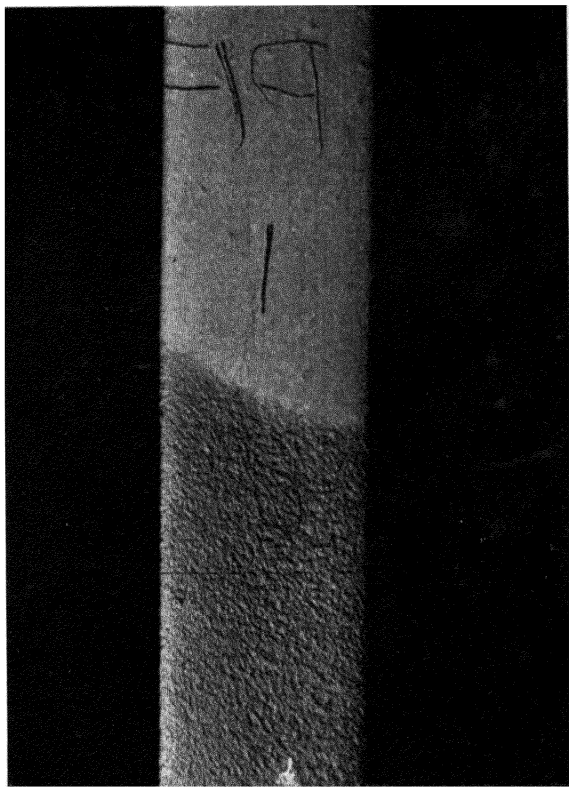


Fig 2. Lüders band in strip of polycrystalline mild steel tested in tension (after LOMER [1952]).

due to the change in the dislocation stress field caused by the presence of locked atoms themselves (saturation of the dislocation) taken into account. Since in Cottrell's original theory only the hydrostatic stress component was considered, interaction occurred only with dislocations that had at least some edge character. CRUSSARD in [1950] also considered the interaction between *shear* stresses and strains of dislocation and impurity atom in body-centred cubic metals, and found about as large an interaction energy with screw dislocations as with edge dislocations. COCHARDT, SCHOECK and WIEDERSICH [1955] studied the interaction between a straight screw dislocation along ⟨111⟩ and a carbon atom in α-iron with tetragonal deformations around it more in detail, starting from equations IV-(64):

$$U = -\sum_{ij} P_{ij}\,\varepsilon_{ij}\,. \tag{3}$$

Here U stands for the energy of interaction between the foreign atom characterized by the stress dipole tensor P_{ij} (see section IV–5) and the strain field ε_{ij} of the dislocation. In a polar coordinate system with the dislocation as polar axis, the only existing strain component is (IV–(50)):

$$\gamma_{z\varphi} = \frac{b}{4\pi r}\,,$$

where r is the distance from the dislocation. Transforming the stress dipole P_{ij} which has only components P_{11}, P_{22} and P_{33} along the cube axes, also to this coordinate system, the component $P_{z\varphi}$ can be derived and inserted into (1). One then finds

$$U = \frac{-b}{3\pi\sqrt{2}}\left[\frac{\sin\varphi}{r}\{P_{11} - \tfrac{1}{2}(P_{22} + P_{33})\} + \tfrac{1}{2}\sqrt{3}\,\frac{\cos\varphi}{r}\,(P_{22} - P_{33})\right] \tag{4}$$

for the interaction energy between screw dislocation and point defect. This formula is comparable to (1), and the interaction is thus certainly not negligible. The P_{ii} are not very well known however (see Chapter XVII). A similar calculation can be applied to face-centred cubic metals. It has been stated repeatedly that impurity atoms in face-centred cubic crystals could not interact with screw dislocations at all, because they should not cause a tetragonal deformation. This is wrong. It has been used as an argument to explain the absence of a sharp upper yield point in face-centred cubic alloys. However, such a yield

point has indeed been observed, e.g. in α-Cu-Zn (COTTRELL and ARDLEY [1953]) and in Cu-Au (HAASEN and KELLY [1957]). These are concentrated solid solutions in which other than elastic interactions may be important. Moreover, in close-packed crystals the dislocation are extended, and at least one of the partials must have some edge character. Therefore some elastic interaction is certainly present.

In thermodynamic equilibrium, the interstitial solute atoms cluster around the dislocations, their concentration as a function of the distance from the axis being given by the formula

$$c = c_0 \exp \{U(r)/kT\}, \tag{5}$$

where c_0 is the mean concentration of the solute in the bulk of the material. (Again U is considered to be independent of the concentration.) CRUSSARD [1950] has derived this formula under somewhat more general conditions than Cottrell and Bilby did, but with the same final result, at least when $c_0 \ll 1$. The equilibrium number of impurity atoms within a distance R of a dislocation can now be derived. If the crystal contains α atoms per unit volume, in an element of volume dv there are $\alpha c dv$ foreign atoms. If this element is a thin prism parallel to the dislocation bounded by two planes through the dislocation axis which subtend an angle $d\theta$, and by two cylindrical surfaces at distances r and $r + dr$ respectively, it contains $\alpha c r dr d\theta$ foreign atoms per unit length. Inserting for c the expression (5), we find for the total number of foreign atoms per unit length of dislocation within a distance R:

$$F(R) = \int_{r_0}^{R} \int_{0}^{2\pi} \alpha c_0 \exp \left\{\frac{A \sin \theta}{rkT}\right\} r \, dr d\theta, \tag{6}$$

where r_0 is the smallest distance from the dislocation for which (5) is valid. We shall put this equal to the Burgers vector b. To form an estimate, we can neglect the θ-dependence, and obtain (CRUSSARD [1950]):

$$F \approx \pi \alpha c_0 \frac{Ab}{kT} = F_0 \frac{A}{kT} \frac{b}{R^2}, \tag{7}$$

where F_0 stands for the number of impurity atoms in the same cylindrical volume if no dislocation were at the centre of it. If the temperature is low, the "atmosphere" of foreign atoms is *saturated:* along the axis of the dislocation the number of foreign atoms approaches the limit

set by the maximum available space. Expressed as a concentration, this limit is of order unity or somewhat smaller. We thus obtain an estimate for the temperature T_1 below which this occurs by putting in (5) $c = 1$ and everywhere $U(r) = U^* = A/b$. U^* is often called the *binding energy* of the impurity to the dislocation, it is the maximum value U can attain. Then

$$T_1 = U^*/k \log (1/c_0) . \tag{8}$$

As an example, consider again α-iron containing carbon in solid solution. From the known sizes of the carbon atoms and the available interstitial-cavities respectively, COTTRELL [1950, 1953]) computed $A = 3 \times 10^{-2}$ dyne/cm^2 (edge dislocation) from which follows $U^* \approx 1$ eV and $T_1 \approx 700$ °K. CRUSSARD [1950] has revised the estimate (using a better elastic approximation) and finds values that are about $\frac{1}{3}$ as large. Substitutional foreign atoms interact with dislocations much less intensively because the lattice dilatations are usually much smaller. For silicon atoms incorporated in aluminium, Crussard estimated U^* to be 0.1 eV and T_1 about 500 °K.

LOUAT [1956] has reconsidered the question of the equilibrium of a "Cottrell atmosphere" around a dislocation. Even in the case of saturation, there should only be one carbon atom per atom plane

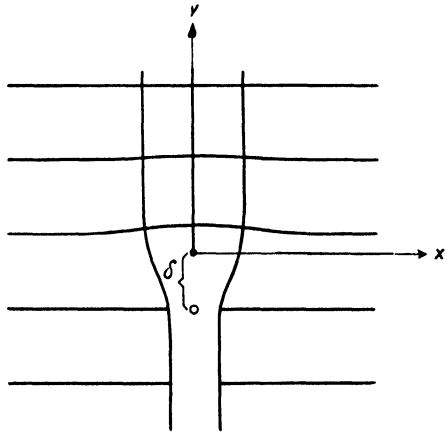


Fig. 4. Position of least energy (open circle) for an impurity atom near an edge dislocation.

within a radial distance of a few Å of an edge dislocation, as follows directly from a consideration of the available volume. This carbon

atom will presumably migrate to the position of greatest dilatation, that is immediately below the dislocation axis (at $\theta = 3\pi/2$), in the atom layer which constitutes the lower side of the slip plane (Fig. 4). A single row of carbon atoms will thus form parallel and close to the dislocation line. It is certainly not permissible to describe *this* carbon distribution by the Maxwell-Boltzmann expression, and Louat uses instead the "Fermi type" formula:

$$c = \frac{c_0 \exp(U^*/kT)}{1 + c_0 \exp(U^*/kT)}, \tag{9}$$

which describes more satisfactorily the occupation probability of the sites of lowest energy (bound positions). A similar formula also governs the occupation by additional electrons of free valence bonds along a dislocation line in a valence semiconductor. (READ [1954], see Chapter XXVIII.)

Various lines of evidence exist that point to the occurrence of a larger number of carbon atoms (10–50) per atomic layer along dislocations in α-iron at low temperatures. As this is in fact a precipitation phenomenon, we shall return to it in Chapter XXI, and consider here only the dislocation blocked by one single row of atoms.

1.3. CHEMICAL INTERACTION

H. SUZUKI [1952] has pointed out that in face-centred cubic metals where dislocations are usually split up into partial dislocations with a stacking fault in between, a completely different, "*chemical*" form of interaction between dislocations and impurity atoms can exist. The few atomic layers which constitute the stacking fault show hexagonal stacking instead of cubic. Therefore, the solid solubility of impurities contained in the matrix can very well differ appreciably "within" the stacking fault and outside. The difference is of a thermochemical nature and may amount to an effective energy of binding of the impurity atom to the extended dislocation of a few tenths of an electron-volt or more, depending on the electronic structure of the "compound".

Long-range interaction is of course absent in this form of binding. The temperature dependence of the equilibrium solute concentration in the dislocation will be very different from that expected for elastic binding. Once an extended dislocation has acquired its equilibrium amount of solute, it may be difficult to move it. The quantitative

theory of this effect lies outside the scope of this book (see also SUZUKI [1957]).

1.4. SNOEK ORDERING

Very recently another type of interaction between impurity atoms and dislocations was proposed (SCHOECK and SEEGER [1958], SUMINO [1956]). This is based on the preference for special lattice positions that foreign atoms may have when present in the stress field of a dislocation.

The material to which this idea applies in the first place is ferritic steel or more simply α-iron containing carbon. As was first pointed out by SNOEK [1941] and will be discussed in full detail in Chapter XVII, the interstitial carbon atoms in the body-centred cubic iron lattice can occupy three types of positions, which differ according to the direction of tetragonality caused by the interstitial foreign atom. Normally the three positions are equivalent, but in a field of stress, as for instance caused by a dislocation, those positions for which the axis of tetragonality complies best with the local stress tensor, will represent least free energy. In this way a state of *order* in the distribution of carbon atoms around the dislocation is created. To move the dislocation, this order must be destroyed, which means a higher yield stress, approximately independent of temperature (see next chapter)

2. Magnitude of Interaction Energy between Dislocations and Point Defects

It is of some importance to form a general idea of the order of magnitude of the energy of interaction (i.e. *binding energy*) between the various kinds of point defects and the two principal types of dislocation. As this interaction also plays an important role in materials other than metals, we shall include these in the discussion also (in this section only).

2.1. ELASTIC INTERACTION

The formula of Cottrell, as extended by Crussard, can, for purposes of numerical computation, be written as

$$U_{\text{elast}} \approx \frac{Gb\varrho_0^3}{r} \{\alpha\varepsilon + \beta(\gamma - 1)\}. \tag{10}$$

where α and β are orientation factors of the type of simple sines and

cosines, thus usually of order unity or less, ε is again the relative difference between the "atomic radii" and γ measures the ratio between the two principal strains produced by the foreign atom perpendicularly to the dislocation axis. The term with $(\gamma-1)$ thus obviously arises from shear strain interaction only; it was not included by Cottrell. For carbon in α-iron $\varepsilon = 0.06$ and $\gamma = 0.12$, and $U_{\text{elast}} \approx 0.3\ b/r$ eV. This value should also be representative for other cubic *interstitial* metallic solutions, since these are, as is well known, characterized by values of ε and γ that differ but slightly from metal to metal, and the elastic moduli do not differ widely either.

For these same reasons, the already mentioned substitutional system Al-Si can be regarded as the prototype of *substitutional* metal solutions. Here, then, U_{elast} is found to be only about $0.05\ b/r$ eV.

Vacancies do not produce any appreciable dilatations in cubic metals, but they form "soft spots" in the lattice. The interaction energy is determined by the total elastic energy due locally to the dislocation: $\sigma^2 b^3/G$, which decreases proportionally to r^{-2}. Its maximum value is a few thousandths of an eV. Interstitial atoms, on the other hand, always produce appreciable dilatations, (section III–4), and in non-close packed lattices they usually also deform the lattice in shear (see e.g. section XVII–2.4.1 for carbon in α-iron). Therefore, the elastic interaction with dislocations is quite strong. Taking $\varepsilon \approx 0.2$ (interstitial in copper, section III–4), one finds $U_{\text{elast}} \approx 0.4\ b/r$ eV.

Not much is known about the actual distortions produced by foreign atoms, vacancies and interstitial atoms in monatomic lattices other than of the body-centred cubic or face-centred cubic type, thus, for instance, in the diamond lattice (see Chapters I and XXIX). In such a relatively "empty" lattice, one would expect the defects to be accommodated rather easily, the volume changes being small. However, the shear strain interaction with dislocations might well be strong in these lattices. Moreover, in the homopolar materials that crystallize in the diamond structure and related structures, the elastic moduli are much larger than in metals. For these reasons we shall tentatively assume about the same mean values for the various elastic interaction energies to apply in the semiconductors germanium and silicon and in metals. We shall consider, however, an interstitial impurity in the former to be more like a substitutional impurity in the latter.

It is well known that in polar crystals the lattice distortions around imperfections can be much larger than in metals, since the "pliable"

Coulomb forces can accomodate much larger internal strains. Thus, in NaCl, a chlorine vacancy has $\varepsilon = 0.07$ (section XXVI–2). However, G is much smaller in these crystals (for the same reason), hence the elastic interaction energy does not surpass 0.1 eV in magnitude, neither for impurities nor for physical defects.

The splitting up of perfect dislocations into extended ones, as occurs for example in close-packed cubic metals, does not appreciably influence the elastic interaction with point defects. Although the effective Burgers vectors are decreased by the splitting and the interaction thus becomes weaker, the mean distance of an impurity from a dislocation is also decreased. In special orientations only, the average total interaction energy might be lowered, for instance when the edge character of the original dislocation is severely reduced in both the partial dislocations.

2.2. ELECTRICAL INTERACTION

Whereas the elastic interaction of dislocations and point defects is, apart from anisotropic corrections, fairly well described in all materials by formula (10), no such simple expression takes account of the electrostatic Coulomb interaction as this is widely different in character and magnitude in various types of solids. For metals, COTTRELL, HUNTER and NABARRO [1953] derived an approximate expression for the electric interaction between an edge dislocation and an impurity atom in a metal, considering only the effect of the variations in charge density that result from the volume dilatation. This charge shift is in the first instance also responsible for the electrical resistance of dislocations (Chapter V). The density variations around an edge dislocation cannot be completely followed by the volume concentration of conduction electrons, since the Fermi energy must everywhere remain the same. This condition is so severe that local departures of charge neutrality are set up to obey it: in the denser parts of the lattice there is an electron deficiency resulting in a net positive charge; in the dilated parts an excess of electrons causes a net negative charge. Thus the charge distribution along an edge dislocation resembles very much that of a line dipole. It interacts with a point defect which can be considered to carry an effective point charge z_{eff}. Owing to *screening* by the conduction electrons, the field due to this charge decays rapidly with the distance. A defect in a metal that carries z electronic charges

e differs $(z-z_0)e$ in charge from the positive ions of charge $z_0 e$ ($z-z_0 = +1$ for a vacancy, and $= -1$ for an interstitial ion in a monovalent metal), but *effective* in the electric interaction is only a small fraction of this. Cottrell *et al.* derived for the monovalent noble metals the following numerical expression for the electrical interaction energy of a vacancy with an edge dislocation:

$$U_{\text{electr}} \approx 0.02(z - z_0)b/r \text{ eV} . \tag{11}$$

A similar expression holds for the interstitial ions. It should be remarked here that the description of the interaction in terms of an effective charge is unsatisfactory from a theoretical point of view.
The energy (11) is small compared to the elastic interaction. Although the computation is certainly very crude, it is improbable that in simple metals the electric interaction energy is actually much higher than that given by (11). The screening is probably less perfect than that adopted to derive (11); however, in that case the charge shifts around the dislocation are also smaller. SEEGER [1956], who also considered the problem, obtained about the same numerical values. About the possibility of chemical binding of impurities to dislocations in close-packed crystals, we refer to section 1.3. There, the relevant binding energy will usually be of the order of 0.1 to 0.5 eV.

In polar crystals the electric interaction is much larger than in metals. As there is no screening, the lattice becomes polarized around the defect (finite dielectric constant). Moreover, a general dislocation is often characterized by a row of ion pairs along its axis (Chapter XXIV), that results in a strong binding energy but weak interaction at large distances. Jogs in such dislocations may carry an effective charge of half the ionic charge (see Chapter XXIV), and *these* attract or repel impurities and point defects strongly at large distances. Numerical estimates are hard to form. However, it is certain that in polar crystals the short-distance electric interaction between impurities and dislocations can easily become comparable in order of magnitude to $e^2/\varkappa a$, the electronic charge squared over the ionic spacing, divided by the dielectric capacity $\varkappa$ (between 1 and 10 usually). This quantity amounts to a few tenths of an eV. Here, in contrast to the situation in metals, the electrical thus surpasses the elastic interaction.

The situation in the homopolar semiconductors is even more complex, and depends inter alia on the concentration of charge carriers,

thus on impurity content and temperature. Dislocations not of pure screw character have in general a row of unpaired valence electrons along their axis (Chapter XXIX). In n-type germanium or silicon, these form a negative charge, neutralized by a positive space charge region, where the material contains fewer conduction electrons than elsewhere. Charged defects, having once arrived in this region, will be strongly attracted or repelled by the dislocation. The electric interaction therefore produces an atmosphere of point defects or impurities around each dislocation which may extend to no less than 10^{-5} cm (READ [1954]). Outside this region the interaction is virtually zero. The effective attraction energy between a positively charged impurity ion and a dislocation is approximately given by

$$U_{\text{electr}} \approx \frac{-e^2}{\varkappa r} f, \tag{12}$$

where f denotes the fraction of valence bonds along the dislocation line that is unsaturated, and $\varkappa$ is again the dielectric constant (16 in germanium, 11 in silicon). Since f is always very small (1% at room temperature) except at very low temperatures, this attraction energy is usually also rather small, about 0.01 eV. However, the formation of an impurity atmosphere is not only a question of energy, but it helps to re-establish charge equilibrium, and will thus in general take place all the same. Once having arrived at the dislocation line proper, most impurity atoms can presumably form easily a sort of *chemical* exchange bond with the unpaired electrons (Fig. 5 shows some examples). The *binding* energy is then very large, of the order of a few eV, much larger than corresponds with purely electrostatical interaction.

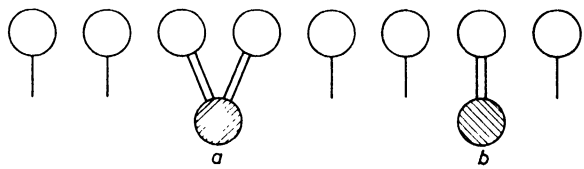


Fig. 5. "Chemical" binding of *a*) divalent (oxygen) and *b*) monovalent (lithium) impurity to a dislocation in a homopolar material, with unpaired electrons (dashes) along its axis.

Table 1 contains the considerations of this section in a concise form; it will be clear that all the numerical values contained in it must

be regarded as very rough approximations: only in the light of this statement can it be understood that we have ventured to assign a set of single data to a whole class of solids. Presented are the *binding* energies (as denoted by an asterisk), that is, the interaction energies at r = one atomic distance from the dislocation line.

TABLE 1

Estimated binding energies between dislocations and physical and chemical point defects in various solids

	Lattice dilatation	U^*_{elast} (eV)	U^*_{elect} (eV)	U^*_{total} (eV)
Interstitial atoms in metals (foreign or parent material)	6–20%	0.2–0.5	0.02	0.2–0.5 + chem. part in f.c.c. (0.1–0.5 eV)
Substitutional atoms in metals	1–4%	0.05–0.1	0.02	0.05–0.10 + chem. part in f.c.c. (0.1–0.5 eV)
Vacancies in metals	1%	0.02	0.02	0.04
Interstitial atoms in semiconductors	5% ?	$\approx$ 0.1	0.01	0.1 + chemical part ? (2–3 eV)
Substitutional atoms in semi-conductors	1–4%	0.05–0.1	0.01	0.05–0.1 + chemical part ? (2–3 eV)
Vacancies in semiconductors	$\geqslant$ 1% ?	0.02 ?	0.01 ?	0.03 ?
Interstitial and substitutional ions in polar crystals	10%	0.05–0.1	0.1–1	0.1–1
Vacancies in polar crystals	5–10%	0.03–0.06	0.1–1	0.1–1

3. Locked Dislocation under Stress. Microcreep

The force required to pull a dislocation from its saturated atmosphere (row of impurity atoms, see section 1) was estimated by COTTRELL and BILBY [1949]) as follows. Taking the x-axis along the direction of the

Burgers vector, the interaction energy per unit length of an edge dislocation which is displaced over an amount x in this direction, with its row of impurity atoms remaining in the original position $x = 0$, $y = -\delta$ (Fig. 4), is, according to (1), given by

$$U(x) = \frac{-A\delta}{(x^2 + \delta^2)b}, \tag{13}$$

The corresponding interaction force is $-\delta U/\delta x$, and this has a maximum at $x = \delta/\sqrt{3}$ amounting to

$$F_l = \frac{3\sqrt{3}A}{8\delta^2 b}, \tag{14}$$

per unit length.

The critical shear stress for break-away is $\sigma_l = F_l/b$ (formula III-(8)), and thus

$$\sigma_l = \frac{3\sqrt{3}A}{8\delta^2 b^2}. \tag{15}$$

For carbon in α-iron,

$$\delta = 2 \times 10^{-8}\,\text{cm},\ b = 2.5 \times 10^{-8}\,\text{cm},\ A = 1.8 \times 10^{-20}\,\text{dyn} \cdot \text{cm}^2$$

(SCHOECK and SEEGER [1958]) and then

$$\sigma_l \approx 10^{10}\,\text{dyne/cm}^2 = 100\,\text{kg/mm}^2.$$

This can be compared to the observed values of the yield strength of α-iron as determined by a large number of investigators. It rises from a value of the order of 10 kg/mm^2 at room temperature to about 100 kg/mm^2 at the temperature of liquid air. In contrast herewith, there is no noticeable temperature dependence any longer *above* room temperature. The rapid rise at low temperature is explained in the Cottrell model as due to the variation of the properties of the impurity atmosphere (see below). At high temperatures, the Snoek mechanism of dislocation locking, which is indeed temperature independent, presumably takes over.

The theoretical value (15) has of course been derived under the assumption of a saturated atmosphere and no thermal activation. However, other observations have recently been obtained that indicate a *constant* yield stress below 90 °K (ERICKSON and LOW [1957]). We shall not discuss this question further, but remark that Louat's

theory mentioned before indeed predicts such an independence when all bound states have been filled. HART [1954] has pointed out that small liberated loops, pinned at their ends by impurities, could act as Frank-Read sources, even if the yield stress σ_l is smaller than the critical stress σ_0 for the operation of these small sources, because the released dislocation has already an appreciable kinetic energy when it moves with a velocity approaching that of sound. The sources, once originated, will suddenly emit a large number of loops; Hart finds for the number of loops emitted:

$$n = \exp\left[2\pi^2(\sigma_l/\sigma_0) - \pi^2\sqrt{2}\right]. \tag{16}$$

The catastrophic character of the plastic flow immediately after yielding is due to this and to the blocking influence of grain boundaries: immediately after yielding the stress drops to the lower yield level and the strain increases by a large amount. The surface of a single crystal after yielding usually shows very deep slip lines, which is also in accord with this picture.

The effect of temperature on the yield stress must be explained both by the influence of thermal fluctuations and by the dilution of the atmosphere. Let us first consider the second influence (following COTTRELL [1953]). A moving dislocation can drag its atmosphere along with it if its velocity v is sufficiently small, so that the foreign atoms can, by diffusion, follow the dislocation. The interaction force with an atom at distance r is approximately $\mathrm{d}U/\mathrm{d}r \approx A/r^2$. Due to this force, the atom acquires a drift velocity

$$v \approx \frac{D}{kT} \cdot \frac{A}{r^2}, \tag{17}$$

where D is the diffusion coefficient of the impurity. Putting $A/kT = l$, the "characteristic radius" of the atmosphere, it is a reasonable assumption to consider the atmosphere as being effectively dragged along with the dislocation when those foreign atoms contained within $r = l/2$ can follow it. Then one obtains from (17) by inserting $r = l/2 = A/2kT$ the *critical speed* of the dislocation, above which it breaks away also from these atoms:

$$v_{\mathrm{cr}} = 4D/l. \tag{18}$$

To estimate the stress necessary to move the dislocation with the

velocity v_{cr}, we notice that the mean force per unit length with which the foreign atoms pull the dislocation back is $A/(l/2)$ times the number of solute atoms per unit area within the atmosphere, which is approximately $2lc_0n \exp\{A/kT \cdot l/2\}$; compare (5). Here n is the number of solvent atoms per cm^3 and c_0 is the mean concentration of impurities.

Thus one finds for the opposing stress:

$$\sigma_{cr} = \frac{4kTl}{b} c_0 n e^2 \approx \frac{40A}{b} c_0 n \,. \tag{19}$$

This approximate expression holds for the break-away stress from a dilute atmosphere, and then, thus at high temperatures, replaces (15). CHALMERS [1936] observed a type of slow creep in tin crystals occurring well below the tensile stress of 150 g/mm^2, which constitutes the "yield stress" for ordinary plastic flow. The creep rate increased linearly with the stress, but it decreased rapidly in time (at constant stress). COTTRELL and JASWON [1949] explained this so-called *microcreep* in terms of the just discussed slow motion of dislocations which must drag their non-condensed atmosphere of impurities along with them. At small stresses they cannot break loose from this cloud and move with a velocity smaller than the critical velocity v_{cr}. In first approximation the actual velocity will be given by

$$v = \frac{\sigma}{\sigma_{cr}} v_{cr} \,, \tag{20}$$

Thus, according to (18) and (19), and noting that the strain rate is given by

$$\dot{\varepsilon} = Nbv \,,$$

where N is the density of dislocations, we have

$$\dot{\varepsilon} = \frac{1}{l^2} \frac{Nb^2}{c_0 n e^2} \frac{D}{kT} \sigma \,. \tag{21}$$

This creep rate is indeed proportional to the stress, and depends exponentially on the temperature (through D). The quantity $l = A/kT$ is of course a measure for the binding energy of the impurities; by comparing (21) with the creep rate observed by Chalmers, a reasonable value for this energy was indeed obtained (a few tenths of an eV). The effect of thermal fluctuations on the actual abrupt yielding process will be discussed in the next section.

4. Thermal Liberation of Dislocations. Delayed Yield

4.1. EXPERIMENTAL

Careful observations have shown that abrupt yielding of a metal sample occurs only after the load has been applied for a certain time. CLARK and WOOD [1949] showed this *delay time* (in mild steel) to depend very sensitively on the stress; it decreased by a factor of 1200 while changing the stress from about the normal yield stress value (as determined in the usual sort of tensile machines) where the delay amounted to a few seconds, to a stress only about 40% higher. Already during the incubation time, that is, *before the delayed yield* (as this phenomenon has been called) there is some plastic creep which subsequently peters out. This has been termed *pre-yield creep* and was studied by HENDRICKSON, WOOD and CLARK [1956] in a well-defined Fe-C alloy; they could describe their observations by the following empirical formula for the creep rate as a function of stress and temperature:

$$\dot{\varepsilon} = \exp\left(A - \frac{B}{\sigma T}\right). \tag{22}$$

The length of the delay time and the strain rate thus depend both strongly (exponentially) on temperature and stress; the phenomenon can only be observed at fairly low temperatures (at or below room temperature in iron). In this respect and also in the stress dependence pre-yield creep differs markedly from the microcreep discussed above,

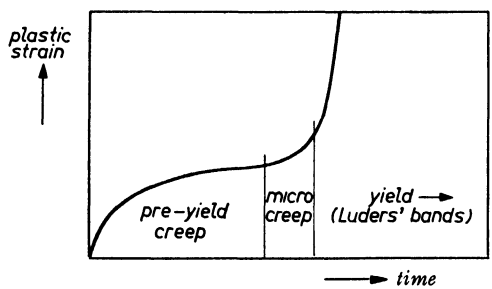


Fig 6. Pre-yield creep and microcreep (schematic). At low stresses and temperatures only the former occurs and a saturation strain is reached.

which varies only slowly with stress and occurs at elevated temperatures.

The difference between microcreep and pre-yield creep is not always

made clear in literature. Fig. 6 illustrates this distinction (see also OWEN *et al.* [1958]). Whereas the rate of microcreep is governed by the *diffusion* of impurities, varies but slowly in time and is *proportional* to stress, pre-yield creep (which is often also called microcreep) is determined by thermally activated breaking loose of impurities from a locked dislocation, and reaches a *saturation* value in time when a new equilibrium is attained; the dependence on the stress is much *stronger* than a simple proportionality. Some microcreep also can occur before actual yielding takes place, but not at stresses below that given by (19).

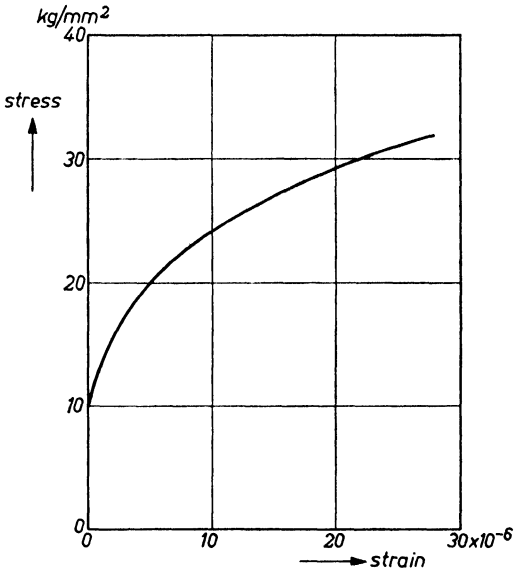


Fig. 7. Applied tensile stress versus saturation value of pre-yield creep strain in a low-carbon steel (after VREELAND *et al.* [1953]).

VREELAND, WOOD and CLARK [1953] have studied the saturation value of the pre-yield creep of a low-carbon steel as a function of the applied stress. They measured the strain at room temperature after application of a stress (smaller than the yield stress) for about 5 minutes, after which the strain had become essentially constant. Stress and "saturation strain" were connected by a curve of the form given in Fig. 7. A lower limiting value of the stress is apparent, at which no creep occurs at all, and an upper limit also, above which yielding occurs. The authors explained the observed relation by con-

sidering more specifically the creep as being due to the activation of Frank-Read sources of suitable length, which goes on at a decreasing rate and finally comes to an end when the applied stress has become too small to counteract the ever increasing back stress exerted by the emitted loops. No breaking away on a large scale can occur. It depends on the distribution of the source lengths what form the stress-strain relationship will take. The authors mentioned a distribution function with a maximum length of only about 1200 atomic distances and a minimum length near 500 atomic distances. These seem rather short lengths, even when the blocking action of the carbon atoms is considered. An explanation along the lines first proposed by Cottrell and discussed below seems therefore more feasible.

Delay times for yield in single crystals of several body-centred cubic, hexagonal close-packed and face-centred cubic metal single crystals have been studied by TIEN-SHIH-LIN *et al.* [1956]. They failed to observe any delay time in copper, but found delay times of the order of 10^{-3} sec in body-centred cubic β-brass and hexagonal zinc at low temperatures. Estimated activation energies (from the temperature dependence of the delay times) were about 0.01 eV in zinc and 0.05 to 0.1 eV in β-brass. Recently also in some close-packed crystals a delay time has been observed, viz. in the same alloys where also a yield point is present. Evidently, dislocations can be locked by impurity atoms also in these crystals. Since the effects are especially strong in *alloys*, it seems probable that chemical interaction is here most important (SHEPHARD and DORN [1956] (Al-Mg alloys) and others). But even in single crystals of copper a yield point has been detected under carefully chosen experimental circumstances (for literature see HAASEN and KELLY [1957]).

4.2. THEORY OF THE RELEASE PROCESS

The occurrence of a delay time can be interpreted in terms of a thermally activated process that is needed for the yield to be initiated. FISHER [1955] assumed that one sees here the effect of the *thermally activated breakaway* of dislocation segments from their pinning points (impurities). This should then be regarded as a *consecutive* process, as a certain number of pinning points have to released *successively*. Consider a dislocation segment of length l, containing a certain number of pinned foreign atoms (Fig. 8). If the line tension of a free dislocation

is $\gamma_0 \approx Gb^2$ (see section III-2), that of a dislocation which has absorbed foreign atoms is certainly lower; it will be denoted by $\gamma(< \gamma_0)$. A shear stress σ acts in the direction of the Burgers vector. The dislocation will

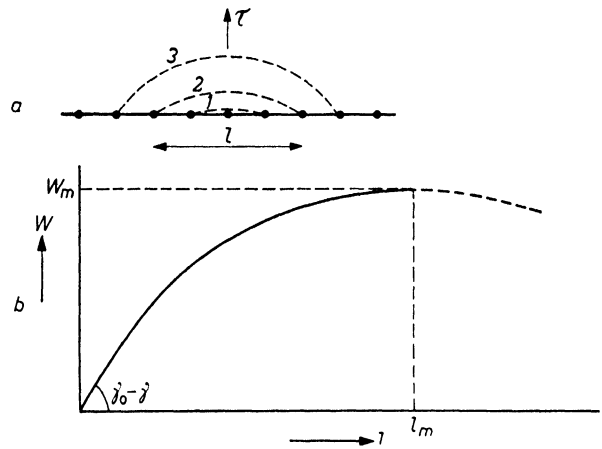


Fig. 8. On the derivation of Fisher's formula.
a) A dislocation breaks away successively from its locking impurities.
b) The energy W to be supplied for this reaches a maximum as a function of the length l of the released portion.

try to bow out and assume the form of a circular arc (section III–6) of radius $R = Gb/\sigma$. The energy of the dislocation in the straight form is $l\gamma$, whereas the energy of the system in the bowed-out situation can be written as

$$l'\gamma_0 - b\sigma A \ ; \tag{23}$$

where l' is the length of the bowed-out section and A is the area covered by it. By applying simple geometry, l' and A can easily be expressed as functions of σ and l, and then it is found that the energy *difference* between the two situations:

$$W = \gamma l - \gamma_0 l' + b\sigma A \ , \tag{24}$$

reaches a maximum (for a given stress) at a certain value of l, which we shall call l_m. The height of this maximum is

$$W_{max} = \frac{\gamma_0^2}{b\sigma} f(\gamma/\gamma_0) \ , \tag{25}$$

per atomic plane, where, if γ/γ_0 differs but little from unity,

$$f(\gamma/\gamma_0) \approx 2\left(1 - \frac{\gamma}{\gamma_0}\right)^{\frac{3}{2}}.$$

W_{max} is now regarded by Fisher as the *activation energy* (to be supplied by the thermal vibration) for the break-away of a dislocation loop from its impurities. He argued that, once the length l_0 which is usually of the order of 10^4 atomic distances has been released by successive breaking loose from the pinning points, the loop will further extend spontaneously. Owing to the assumption that a *consecutive* process underlies the phenomenon, the dependence of the rate of the process on the stress is very strong. This is expressed by the fact that the activation energy is inversely proportional to the stress. According to Fisher's theory the delay time as well as the inverse strain rate during pre-yield creep, are thus both expected to depend on the stress and the temperature through one *exponential factor*, of the form

$$e^{C/\sigma T} \tag{26}$$

where

$$C = \frac{{\gamma_0}^2}{bk} f(\gamma/\gamma_0) . \tag{27}$$

Such a strong stress dependence has not often been observed in practice, although it seemed to be present in the original results obtained by Hendrickson *et al.* (formula (22)), see also below).

From this theory also the dependence of the upper yield stress on the temperature under given experimental conditions can be deduced. This follows directly from the result contained in (26), that the *product* of stress and temperature governs the yield process. Therefore, when all other circumstances remain the same, the yield stress should be inversely proportional to the temperature:

$$\sigma_y \sim \frac{1}{T}. \tag{28}$$

In practice, a much stronger temperature variation is usually observed (see below).

From the observed stress and temperature dependence of the delay time, an estimate for $f(\gamma/\gamma_0)$ and thus for

$$\frac{\gamma_0 - \gamma}{\gamma_0}$$

can be derived. This quantity is directly proportional to the binding energy of the impurity atom to the dislocation, the proportionality constant being the inverse line tension itself. From all experiments carried out on α-iron so far, the ratio $(\gamma_0-\gamma)/\gamma_0$ was derived as only 0.002 to 0.003, from which it follows that the binding energy for a carbon atom to a dislocation should only amount to 0.01 to 0.02 eV (unless one assumes carbon atoms to be adsorbed only here and there along a dislocation instead of in every atomic plane). This seems an extremely small value, as compared to the data in Table 1 and with other experimental evidence on yielding and locking of dislocations by foreign atoms.

COTTRELL [1957] has studied the rate of propagation of so-called Lüders bands or stretcher strains (Fig. 2), the specific surface markings occurring during the constant part of the stress-strain curve of Fig. 1, immediately after yielding. They propagate through the specimen with a finite velocity during the lower yield-point extension. The form and behaviour of the Lüders bands strongly suggest the simultaneous motion of large numbers of dislocations freed after yielding (see also LOMER [1952]). Recent experiments by VREELAND *et al.* [1957] showed the delay time for yielding and the inverse rate of propagation of Lüders bands to be exactly the same functions of the temperature and the stress. However, this universal function, although containing an activation energy and showing an exponential stress variation with the stress, differed from that proposed by Fisher in that the activation energy itself was certainly not inversely proportional to the stress, but rather independent of it; the empirical formulae read:

$$\text{delay time: } t = t_0 e^{U/kT}\, e^{\sigma_1/\sigma}$$

$$\text{propagation rate: } v_L = v_0 e^{-U/kT}\, e^{-\sigma_1/\sigma}. \qquad (29)$$

Similar formulae have been derived empirically by SHEPHARD and DORN [1956] in their study of substitutional Al-Mg alloys. An analysis of the mechanism of propagation of a Lüders band has been given by HART [1955].

A theoretical explanation of these formulae has not yet been given, and other semi-empirical formulae for the stress and temperature dependence of the delay time for yield have been proposed (CAMPBELL [1953], COTTRELL [1957]). Whereas the latter dependence is always exponential, through an energy of activation, the variation of this

energy with stress is invariably much slower than would follow from Fisher's formula. In germanium and silicon similar pre-yield phenomena can be studied with unusual accuracy (see Chapter XXX). There, a linear dependence of the activation energy on the stress, of the form

$$Q = Q_0 - c\sigma\,, \tag{30}$$

was derived from the experiments (VAN BUEREN [1958]).

COTTRELL [1957] has, in the light of most of the evidence obtained recently, reconsidered the assumptions underlying Fisher's theory, and especially the question whether indeed a consecutive release process should be needed (because this infers the strong stress dependence). Arguing that the concept of line tension loses its meaning for a very small dislocation loop, the strain field of which—the physical origin of the line tension—extends much further than the radius of curvature of the loop, Cottrell criticizes Fisher's approach, and considers anew the elementary process of the release of a small piece of

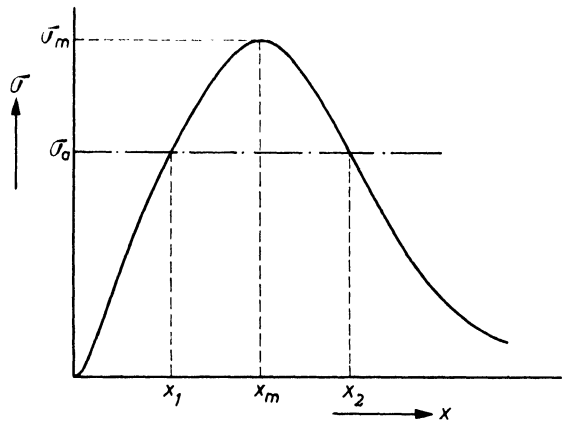


Fig. 9. The distance x between a segment of dislocation of unit length and an impurity atom determines the stress σ with which the former is pulled back to the latter. At a certain distance x_m this stress is maximum. In the simultaneous presence of an applied shear stress σ_a there are two equilibrium positions of the dislocation: x_1 (stable) and x_2 (labile).

dislocation containing only one impurity atom. Using formula (13), the difference in *core* energy rather than in *line* energy of the dislocation segment, between the locked position and when removed a certain distance from the impurity (Fig. 9), is estimated. The relation

between the stress σ which tries to pull back the dislocation to the impurity atom and the distance x, can be found by differentiating this energy difference. The stress being given, there are two positions of equilibrium of the dislocation. The first is a position of *stable* equilibrium, approximately given by

$$x_1 = \frac{1}{2}\frac{\sigma}{\sigma_m} x_m \; ; \tag{31}$$

here x_m denotes the shift in position to where the interaction shear stress has its maximum sustainable value σ_m. There exists also a position x_2 of *unstable* equilibrium for $\sigma < \sigma_m$, namely that in which the dislocation has already *overcome* the interaction barrier σ_m. This position depends much more strongly on the applied stress, approximately as

$$x_2 = 2\sqrt[3]{\frac{2\sigma_m}{\sigma}}\, x_m \, . \tag{32}$$

This formula is due to HAASEN [1959]. Cottrell, by using a piece-wise linear approximation to (13), derived a linear dependence of x_2 on σ. We shall follow Haasen's more accurate computation here.

Beyond the position x_2, the dislocation segment can, under the action of a slight increase in the applied stress, leave its impurity cloud altogether, since the applied stress always exceeds the interaction stress pulling the dislocation back. The difference in total *potential*

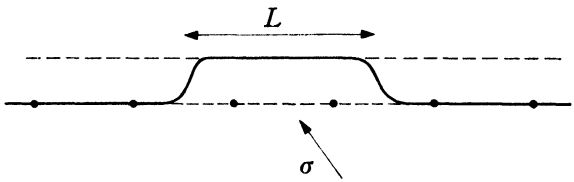


Fig. 10. Kink in a dislocation line bound to impurity atoms, formed under the combined action of an applied stress and thermal vibrations. The length L of the kink is determined by the mutual attraction between its side parts.

energy of a dislocation of unit length with Burgers vector b in the direction of the stress, between these two equilibrium positions will be called $W(\sigma)$, and the distance between the positions d. The latter distance is always of the order of x_m, thus of b. In practice, an

originally straight locked dislocation will, in the presence of stress, develop one or more *kinks* as illustrated in Fig. 10. In these kinks the dislocation has jumped from the stable to the unstable equilibrium position. The mean length of a kink is determined primarily by the counterbalance of two opposing forces: that exerted by the applied stress, which tends to shift the greatest possible length of dislocation line (i.e., tends to pull the sides of the kink apart), and the attractive force between the two sides of the kink (compare also the discussion on the low-temperature dislocation damping in section XVII–2.2.1). Equilibrating these two forces it is found that the length of a stable kink will be of the order of its depth d, thus of the order of the Burgers vector. The *activation energy* for the thermal generation of such kinks is now simply $Q = W(\sigma)d - \sigma bd^2$ — difference in line tension.

Neglecting the line tension altogether, we find after some calculation:

$$Q = 4\sigma_m b x_m^2 \left\{ \frac{4}{3}\sqrt[3]{\frac{2\sigma_m}{\sigma}} - 3\sqrt[3]{\frac{\sigma}{2\sigma_m}} \right\}. \tag{33}$$

In order of magnitude, for $\sigma \ll \sigma_m$:

$$Q = \tfrac{16}{3}\sigma_m b x^2{}_m \sqrt[3]{\frac{2\sigma_m}{\sigma}}. \tag{34}$$

$U^* \approx \sigma_m b^3$ equals the binding energy (per unit length of dislocation) of the impurity atom to the imperfection. Taking a reasonable value for U^* (0.5 eV), one finds quite generally σ_m to be of the order of $G/100$. The quantity before the bracket in (33) is again of the order of U^*, hence we can write

$$Q = \alpha U^* \left(\frac{\sigma_m}{\sigma}\right)^{\frac{1}{3}}, \tag{35}$$

where α is a numerical factor of order unity.

The formula originally derived by Cottrell reads

$$Q = 9\sigma_m b x_m^2 \left(1 - \frac{\sigma}{\sigma_m}\right)^3, \tag{36}$$

but in this the long range interaction as present in formula (13) has been neglected.

About a comparison of this theoretical $\sigma^{-\frac{1}{3}}$ dependence of activation energy on stress with the various observations, we must refer to the

literature. It may yet be remarked that this dependence varies very sensitively with the form of the interaction law used. When chemical or electrical interaction dominates, a completely different stress dependence must be expected (see for instance section XXX–5.1, where the application of the Cottrell theory to homopolar materials is discussed).

REFERENCES

BUEREN, H. G. VAN (1958); Physica **24**, 831.
CAMPBELL, J. D. (1953); Acta Metall. **1**, 706.
CHALMERS, B. (1936); Proc. Roy. Soc. **A156**, 427.
CLARK, D. S. and D. S. WOOD (1949); Proc. ASTM **49**, 717.
COCHARDT, A., G. SCHOECK and H. WIEDERSICH (1955); Acta Metall. **3**, 533.
COTTRELL, A. H. (1950); Report Conference on Plastic Deformation of Metals, Pittsburgh, p. 60 (Carn. Inst. Techn. and O.N.R., Washington).
COTTRELL, A. H. (1953); Dislocations and Plastic Flow of Crystals (Clarendon Press, Oxford).
COTTRELL, A. H. (1957); Conf. Mech. Engrs. Birmingham; On high rates of strain.
COTTRELL, A. H. and W. ARDLEY (1953); Acta Metall. **1**, see also PIERCY, G. R., R. W. CAHN and A. H. COTTRELL (1955); Acta Metall. **3**, 331.
COTTRELL, A. H. and B. A. BILBY (1949); Proc. Phys. Soc. **A62**, 49.
COTTRELL, A. H., S. C. HUNTER and F. R. N. NABARRO (1953); Phil. Mag. **44**, 1064.
COTTRELL, A. H. and T. JASWON (1949); Proc. Roy. Soc. **A199**, 104.
CRUSSARD, C. (1950); Métaux et Corrosion **25**, nr. 301, 203.
ERICKSON, J. S. and J. R. LOW (1957); Acta Metall. **5**, 405.
FISHER, J. C. (1955); Trans. ASM **47**, 451.
HAASEN, P. (1959), in: Internal Stresses and Fatigue in Metals, p. 205 (Elsevier, Amsterdam).
HAASEN, P. and P. KELLY (1957); Acta Metall. **5**, 192.
HART, E. W. (1954); Acta Metall. **2**, 416.
HART, E. W. (1955); Acta Metall. **3**, 146.
HENDRICKSON, J. A., D. S. WOOD and D. S. CLARK (1956); Acta Metall. **4**, 593.
LOMER, W. M. (1952); J. Mech. Phys. Solids **1**, 64.
LOUAT, N. (1956); Proc. Phys. Soc. **B69**, 459.
OWEN, W. S., M. COHEN and B. L. AVERBACH (1958); Transact. ASM **50**.
PAXTON, H. W. and I. J. BEAR (1955); J. Metals **7**, Transact. Suppl. II, 989.
PIERCY, G. R., R. W. CAHN and A. H. COTTRELL (1955); Acta Metall. **3**, 331.
READ, W. T. (1954); Phil. Mag. **45**, 775.
SCHOECK, G. and A. SEEGER (1958); Acta Metall. **7**, 469.
SEEGER, A. (1956); unpublished.
SHEPHARD, L. A. and J. E. DORN (1956); J. Metals **8**, Transact. Suppl. II, 1229.
SNOEK, J. L. (1941); Physica **8**, 711.
SUMINO, K. (1956); Jap. J. Physics **11**, 897.
SUZUKI, H. (1952); Rep. Tohoku Univ. Res. Inst. **A4**, 455.
SUZUKI, H. (1957), in: Dislocations and the Mechanical Properties of Crystals (Wiley, New York).
TIEN-SHIH-LIN, I. R. KRAMER and M. A. STEINBERG (1956); Acta Metall. **4**, 414.
VREELAND, T. (1957); Dislo ations and Mechan. Prop. of Crystals (Wiley, New York).
VREELAND, T., D. S. WOOD and D. S. CLARK (1953); Acta Metall. **1**, 414.

CHAPTER X

AGEING AND HARDENING OF ALLOYS

1. Strain Ageing

If a metal A contains atoms of another element B in supersaturated solid solution, the latter usually tends to precipitate in the form of a compound with the solvent metal. This formation of a second phase occurs by the successive processes of nucleation and growth, and the rate of the latter is ultimately controlled by the diffusivity of B through A. When the temperature is low this diffusivity is often very small, and the supersaturated solution is *frozen in*. Such a situation can for instance arise after quenching; a well-known example is iron with a few hundredths per cent of carbon, quenched from above 720 °C in water. Above the quenching temperature the carbon dissolves readily in the, then stable, face-centred cubic austenite. Rapid quenching results in the transformation of the austenite into body-centred cubic ferrite, in which the solubility of carbon is extremely low, less than $10^{-5}\%$ at room temperature. The excess carbon thus tends to precipitate in the form of cementite, Fe_3C, as is clear from the phase diagram (Fig. 1). However, the diffusivity of carbon is also extremely low at room temperature, and therefore most of the carbon remains in supersaturated solid solution.

If dislocations are also present, an interaction occurs with the foreign solute atoms in the way described in the previous chapter. Let us simplify the actual situation, and assume that the interaction energy is given by

$$U = \frac{A \sin \theta}{r}, \tag{1}$$

as holds for an edge dislocation interacting through its hydrostatic stresses alone. In Fig. 2 the equipotential lines of this form of inter-

action are depicted. They are circles centred on the y-axis (extra half plane) and passing through the (z–) axis of the dislocation. On a given

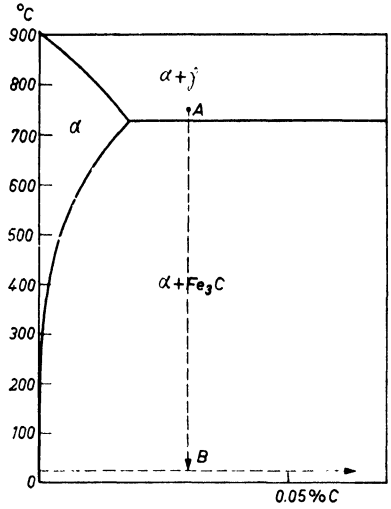


Fig. 1. Portion of the iron-carbon diagram. After quenching the austenitic alloy A to room temperature, the excess carbon tends to precipitate as cementite in the α-iron matrix.

solute atom acts a force $F = -$ grad U, which gives rise to a drift velocity (in the absence of concentration gradients):

$$v = \frac{D}{kT} F = -\frac{D}{kT} \text{ grad } U . \tag{2}$$

Of course, this drift velocity is not the actual velocity of the carbon atoms, only the mean velocity that is observed during a large number of "random" jumps. When, for the purpose of simplification, the drift motion of the solute atoms is regarded as a real velocity v, we can draw in Fig. 2 the dotted flow lines, which are circles orthogonal to the equipotential circles, and are centred on the x-axis (glide direction). In this approximation the solute atoms drift along the circles of flow until they have reached the point of maximum binding energy (see Chapter IX) immediately below the dislocation axis. The velocity, being proportional to the diffusion coefficient, is of course very small at low

temperatures: from the known value of D at 20 °C it follows that a carbon atom in α-iron jumps about once per second from one interstitial site to a neighbouring one.

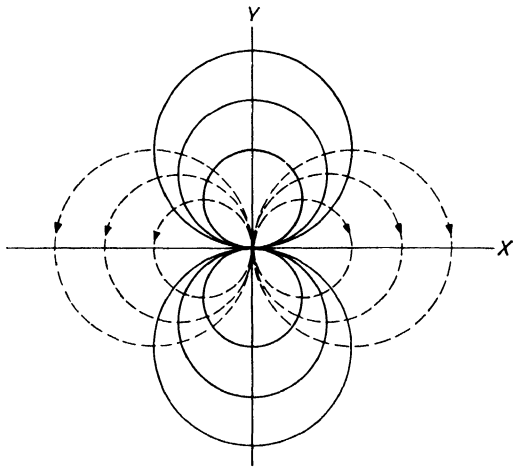


Fig. 2. Drawn: equipotential lines for the elastic interaction potential (1) between an edge dislocation along the z-axis with Burgers vector in the x-direction, and solute atoms. Dotted: flow lines perpendicular to the former, indicating the course taken by these atoms in drifting to the dislocation (in the absence of lateral diffusion). (After COTTRELL and BILBY [1949].)

On the basis of this picture, COTTRELL and BILBY [1949] have computed the rate of agglomeration of solute impurity atoms in an edge dislocation. As grad U can be approximated by A/R^2 on a flow circle of radius R, the velocity of the atoms on this circle is about constant, $(AD/kT)/R^2$, and the average time in which all flow circles with radii smaller than R are completely drained of foreign atoms, is easily found to be of the order

$$t \approx \frac{kTR^3}{AD}. \tag{3}$$

The number of atoms (per unit length of dislocation perpendicular to the plane of drawing in Fig. 2) originally present in the drained area, is proportional to R^2 and thus to $t^{2/3}$. Hence the concentration of solute atoms in the dislocation increases proportionally to $t^{2/3}$. More accurately

Cottrell and Bilby derived

$$n(t) = 3\left(\frac{\pi}{2}\right)^{1/3} n_0 \left(\frac{ADt}{kT}\right)^{2/3}, \tag{4}$$

where n_0 is the total number of solute atoms per unit volume and $n(t)$ is the number precipitated in the dislocation after the time t. Formula (4) has been verified remarkably accurately in a great variety of experiments, for instance concerning the decrease of the Snoek-damping in quenched and subsequently deformed iron (see section XVII–2.4) or the decay of the excess resistivity due to solute carbon in iron. From these experiments the value of A and thus of the binding energy of a carbon atom to an edge dislocation in iron was obtained as about 0.5 eV (see also Chapter IX). HAM [1959] has revised the derivation of (4), taking into account *lateral* diffusion. He obtained a result very similar to (4). The $t^{2/3}$-law applies in the beginning of the process only, to be replaced by an exponential time dependence later on.

The slow precipitation of carbon atoms into dislocations gives rise to the phenomenon of *strain-ageing*. A freshly quenched iron-carbon alloy does not show a sharp yield point, since all dislocations are still free of carbon (or nitrogen atoms, which behave similarly). After resting the quenched material many hours at say 50 °C, a temperature at which the carbon or nitrogen diffusion is already sufficiently rapid, a yield point has again developed because the dislocations have gradually become locked by the impurities. The material has now also become harder.

The same effect occurs in ordinary yield tests. Unloading a freshly yielded iron specimen at room temperature, and reloading it immediately afterwards, results in the absence of a yield point. This returns only after keeping the unloaded material for several hours at a sufficiently high temperature. The explanation is the same as before: directly after yielding the dislocations are free of carbon and no yield phenomenon is to be expected. Only after the carbon or nitrogen atoms have had the opportunity to segregate into the dislocation lines, or into precipitates elsewhere in the material, does the yield point return.

2. Precipitation Hardening

When an alloy contains a second phase, particles of this phase, consisting of more than a few molecules, are present throughout the

matrix of the parent metal. These *precipitates* usually produce internal strains in the metal, as, for instance, the specific volume of the second phase will usually differ from that of the matrix. These strains interact with the dislocations and *harden* the material. When the precipitates are spherical, this interaction is completely analogous to that between dislocations and point defects as far as elastic strains are concerned; only it occurs on a different scale.

Once a precipitate nucleus has reached a certain size, it will, on the average, grow by accretion of new solute atoms, because a mutual attraction exists between the precipitate and the solute atoms. CRUSSARD [1950] derived from simple elasticity considerations the following expression for the energy of interaction between a spherical precipitate of radius R' and a solute atom:

$$U_{\text{int}} = \frac{3\pi}{\varkappa} R^3 C_B^2 \varepsilon^2 \frac{R'^6}{r^6} . \tag{5}$$

Here $\varkappa$ is the compressibility of the matrix, R the atomic radius of the parent metal A, C_B the concentration of the atoms of B and ε the relative difference in "available radius" of the atoms of A and B (compare section IV–3). About the sign of the interaction energy there has been some discussion in literature (see ESHELBY [1958]).

The classic example of an alloy in which precipitation hardening occurs is aluminium containing a few (2–4) percents of copper (the prototype of duralumin alloys). Actually *this* system is rather complex. The copper finally precipitates in the form of $CuAl_2$, but beforehand a number of coherent precipitation stages occur. For this case Crussard computed the interaction energy between a copper atom and a precipitate spherule at contact distance as 90 cal/mole. This is much smaller than the interaction energy between a dislocation and a copper atom, which is of the order of 1000 cal/mol (see Chapter IX, Table 1).

Dislocations will thus act as better collectors for foreign atoms, than spherical precipitates. They will provide *nuclei* for the precipitation of the second phase. This effect of dislocations will be discussed more fully in Chapter XXI, where also the effect of the *form* of the typical pre-precipitation structures in this alloy, commonly called *Guinier-Preston zones*, will be described.

During the process of precipitation the hardness and flow stress first increase together with the size of the precipitate, but later they decrease again. The hardness usually reaches a maximum when the

average distance between precipitate particles is between 20 and 50 atomic distances. When the precipitate coarsens further, this distance increases and the mechanical hardness decreases again (*over-ageing*).

MOTT and NABARRO [1948] explained the characteristics of precipitation hardening along the following lines (considering elastic interaction only). The internal stresses around a (spherical) precipitate particle oppose the motion of dislocations. Since the activation energy for the motion of dislocations against such stresses is very large (see below), the applied stress must be about equal to the average internal stress due to the precipitate particles in order that dislocation movement on an appreciable scale can occur at moderate temperatures.

The average absolute magnitude of this internal stress can be estimated as

$$\bar{\sigma}_i \approx 2GC_B\varepsilon \tag{6}$$

in the notation of formula (5), where C_B stands for the *original* solute concentration. It is thus *independent* of the size and mutual distance of the precipitates. This estimate is obtained as follows.

The average stress is (see Chapter IV):

$$\bar{\sigma}_i \approx G\varepsilon \frac{R'^3}{(\bar{r})^3}. \tag{7}$$

As $\bar{r} = \frac{1}{2}N^{\frac{1}{3}}$, where N is the number of precipitate particles per unit volume, $\bar{\sigma} \approx 8G\varepsilon NR'^3$. The concentration of atoms B being $C_B = \frac{4}{3}\pi R'^3 N$, (6) follows directly.

To explain the nevertheless observed dependence of the mean flow stress on the precipitate spacing, one must, according to Mott and Nabarro, consider the *flexibility* of a dislocation line. This enables a dislocation to "wriggle" through between the precipitates at a smaller stress than the limit set by $\bar{\sigma}_i$. Given a certain local stress σ, a dislocation can bend to a minimum (local) radius of curvature that is approximately given by

$$r_c = \frac{1}{2}\frac{Gb}{\sigma} \tag{8}$$

(compare Chapter III); smaller loops are instable. The flexibility of a dislocation, although extant, is thus limited. Hardening will now be maximum when the mean distance between precipitates, Λ, becomes of the same order as the minimum radius of curvature r_c in the field of internal stress.

Mott and Nabarro consider three different situations (see also COTTRELL [1953]).

1) $\Lambda \approx r_c$ (Fig. 3*a*). The dislocation is bent into loops that move

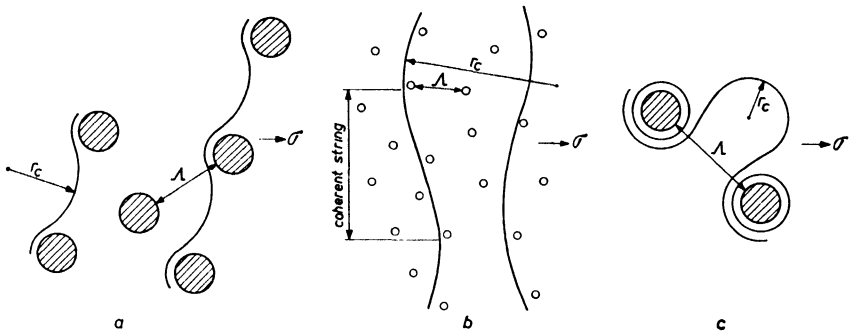


Fig. 3. Three cases of precipitation hardening.
a) The mean distance Λ between precipitates is of the same order as the minimum radius of curvature r_c to be attained by a dislocation in the field of internal stress σ (optimal hardening). *b*) $\Lambda \ll r_c$ (solid-solution hardening). *c*) $\Lambda \gg r_c$ (over-aged alloy).

largely independently of each other. The flow stress is indeed the average internal stress (6): $\sigma_f = 2GC_B\varepsilon$. The critical dispersion size follows from inserting σ_f into the formula (8) for r_c. One has

$$\Lambda_c \approx \frac{b}{4\varepsilon C_B}. \tag{9}$$

In practice, the hardest age-hardened alloys have $\sigma_f \approx G/100$, from which, as an average, taking $\varepsilon = 0.2$ and $C_B = 2.5\%$, Λ_c is found to be about 50 atomic spacings, in good agreement with the observed critical spacings.

2) $\Lambda \ll r_c$ (Fig. 3*b*). This situation refers to a very finely dispersed precipitate or, in the extreme, to the presence of single solute atoms. MOTT [1954] has refined the original theory of Mott and Nabarro for this latter case. The average wavelength of the stress field is now equal to $b/c_B^{\frac{1}{3}}$. Because this is usually *very* small compared to the radius of curvature of the dislocations, an appropriate average of the internal stress acting on a dislocation segment of length Λ is required, which can for all purposes be identified with the flow strength of the hardened material. It was found to be

$$\sigma_f = \tfrac{1}{2}G\varepsilon^{4/3}\, C_B \cdot f(C_B)\,, \tag{10}$$

where $f(C_B) = \{2C_B^{2/3} (\log C_B)^4\}^{1/3}$ is a very slowly varying dimensionless function of C_B, of order of magnitude 10.

The hardening due to finely dispersed precipitates is thus again proportional to the concentration of solute atoms, the proportionality factor, however, being $\sim 5\, G\varepsilon^{4/3}$ instead of $2G\varepsilon$ in the former case (6). Although smaller than the latter value the former is still much too high to explain to observed hardening in perfect solid solutions to which (10) should apply best. The observed values of the flow stress in several copper-base solid solutions are still 10 times smaller. In view of the observed temperature dependence of the flow strength of solid solutions it is possible that thermal activation can facilitate appreciably the slip process, as it can help the short segments Λ to overcome their obstacles (in the first example this was impossible). The activation energy for motion in the absence of stress over a distance Λ was estimated by COTTRELL [1953] to be $Gb^3\varepsilon^{2/3}$ which is indeed rather small (less than 0.1 eV) when ε amounts to a few percents.

3) $\Lambda \gg r_c$ (Fig. 3c).

This case refers to over-aged alloys. The mean distance between the precipitate particles is now so large that a dislocation passes them by just bulging through them, leaving around the particles small dislocation rings, which are formed by the coalescence of loops which have passed through (Fig. 4). The stress needed for this is simply the stress to bend the dislocation to a radius $\Lambda/2$; thus, according to (8):

$$\sigma = Gb/\Lambda\,. \tag{11}$$

This should be the flow stress of an over-aged alloy (OROWAN [1947]).

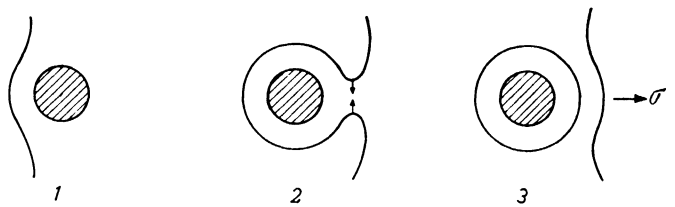


Fig. 4. Successive stages by which a dislocation can surmount an obstacle which is larger than the radius of curvature attained by the dislocation in the field of applied stress σ. (Compare for instance the mechanism of the formation of closed loops in a Frank-Read source, Fig. III–16.)

Owing to the presence of the small dislocation loops around the precipitate particles, the available space Λ for other dislocations to slip

through decreases progressively with the deformation. The over-aged alloys are thus expected to show increasing hardening with increasing strain (cf. WASSERMANN *et al.* [1958], Al-Cu alloys). Indications of this are indeed found in many experiments. The analysis is applicable soon after the dispersion Λ exceeds the critical value (9), because then the Orowan mechanism, which has a yield strength $<4\varepsilon C_B G$ associated with it (according to (11) and (9)), becomes more favourable than the Mott-Nabarro flow stress $2G\varepsilon C_B$, which is independent of the dispersion size.

The theories discussed above do not apply to coherent precipitates, and certainly not to non-spherical coherent Guinier-Preston zones in Al-Cu alloys. There the stresses around the precipitates have the form (7) or resemble dislocation stresses respectively. We return to this in Chapter XXI. Some new interesting aspects of hardening have been discussed recently by FRIEDEL [1959], to which we refer.

3. Solid-Solution Hardening. Effect of Order

FISHER [1954] has pointed out that an alloy showing short-range order (see Chapter XXI) should offer an additional resistance these motion of dislocations, because in moving through the lattice these imperfections destroy completely the short-range order across their slip plane. This can be described mathematically by noting that an area over which a dislocation has passed will have increased in surface energy by a certain amount γ. Expressed as a stress, the short-range order increases the flow stress thus by the amount

$$\Delta\sigma_t = \gamma/b\,. \tag{12}$$

Fisher estimated $\Delta\sigma_t$ for cubic α-brass to be a few kg/mm^2, which is an appreciable fraction of the observed yield strength of α-brass single crystals.

The general problem of *solid-solution hardening* is too extended to be dealt with here more than summarily. Excellent review papers exist (PARKER and HAZLETT [1954], COTTRELL [1954], H. SUZUKI [1958], SEEGER [1958]). The subject is closely connected with that of order-disorder transitions, to be discussed in Chapter XXI.

A solid solution showing either short-range or long-range order will always be harder than a non-ordered alloy, because the dislocations, when moving through the lattice, destroy the symmetry and therewith

increase the free energy of the crystal. Two influences must be distinguished, one which produces a sharp yield point because it determines the stress required to *initiate* the motion of dislocations, and another that acts on the flow stress also in later stages, since it is connected with the general *mobility* of the dislocations in the lattice. The short-range order effect mentioned above is typically a yield strength effect; once the very first dislocations have moved over a glide plane, the short-range order has been destroyed and the subsequent dislocations do no longer suffer the effect of this (ARDLEY and COTTRELL [1953]) (unless all dislocations are of exactly the same type and parallel).

Other hardening mechanisms which affect predominantly the yield stress, or are even responsible for the occurrence of a sharp yield point, have already been discussed previously (Chapter IX); they are impurity locking, chemical (Suzuki) locking in face-centred cubic metals (section IX–1.3) and Snoek ordering of interstitial impurities (section IX–1.4).

An indication about the possible origin of an observed yield point in an alloy can sometimes be obtained by measuring its *temperature dependence*. Cottrell (impurity) locking becomes rapidly less efficient at high temperatures where the impurity cloud around the dislocations begins to dissolve, but Suzuki hardening is usually considered to be independent of temperature. The large area over which the accumulation of the impurities takes place should not permit activation by thermal vibrations. This is not necessarily true, however, as is borne out for instance by the temperature dependence of the impurity concentration in grain boundaries or at the surface, which indicates that also the relative solid solubility in the dislocation itself may vary strongly with the temperature. We cannot go into this problem here further, and only want to point out that chemical hardening need not be independent of temperature at all. Short-range order effects are essentially independent of the temperature up to the *transition temperature* (if present), above which the ordening disappears.

Long-range order does not affect the yield stress, but the presence of distinct kinds of atoms in the lattice may nevertheless influence the mobility of dislocations, whether this distribution of the various atoms is ordered or not. We have already mentioned in Chapter VI that the presence of foreign substitutional impurities extends the length of the easy glide range in copper single crystals. This and similar effects will usually be different in the ordered and disordered state, thus

above and below the transition temperature, in an alloy showing long-range order.

Perfectly ordered alloys do not distinguish themselves much from pure metals with regard to the flow stress, as long as the dislocations which move in them do not affect the ordening. For this the nature of dislocations should be different in a superstructure from in a monoatomic lattice. "Normal" dislocations should become paired two by two to regain the property of transforming the lattice into itself; single, they are only a kind of partial dislocation in the superlattice. One thus should have a kind of super-extended dislocation, the "stacking fault" of which (an *antiphase* region) differs from a normal stacking fault in that now only the phase of order is changed, not the structure. If the alloy is divided into *domains* showing different phases of ordering, the motion of these dislocations from one domain to the other may be appreciably hindered.

ARDLEY [1955] has investigated the effect of the degree of order in Cu_3Au single crystals (transition temperature $T_c = 385$ °C) on the plastic properties, and found the critical shear stress of the ordered alloy to *increase* with temperature between 300° and T_c. This very remarkable effect was explained as due to the gradual decrease of the degree of long-range order in this temperature region: in a partly disordered crystal a superextended dislocation leaves in its wake a region with a lower degree of order than in front. At the critical temperature the flow stress falls abruptly by many tens of percents, and far above this temperature (near 600 °C) it starts to decrease again very rapidly, owing to the disappearance of the final traces of short-range order.

It is hardly necessary to remark at the end of this chapter that, once actual regions of second phase have been formed, precipitation hardening occurs (section 2), which does not lead to a yield phenomenon but affects the flow stress in all stages of deformation as discussed previously.

REFERENCES

ARDLEY, G. W. (1955); Acta Metall. **3,** 525.
ARDLEY, G. W. and A. H. COTTRELL (1953); Proc. Roy. Soc. **A219,** 328.
COTTRELL, A. H. and B. A. BILBY (1949); Proc. phys. Soc. **A62,** 49.
COTTRELL, A. H. (1953); Dislocations and Plastic Flow of Crystals (Clarendon Press, Oxford).
COTTRELL, A. H. (1954); Trans. ASM **46A,** 131. (Relations of properties to microstructure).
CRUSSARD, C. (1950); Métaux et Corrosion **25** (301), 203.
ESHELBY, J. (1958); Ann. d. Physik **1,** 116.
FISHER, J. C. (1954); Acta Metall. **2,** 9.
FRIEDEL, J. (1959); Internal Stresses and Fatigue in Metals, Symp. Detroit, p. 220 (Elsevier, Amsterdam).
HAM, F. S. (1959); J. applied Physics **30,** 915.
MOTT, N. F. (1950); Imperfections in Nearly Perfect Crystals, p. 173 (Wiley, New York).
MOTT, N. F. and F. R. N. NABARRO (1948); Rep. Conf. Strength in Solids, Bristol, p. 1 (The Physical Society, London).
OROWAN, E. (1947); Symposium on Internal Stresses, London.
PARKER, E. R. and T. H. HAZLETT (1954); Trans. ASM **46A,** 30 (Relation of properties to microstructure).
SCHOECK, G. and A. SEEGER (1958); Detroit Conference on Mechanical Properties of Crystals, see also Acta Metall. **7,** 469.
SEEGER, A. (1958); Encyclopedia of Physics, Part VII–2, section 60 etc. (Springer, Berlin).
SNOEK, J. L. (1941); Physica **8,** 711.
SUZUKI, H. (1952); Science Rep. Tohoku Univ. Japan, **A4,** 455.
SUZUKI, H. (1958), in: Dislocations and the Mechanical Properties of Crystals p. 361 (Wiley, New York).
WASSERMANN, G. and H. RICHTER (1958); Z. Metallk. **49,** 549.

CHAPTER XI

ANNEALING, POLYGONIZATION AND RECOVERY

1. Effects of Annealing

A plastically strained metal is in a state of internal stress, and thus its free energy is enhanced over that in the stress-free state. The situation is one of thermodynamically unstable equilibrium. The material tries to return to a state of smaller free energy, but this is in general only possible at elevated temperatures, where diffusion and other thermally activated processes can help in this tendency. A suitable *annealing* treatment, that is, keeping the deformed metal for a certain time at a high temperature (very roughly about half its absolute melting temperature or higher) indeed usually results in a decrease of the hardness and an increase of the ductility, a removal of internal strains and a partial *recovery* of the imperfect lattice structure of the metal.

The atomic mechanisms responsible for these annealing effects are several. One of the most important of them is diffusion, particularly self-diffusion, which is in most metals, as we shall see in Chapter XVIII, brought about by vacancy migration. The presence, in thermal equilibrium, of a sufficient amount of vacancies, together with the possibility for their migration, enables the *climb motion* of dislocations to take place (Chapter III). This can result inter alia in a partial annihilation of the many dislocations formed during cold-work, an a rearrangement of the remaining dislocations in energetically more favourable arrays (characterized by the absence of long-range stresses). A specific internal structure develops, which has been called the *polygonized* structure. The whole process bears the name *polygonization*. It often (but not invariably) accompanies the release of internal strain called recovery, and is then responsible for a large part or all of it.

Occasionally observed during an annealing treatment is the *migration of crystal boundaries* under the effect of the internal strains. This

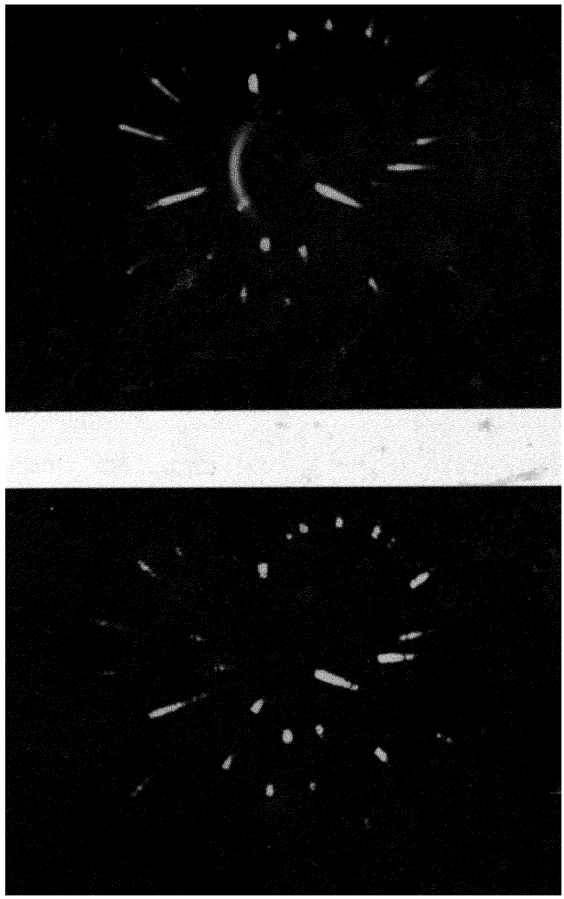


Fig. 1. *a*) Asterism in Laue-diagram of a 7% extended aluminium single crystal. *b*) Same diagram after 36 hours annealing at 590° C (after CAHN [1949]).

results in sweeping the metal partially free of imperfections. This process is often considered as a form of *recrystallization* rather than of recovery (BECK and SPERRY [1950]). The distinction between the latter process and those called recrystallization and *grain growth* is often not very clear.

During recrystallization certain nuclei form in the deformed matrix which grow out into complete, practically stress-free crystals at the cost of the old grains. The process is thus one of nucleation, followed by one of growth. Although in practical cases recrystallization and recovery usually take place at different temperatures, both mechanisms can in a certain sense be regarded as competitive. A characteristic aspect of recrystallization proper is that strong dependence on traces of impurity exists. Heavily cold-worked super-pure (zone-melted) aluminium has been shown to undergo recrystallization already below room temperature (CHAUDRON [1954]), before recovery takes place. Normally, however, recovery occurs first and prevents subsequent recrystallization. The latter process is further severely influenced by the presence of internal strain and imperfections such as grain boundaries and subboundaries. We shall discuss recrystallization in a separate chapter (XX).

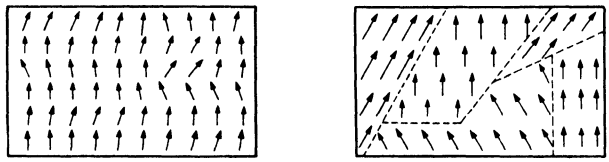
2. Polygonization

As has already been mentioned, polygonization consists in the spatial rearrangement and partial annihilation of dislocations. It was originally observed in the Laue diffraction pattern of cold-worked and subsequently annealed metals. In cold-worked metals this pattern is characterized by *streaks* instead of *dots* as in a perfect crystal; the streaks are preferentially oriented radially away from the spot due to the central beam (Fig. 1*a*). This is called *asterism*, and for a detailed description and explanation we refer to Barrett's book (see also Chapter XVI). The elongated Laue-spots indicate a continuous range of small orientation differences in the irradiated portion of the crystal (Fig. 2*a*). This is not surprising, in view of the many dislocations with their internal strains present in the cold-worked material.

On annealing the metal, in most cases the asterisms are seen to break up into strings of separated spots (Fig. 1*b*). These still retain the envelope form of the streaks, but the discrete structure points to the existence, not of a continuous range of orientations but of a discrete

number of distinct orientations, differing by small angles (of the order of a few minutes of arc, Fig. 2*b*).

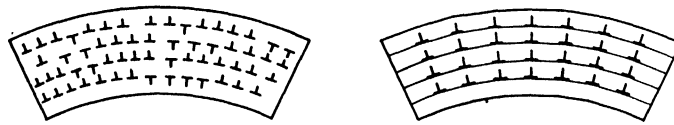
CAHN [1949] was able to demonstrate this breaking-up process in a very pronounced form by bending single crystals of zinc to a radius of about 1 cm and then annealing these just below the melting point.



a *b*

Fig. 2. *a*) In a cold-worked metal the lattice orientation changes continuously from place to place. *b*) After annealing, separate blocks of definite but mutually different orientations occur.

The latter author has proposed the following dislocation mechanism to explain the observed facts. For simplicity we consider a bent crystal (constant radius of curvature) in which only edge dislocations parallel to the axis of bending have been formed. (In general also other dislocations will have been generated, although the edge dislocations alone observed in the Laue diffraction could account for the required lattice deformation.) Parallel dislocations of both signs will be present after the cold work (Fig. 3*a*) but there will exist a net difference, such



a *b*

Fig. 3. *a*) More or less random distribution of dislocations of different sign (but with net excess) in a cold-bent crystal. *b*) After polygonization the dislocations have formed a subboundary structure by the processes of rearrangement and annihilation.

as to produce the observed curvature. It is easily found (see e.g. COTTRELL [1953]) that this net difference must be everywhere the same in the material:

$$\Delta = \frac{1}{\varrho b} \text{ cm}^{-2}, \tag{1}$$

where ϱ is the radius of curvature. Formula (1) is an example (a very simple one) of the dislocation distribution derived from a given distortion (see section IV–5 and NYE [1953]). On annealing, the dislocations leave their glide planes by *climb* and annihilate each other until only about the net density (1) of one sign is left (hence the radius of curvature remains unaltered). These arrange themselves in the form of so-called *subboundaries* of the tilt type (compare Chapter XIX on grain boundaries). This situation is depicted in Fig. 3*b*, and indeed corresponds to the existence of a relatively small number of fairly perfect regions with discrete mutual orientation differences (*subgrains*), in agreement with the X-ray evidence.

The correctness of Cahn's model has been strikingly verified by the recently developed etchpit techniques for the indication of dislocations (Chapter V). Examples have been given by LACOMBE *et al.* [1947] of polygonized aluminium crystals. In the specimen as bent, the etchpits are numerous and distributed at random. In the polygonized specimen they lie along straight lines, which delineate the subboundaries. The observed orientation differences between the subgrains agree very well with the number of dislocations in these boundaries, as will be explained in Chapter XIX.

Polygonization has been observed to occur in many different substances, and the dislocation mechanism seems fairly well substantiated. It has been extensively described by AMELINCKX and DEKEYSER [1959].

3. Recovery as a Thermally Activated Process

It will be clear that when a metal has polygonized completely it has become softer and more ductile than it was when in the cold-worked state. Work-hardening has been practically removed, because once the dislocations are aligned, their long-range stresses are no longer present (see Chapter XIX). Further the number of other lattice imperfections has been severely reduced by diffusion; only the subboundaries have remained.

It is a common observation, however, that softening of a metal need not be accompanied by polygonization, and occurs even in those cases where polygonization does apparently (within the accuracy of the X-ray or microscope technique) not take place at all (for instance when no excess of dislocations of one sign is present, as in cubic crystals deformed in easy glide without lattice curvature). To judge from these

observations, the softening would appear to be primarily associated with the mutual annihilation of parts of the dislocations, by (thermally activated) motion over barriers or across screws, and not so much with their lining up into subboundary dislocation walls.

However, X-ray observations of the recovery of aluminium (NOWICK [1955]) showed that even *before* noticeable softening occurs already the internal strains in the metal have appreciably decreased, either through local rearrangement or through annihilation. X-ray methods are obviously considerably more sensitive in detecting variations in the dislocated state than mechanical observations. LUTTS and BECK [1956] have studied the kinetics of this "first" recovery stage in more detail. They found that hardness as well as X-ray line sharpening follow a temperature-dependent time law, with an activation energy that is approximately 1 eV in aluminium, which is lower than that of self-diffusion (1.4 eV) in this metal. After the first stage has been fully completed (completely sharp X-ray diffraction lines), there remains yet about half of the original work-hardening. PERRYMAN [1956] showed that the addition of other atoms (Mg in Al) decreases the activation energy for recovery. He explained this as the result of the trapping of vacancies, formed during the deformation process, by the foreign atoms. These vacancies, released again during the anneal, accelerate the recovery process, whatever its precise nature may be and speed up the mutual elimination of dislocations.

This mutual annihilation was considered by COTTRELL and AYTEKIN [1950] and by KUHLMANN *et al.* [1949] as a process the activation energy of which is *a function of the internal stress* in the metal. Whether the important mechanism is climb, as is certainly necessary for real polygonization to occur, or activated glide motion over barriers, is not important for this formal description. Here, we shall assume that barriers occur existing of dislocations of preponderantly screw character. As explained in section III-5.3, the activation energy for the crossing of two perpendicular screw dislocations is of the form:

$$U = U_j - \sigma b^2 l_0 , \tag{2}$$

where U_j is the energy to form a jog, and l_0 is the mean distance between crossing points. σ is here to be regarded as the average internal shear stress along the Burgers vector of the dislocation. Assuming that *this* activation energy governs the recovery process, the rate of

decrease of the mean magnitude of the internal stress is proportional to exp $(-U/kT)$. Thus one has

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = -C \exp\left(-\frac{U_1 - \sigma b^2 l_0}{kT}\right). \tag{3}$$

This equation has as a solution:

$$\sigma = \sigma_0 - \frac{kT}{b^2 l_0} \log\,(1 + t/t_0)\,, \tag{4}$$

where the initial stress σ_0 and the constants C and t_0 are interrelated by

$$\sigma_0 = -\frac{kT}{b^2 l_0} \log \frac{C b^2 l_0 t_0}{kT} - \frac{U_1}{b^2 l_0}\,. \tag{5}$$

A logarithmic time law of the form (4) for the time decay of the flow stress and the hardness has indeed been observed by Cottrell and Aytekin and by others. The flow stress finally attained depended further on the initial amount of cold-work, through l_0 as must be (MOTT [1953]). For zinc single crystals, Cottrell and Aytekin obtained $U_1 \approx 1.4$ eV and $\sigma_0 b^2 l_0 \approx 0.5$ eV. Taking $b = 3 \times 10^{-8}$ cm, $l_0 = 10^{-5}$ cm, one has $\sigma_0 \approx 10^8$ dyn/cm^2 = 1 kg/mm^2 which is reasonable, as is the value of U_1.

Of course, the process characterized by an activation energy of the form (2) need not be the mutual crossing of dislocations at all. It is difficult to find another process however, which has definite advantages over the former.

The climb process itself, leading to the formation of subboundaries, is also associated with an energy of activation, but this does not depend on stress except in a typical discontinuous way indicated below. Climb (see section III–5.2) of a dislocation with some edge-character necessitates a change of dimensions of the "extra half plane", that is, it requires the addition to or removal from it of atoms. This occurs by diffusion, i.e. self-diffusion in the metal. In most metals self-diffusion is a vacancy mechanism; then climb of a dislocation occurs by *absorption* or *emission* of *vacancies* at dislocations.

The sites along a dislocation which are best suited for these processes are the jogs (Chapter III). The number of jogs per unit length of a dislocation is in thermodynamic equilibrium given by:

$$n = n_0 \exp\,(-U_1/kT)\,, \tag{6}$$

where n_0 is the number of atomic sites per unit length of dislocation. In thermodynamical equilibrium the density of vacancies in the metal is also given by an exponential expression in which occurs the free energy of formation of a vacancy, U_v. Finally, the migration of vacancies is again governed by an activation energy U_m. Climb in thermal equilibrium will therefore be characterized by the activation energy

$$U_c = U_j + U_v + U_m = U_j + U_s ,$$

where U_s is the energy of activation for self-diffusion.

In plastically deformed metals the concentration of jogs is usually much larger than that corresponding to thermal equilibrium, because during the recovery process the dislocations do not only climb but also glide through the forest of other dislocations, thereby acquiring a full stock of jogs. Since during this process vacancies are also formed in large numbers, for the rate of climb the migration of these vacancies is the only determining process, and this is characterized by the activation energy U_m only. This situation is realized when the internal stresses are large, or when an additional external stress is applied, thus for instance in *creep* and *fatigue* tests. Indeed has it been found that polygonization during creep or fatigue occurs at much lower temperatures than in static tests. To give an example: whereas predeformed copper specimens usually do not polygonize on annealing, they did so during a fatigue test at a slightly elevated temperature (BROOM and HAM [1957]). The so-called *cell*-formation in aluminium during creep tests at about 250 °C, as observed by WOOD and co-workers [1949], constitutes another well-known example (section XVI–2.3.4).

During *stress-free* annealing of previously deformed metals, vacancies must be formed and displaced, but the dislocations usually contain already so many jogs at which this can occur that the activation energy for *self-diffusion* alone determines the rate of climb and thus of polygonization. As SEEGER [1955] has pointed out, this conclusion is no longer valid in slightly deformed metals of medium stacking fault energy like copper (cf. Chapters VII and VIII), in which the energy of formation of a jog is so very large that the equilibrium number of jogs is limited. Then the *sum* of the energies of jog formation and self-diffusion governs the recovery rate, and this is usually so large that only little recovery occurs in practice. Thus the stacking fault energy plays an important role in determining whether polygonization after slight deformation

shall occur or not; the different behaviour of Al and Cu must be explained in this way. (See also section XII–3).

Polygonization bears a superficial resemblance to *fragmentation*, that is, the breaking-up of the lattice into very fine coherent domains, of about 10^{-4} cm diameter, which occurs eventually after cold-working *without* annealing. The structure obtained differs from the cells formed during creep (at higher temperatures) in the size of the fragments. This is of the order of one micron or less in fragmentation, and of 10 microns during cell-formation (HIRSCH [1957]).

Fragmentation has been observed at −180 °C in zinc and in bent tungsten single crystals at room temperature which is relatively a very low temperature (0.1 T_m) for this metal. To explain the process it has been assumed that polygonization on a very fine scale occurs by migration of the vacancies, formed during the plastic deformation, to dislocations in their near neighbourhood. The latter are then caused to climb over *small* distances. However, the occurrence of fragmentation already at very low temperatures makes it improbable that this process has anything to do with vacancy migration. Presumably it is a purely geometrical dislocation phenomenon, to be compared with kinking (section XIII–4).

BECK [1954] has pointed out that as far as the energy stored in a fragmented metal during cold-work is concerned, this is released for the larger part during the pre-polygonization stage of recovery, whereas the rest disappears during migration of the subboundaries, formed during the cold-work; this may be considered as recrystallization). A certain part of the work-hardening is thus presumably connected with the formation of the fragmentation structure.

4. Grain Boundary Migration

A quite different recovery mechanism, occurring in polycrystalline metals, is the migration of grain boundaries. In the course of this migration, heavily strained regions of the material find their excess dislocation and other imperfection content removed by the migrating boundary. As mentioned, in a *fragmented single* crystal the fragmentation boundaries migrate also: this special kind of recovery by *subboundary* migration was called *recrystallization in situ* by CRUSSARD [1944].

As the atomic motions during grain boundary migration are similar

to those during self-diffusion, it was suggested by BECK [1954] and others that the activation energy for this process would be that of self-diffusion. However, in practice, much higher energies are found. MOTT [1948] (see also section XIX–3) tried to correlate the energy of activation with that of fusion, L, by assuming the boundary to consist of disordered ("molten") islands imbedded in regions of good fit. The observed activation energy would then be equal to nL, where n is the mean number of atoms that together form a melting and subsequently recrystallizing region during the migration of a boundary. Taking $n \approx 20$, agreement can be obtained with the experimental data.

More will be said about this in the Chapters XIX and XX on grain boundaries and recrystallization respectively.

REFERENCES

AMELINCKX, J. and W. DEKEYSER (1959); Solid State Physics **8**, 441.
BARRETT, C. S. (1953); The Structure of Metals (McGraw-Hill, New York).
BECK, P. A. (1954); Adv. in Physics **3**, 245.
BECK, P. A. and A. SPERRY (1950); J. applied Physics **21**, 150.
BROOM, T. and R. K. HAM (1957); Vacancies and other Point Defects (London, Inst. of Metals) p. **41**.
CAHN, R. W. (1949); J. Inst. Metals **79**, 129.
CHAUDRON, G. (1954); Nature (London) 197, 923.
COTTRELL, A. H. (1953); Dislocations and plastic flow of crystals (The Clarendon Press, Oxford).
COTTRELL, A. H. and V. AYTEKIN (1950); J. Inst. Metals **77**, 389.
CRUSSARD, C. (1944); Rev. Métall. **41**, 118 and 139.
HIRSCH, P. B. (1957); Progress in Metal Physics **6**.
KUHLMANN, D., G. MASING and J. RAFFELSIEPER (1949); Z. Metallk. **40**, 241.
LACOMBE, P. and L. BEAUJARD (1947); Rev. Métall. **44**, 65; **45**, 317.
LUTTS, A. H. and P. A. BECK (1956); Trans. AIME **210**, 1226.
MOTT, N. F. (1948); Proc. Phys. Soc. **60**, 391.
MOTT, N. F. (1953); Phil. Mag. **44**, 753.
NOWICK, A. S. (1955); Acta Metall. **3**, 312.
NYE, J. F. (1953); Acta Metall. **1**, 153.
PERRYMAN, E. C. W. (1956); Trans. AIME **210**, 1247.
SEEGER, A. (1955); Rep. Conf. Defects in Cryst. Solids, Bristol, p. 391 (The Physical Society, London).
WOOD, W. A. and W. A. RACHINGER (1949); J. Inst. Metals **76**, 372.

CHAPTER XII

CREEP

1. Introduction

By *creep* we denote here the time dependent part of the plastic deformation of metals. (We have already encountered several examples of creep phenomena, e.g. the pre-yield creep discussed in section IX–4.) A metal loaded below its flow stress at a sufficiently elevated temperature usually shows a gradual plastic deformation. This indicates that dislocations and other imperfections, although unable to execute large-scale motions, are yet not completely immobile at stresses below the critical stress needed to move them over their barriers.

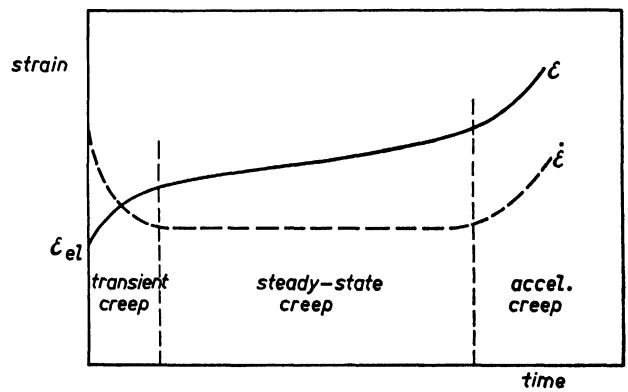


Fig. 1. Schematic creep curve (full-drawn) of a face-centred cubic metal under constant stress. The variation of the creep *rate* has been indicated by the dotted curve.

A typical creep curve (after Andrade) of a metal under constant stress is shown in Fig. 1. After the instantaneous elastic extension ε_{el}, the strain increases gradually, although at a decreasing rate. This is

called *primary* or *transient creep*. Afterwards the strain rate becomes approximately constant and we have entered the region of *secondary* or *steady-state creep* (quasi-viscous creep). After a considerable strain has been reached the creep rate may increase again: *tertiary* or *accelerated creep*, ending finally in rupture. This trend is common to many crystalline materials. Particular forms of creep occurring under special circumstances only are connected with the sliding of grains past each other: *grainboundary creep*, and with the migration of vacancies: *vacancy creep*.

Creep as a form of plastic flow is usually (but not always) caused by the motion of dislocations, but differs from ordinary plastic deformation in that *thermal activation* is now an essential factor, whereas this plays but a negligible role in ordinary deformation. Therefore, creep shows most conspicuously at elevated temperatures, where recovery may take place simultaneously. However, it must be emphasized that creep as such is *not* a typical high-temperature phenomenon, it can as well occur at room temperature or below. For experimental and technical particulars of this complex but important subject, we must necessarily refer to the vast literature of the field (see for instance SULLY [1956], SCHOECK [1956]). Here we shall discuss only the interpretation of the creep phenomenon in terms of the properties of lattice imperfections.

2. Transient Creep

The early part of the creep curve, that which is characterized by a decreasing strain rate, can be empirically described by a time law of the form

$$\dot{\varepsilon} = At^{-n}, \quad 0 < n < 2. \tag{1}$$

In the special case that n is unity, one speaks of *logarithmic creep*, because after integration $\varepsilon \sim \log t$. The value $n = \frac{2}{3}$ was commonly found in the past in polycrystalline metals and many other materials (ANDRADE [1914]), and this particular dependence has therefore also acquired a special name: *Andrade creep*, after its foremost investigator. The coefficient A in (1) depends on the temperature and the loading stress in a rather undefined manner. Transient creep is conspicuous in previously cold-worked polycrystalline or in precipitation-hardened samples, thus in samples in which numerous obstacles for dislocations exist. For a review of transient creep in face-centred cubic metals see,

for instance, BLANK [1958]. A very detailed experimental investigation of the fundamental properties of creep in polycrystalline aluminium and copper has been carried out by WYATT [1953].

The characteristics of transient creep led first to its description as an *exhaustion* phenomenon. MOTT and NABARRO [1948] thought of it as due to the stress-induced motion of dislocation segments held up by obstacles that could be surpassed by the aid of relatively slight thermal activation. Each segment could only once be active in this way, and in the course of time the number of available suitable segments will become exhausted. The strain rate decreases accordingly. COTTRELL [1952] has applied the theory in a modified form to Wyatt's results.

To take account of the observed increase of the dislocation density during transient creep and the occurrence of a certain degree of work-hardening, the exhaustion theory has been replaced by a theory of thermally activated dislocation motion of a somewhat wider scope (MOTT [1953]). It is assumed that the motion of all dislocations is restricted by the processes of dislocation intersection, formation of constrictions, etc., all of which are characterized by a strain-dependent energy of activation

$$U(\sigma) = U_0 + \alpha\varepsilon \tag{2}$$

(cf. section III–5.3). This agrees with the observations by Wyatt (see below).

Under the usual creep conditions the thermally activated motion of dislocations gives rise to a creep rate

$$\dot{\varepsilon} = \nu_0 AbN \exp(-U/kT), \tag{3}$$

where N is the number of dislocation segments available and A the mean area covered by them in one jump; ν_0 is again the atomic vibration frequency.

To explain the observed time dependence of the creep rate, both the quanties N and U are considered as functions of time. This includes not only the possibility that the number of available segments becomes exhausted: N decreases gradually to zero, but also that the dislocation density increases and work-hardening takes place simultaneously: N and U both *in*crease with time. The latter alternative agrees best with observations. To be able to integrate (3) in closed form Mott considered U to depend linearly on the creep strain (this is equivalent to assuming

the internal stress to be proportional to the strain). Taking first N constant, Mott thus obtains the differential equation

$$\dot{\varepsilon} = \nu_0 \, AbN \exp\{-U(t=0)/kT\} \cdot \exp(-\alpha\varepsilon/kT),$$

the solution of which reads

$$\varepsilon = \frac{kT}{\alpha} \log\left\{\frac{\nu_0 AbN\alpha}{kT} t \cdot \exp - \frac{U(t=0)}{kT} + 1\right\}, \tag{4}$$

where α is a proportionality factor of the dimension of an energy. Thus indeed a logarithmic time law is obtained under these conditions. Agreement with experiments is found by taking the *activation volume* v of the order of 10^{-19} cm$^3 \approx 10^3 - 10^4 \, b^3$, in fair accordance with the underlying model.

Exponents in (1) smaller than unity are formally obtained by assuming N to *in*crease progressively with the strain (and U proportional to strain as before). Such a gradual increase of the dislocation density is commonly observed, and the occurrence of exponents n smaller than unity is therefore not surprising, although there is no preference in the theory for $n = \frac{2}{3}$ (Andrade creep). Some discussion on possible other forms of the function $U(\sigma, t)$ can be found in Mott's and Wyatt's papers.

Wyatt observed the exponent n in transient creep of cold-worked aluminium to be unity only at very low temperatures (−196 °C). At high temperatures the exponent changes to $\frac{2}{3}$. This is a general observation, and was first explained by MOTT [1953] as the result of the occurrence of *recovery* simultaneously with the creep process. This recovery, in the form of dislocation *climb*, makes it possible for otherwise strongly bound dislocations to escape from piled-up groups or other places (see previous chapter) and also to take part in the creep process. The dislocation density thus now increases, whereas it remained more or less constant during low-temperature creep. The fundamental assumption leading to the exponent $n = \frac{2}{3}$ is now (MOTT [1953]) that the changes in the internal state due to the "escape" of a dislocation from a pile-up are distributed *randomly* in sign. To make this clear we shall consider a particular case.

Let σ_1 be the change in the local internal stress caused by the escape of one dislocation. $|\sigma_1|$ will usually be much smaller than σ_a, the change of stress that would be necessary to produce a "creep jump" of another dislocation segment. Then, if the sign of σ_1 is arbitrary, the

number of escapes necessary to produce a creep jump is of the order $(\sigma_a/\sigma_1)^2$. Assuming a linear stress-strain curve to apply (it is only the early part that interests us here) σ_a is proportional to the strain, hence the *rate* of creep, inversionally proportional to the number of escapes, is:

$$\dot{\varepsilon} \sim {\sigma_a}^2 \sim \varepsilon^{-2} , \tag{5}$$

which is equivalent to

$$\dot{\varepsilon} \sim t^{-2/3} . \tag{5a}$$

It is thus the fundamental statistical character of the release process that is responsible for the time dependence of Andrade creep. In view of the universal occurrence of the $t^{-2/3}$ law (even in non-crystalline materials), this supposition seems most plausible, although the exact imperfection mechanism responsible may differ widely from case to case, and from material to material. Mott has applied the theory to metals, and compared it to the stress and temperature dependence of Andrade creep of steel observed by FELTHAM [1956]. The agreement is fair, but rough. KENNEDY [1953] has reviewed Andrade creep from a wide variety of data pertaining to non-metallic solids, with similar results.

Wyatt's experiments on pure metals indicate clearly a linear dependence of the activation energy on the activation stress, in accordance with the assumption (2), but observations by DAVIS and THOMPSON [1950] on precipitation-hardened copper alloys pointed to a $\sigma^{2/3}$-dependence. Other observers find the exponent n to be approximately equal to T/T_m, where T_m is the absolute melting temperature. These questions have not yet been answered satisfactorily.

3. Steady-State Creep

We have already mentioned in the previous section that the occurrence of deformation, thus work-hardening, and recovery simultaneously leads to a time law for the rate of creep which is characterized by an exponent smaller than unity. The most common type of creep phenomenon, extensively observed because it is of great technical importance, is steady-state creep, which is characterized by smallest possible exponent, viz. $n = 0$:

$$\dot{\varepsilon} = \text{const} ,$$

thus

$$\varepsilon = kt . \tag{6}$$

The proportionality factor k is usually an exponential function of temperature (containing an activation energy) and an only slightly less sensitive function of stress. In practice this kind of creep takes place in heavily loaded constructions at elevated temperatures.

It is not surprising that, to explain steady-state creep, the fundamental assumption has been put forward that here work-hardening and recovery should exactly *balance*, the effect of the increase of internal stress being just compensated by the thermally activated escape of dislocations over barriers.

This idea, originally advanced by OROWAN [1947], has been developed by MOTT [1953] and worked out more quantitatively by WEERTMAN [1955]. The latter author considers the stress-induced *climb* of dislocations as the rate-governing process in steady-state creep. Whether this climb occurs over Lomer-Cottrell barriers in front of piled-up groups, or over precipitate particles, is not essential to the theory. The activation energy for climb being equal to that for self-diffusion when this is considered as a vacancy mechanism, one should expect the activation energies for steady state creep (thus occurring in k of (6)) and for self-diffusion to be the same. This is borne out by many experiments on metals (see DORN [1954, 1956], LIEH HUANG *et al.* [1956], SHERBY *et al.* [1957], WEERTMAN [1957]), and also on other types of crystals. LYTTON *et al.* [1958] found in single crystals of aluminium at least three distinct creep mechanisms to take part, with different activation energies: climb (above 600 °K), cross slip (between room temperature and 370 °K) and the fundamental dislocation friction (below 100 °K).

When considered more carefully, the activation energy for climb should depend on the stacking fault energy, as has already been discussed in the previous chapter (see also THORNTON and HIRSCH [1958]. In a metal with a high stacking fault energy the thermal equilibrium concentration of jogs in dislocations is so high that the formation of sufficient point defects is not limited by it. In a metal with a low stacking fault energy the jog concentration is determined by the geometry of the deformation rather than by the temperature. In both cases the activation energy for climb is equal to the sum of the energies of formation and migration of the defects, that is, to the energy for self-diffusion. In between, however, the energy of formation of jogs and constrictions cannot be neglected and the climb energy is enhanced over the diffusion energy.

The stress dependence of the creep rate was derived by Weertman

from a detailed consideration of the atomic mechanism of climb. The result is a powerlaw-type stress dependence, which together with the temperature dependence discussed above leads to

$$\dot{\varepsilon} = \text{const}\, \frac{\sigma^p}{kT} \exp\left(-U_s/kT\right) \tag{7}$$

for not too high stresses. The exponent p lies between 3 and 4, depending on the structure of the material. Such a powerlaw, with exponent around 4, had indeed been observed in a number of metals and non-metals. Part of the stress dependence (σ^2) arises from the dependence of the rate of climb on the stress, the other part from the variation of the necessary climb distance with the stress.

At very high stresses (above 10 kg/mm^2 in most metals) the creep rate sometimes begins to increase approximately exponentially with the stress. No theoretical explanation has been given for this yet.

Weertman's interpretation of the stress dependence differs from an earlier one by COTTRELL and AYTEKIN [1950], who considered the stress variation as due to the presence of a stress-dependent energy of activation. The rate-governing process would not be climb, but the formation of constrictions and jogs, and the considerations of section XI–3 would apply. An exponential stress dependence has not been observed at low stress levels, however.

4. Creep in Nearly-Perfect Non-Metallic Crystals

In a number of polar and homopolar materials a type of temperature-activated deformation is observed that bears some resemblance to high-temperature creep in metals. However, these non-metals are usually brittle at low temperatures, and only become ductile on heating. Examples are germanium, silicon, InSb and a number of ionic crystals.

In the former materials, owing to the homopolar character of the atomic bonding forces, the dislocations can only move with difficulty. In other words, the Peierls' force is high, and the dislocations need to be thermally activated in order to be able to move at all. A discussion of this type of creep, which is interesting because it gives information on the properties of individual dislocations, will be postponed to Chapter XXX dealing with homopolar crystals.

5. Grainboundary Creep and Vacancy Creep

Polycrystalline metal samples stressed at a high temperature often show the phenomenon of grainboundary sliding. Probably this process is common to polycrystalline materials, together with other forms of creep (McLEAN [1957]). The sliding is again a thermally activated phenomenon. MOTT [1948] proposed the elementary process to be the disordering ("melting") of groups of atoms in the boundary; activation energy is the latent heat of melting of the group. More about this will be said in Chapter XIX and in Chapter XVII on internal friction. (See also the previous chapter and the book by FRIEDEL [1956].) WEERTMAN [1955] proposed a sliding mechanism based on dislocation climb.

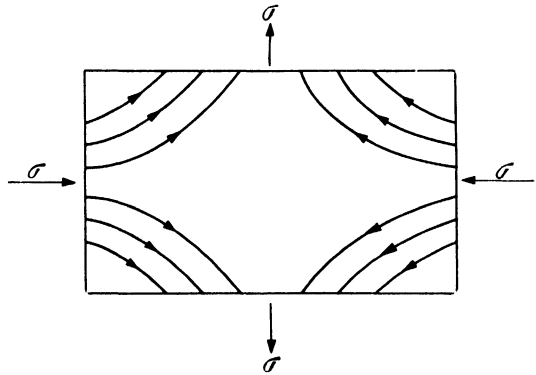


Fig. 2. To illustrate vacancy creep. The arrows denote the motion of the *atoms*.

Creep can in principle be the outcome also of the stress-induced internal transport of material through *diffusion* of point defects in one or more preferred directions (NABARRO [1948]). Consider a block of material acted upon by the system of stresses σ of Fig. 2 (which is equivalent to a single *shear* stress $\sigma\sqrt{2}$ under 45° with the edges). If vacancies are generated at the tensile faces, move along the arrows to the compressed faces and become trapped there, this is equivalent to a transport of atoms in the opposite direction. By such a transport the stresses are partly relieved. Conversely, the stress system will give rise to a directed flow of vacancies, if the condition for sufficiently easy formation and absorption of vacancies is realized, that is, at *very* high temperatures only. To find the equilibrium creep rate, Nabarro reasoned as follows. The formation of a vacancy in a stress-free region

requires an energy U_v. In the presence of a suitable stress, such as is present at the tensile faces of the block in Fig. 2, this energy is decreased. The area which each atom exposes to the stress σ be a^2; then the work done by the stress on removing the atom is σa^3, and the effective energy of formation of a vacancy has become $U_v - \sigma a^3$. In thermal equilibrium, the concentration of vacancies on the tensile faces is thus higher than that on the compressed faces by the factor

$$\exp\left(\frac{2\sigma a^3}{kT}\right).$$

This concentration difference gives rise to a diffusion flow that is in good approximation ($\sigma a^3 \ll kT$) given by

$$n = \beta \frac{D}{a^3 l} \cdot \frac{2\sigma a^3}{kT} \text{ (atoms per unit area per second)}, \tag{8}$$

where l is the dimension of the block and D the coefficient of self-diffusion; β is a numerical factor.

As each transported vacancy produces an increment of strain of a^3/l, the strain rate is

$$\frac{d\varepsilon}{dt} = \beta \frac{2\sigma a^3 D}{l^2 kT}. \tag{9}$$

The proportionality constant β has been computed by HERRING [1950] to be about 12. A more exact calculation yields about 10 times as much, and this is in better agreement with several observations on the rate of creep at very high temperatures (within a few degrees from the melting point). As follows from formula (9) the effect will be most conspicuous in wires of small cross-section stressed along their axis. The efficiency of the surface as a source and sink of vacancies has in this theory been taken for granted. In actual cases, grain boundaries and polygonization walls are much more efficient generators, and for l the grain (or subgrain) diameter should be taken. Internal friction experiments on Al by BOULANGER *et al.* [1956] seem to confirm this. Vacancy creep appears in normal cases to be less important than other creep processes; it determines in a sense an upper limit to the creep strength of metals. Vacancy creep appears to play an important role in the deformation of non-metallic materials and in *sintering*. CHANG [1958] has studied the creep and variation of the elastic constants

of Al_2O_3, and found the data to fit the Herring-Nabarro formula to a reasonable degree.

During creep, the structure of a metal changes, as can be observed by X-ray and microscopical methods. Not only do grains sometimes slide along each other, but polygonization accompanies the recovery occurring during Andrade and high-temperature steady-state creep. The formation of subboundaries lining up to form so-called *cells*, is one of the most conspicuous phenomena (WOOD and RACHINGER [1949]); it has already been mentioned in section XI–3. A good review of this subject was given by Mme WEILL [1952]. The slip band configuration of a metal deformed in creep is much less clear than that of a metal strained above its flow stress. Very fine slip lines (fine slip) are sometimes observed in the earlier stages (MCLEAN [1957]), which make place for large and irregular bands in later stages. The inhomogeneity of the deformation generally increases during the creep process, possibly leading up to the as yet not satisfactorily described phenomenon of accelerated creep.

A rather extended microscopical X-ray investigation of the structural changes in aluminium during creep has been published by WYON [1953], whereas the phenomenology and the substructural effects of creep in ferritic iron and stainless steel have been studied by TAMHANKAR *et al.* [1958]. Important differences between the latter two systems were noted, the austenitic steel showing little polygonization and much grainboundary sliding, the iron polygonizing rapidly, while simultaneously grainboundary *migration* took place (see Chapter XI–4).

REFERENCES

ANDRADE, E. N. DA C. (1914); Proc. Roy. Soc. **A90**, 329.
BLANK, H. (1958); Z. Metallk. **49**, 27.
BOULANGER, C. and C. CRUSSARD (1956); Rev. Métall. **53**, 715.
CHANG, R. (1958); unpublished.
COTTRELL, A. H. (1952); J. Mech. Phys. Solids **1**, 53.
COTTRELL, A. H. and V. AYTEKIN (1950); J. Inst. Metals **77**, 389.
DAVIS, M. and N. THOMPSON (1950); Proc. phys. Soc. **B63**, 847.
DORN, J. E. (1954); J. Mech. Phys. Solids **3**, 85.
DORN, J. E. (1956); Creep and Recovery (A S M, Cleveland).
FELTHAM, P. (1956); Proc. phys. Soc. **B69**, 1173.
FRIEDEL, J. (1956); Les Dislocations (Gauthier-Villars, Paris).
HERRING, C. (1950); J. appl. Physics **21**, 437.
KENNEDY, A. J. (1953); J. Mech. Phys. Solids **1**, 172.

LIEH HUANG, H. I., O. D. SHERBY and J. E. DORN (1956); Transact. AIME **210**, 1385.
LYTTON, J. L., L. A. SHEPHARD and J. E. DORN (1958); Trans. Met. Soc. AIME **212**, 220.
MCLEAN, D. (1952); J. Inst. Metals **80**, 507.
MCLEAN, D. (1957); Grainboundaries in Metals (Clarendon Press, Oxford).
MOTT, N. F. (1948); Proc. phys. Soc. **60**, 391.
MOTT, N. F. (1953); Phil. Mag. **44**, 742.
MOTT, N. F. and F. R. N. NABARRO (1948); Rep. Conf. Strength of Solids, Bristol, p. 1 (The Physical Society, London).
NABARRO, F. R. N. (1948); Rep. Conf. Strength of Solids, Bristol, p. 50 (The Physical Society, London).
OROWAN, E. (1947); J. West Scotland Iron Steel Inst. **54**, 45.
SCHOECK, G. (1956); Creep and Recovery (ASM, Cleveland) p. 199.
SEEGER, A. (1954); Rep. Conf. Defects in Cryst. Solids, Bristol, p. 391 (The Physical Society, London).
SEEGER, A. (1957), in: Dislocations and Mechanical Properties of Crystals (Wiley, New York) p. 238.
SHERBY, O. D., J. L. LYTTON and J. E. DORN (1957); Acta Met. **5**, 219.
SULLY, A. H. (1956); Progress in Metal Physics **6** 135.
TAMHANKAR, R., J. PLATEAU and C. CRUSSARD (1958); Revue Métall. **55**, 383.
THORNTON, P. R. and P. B. HIRSCH (1958); Phil. Mag. **3**, 740.
WEERTMAN, J. (1955); J. appl. Phys. 5.
WEERTMAN, J. (1957); J. appl. Ph 5.
WEILL, MME (1952); Mém. Artillerie Royale **26**, 41.
WOOD, W. A. and W. A. RACHINGER 9); J. Inst. Metals **76**, 372.
WYATT, D. H. (1953); Proc. phys. So 4.
WYON, G. (1953); Métaux et Corrosi 43, 203 and 244.

CHAPTER XIII

SPECIAL TOPICS OF THE PLASTIC DEFORMATION OF METALS

1. Bauschinger Effect

Originally observed in polycrystals but later found to be present also in single crystals, the *Bauschinger effect* denotes a certain dependence of the flow stress and rate of work-hardening on the strain history of the metal: during plastic deformation the flow stress for strain in the *reverse* direction (for instance compression) is smaller than for strain in the *original* direction (for instance tension). The rate of work-hardening *in*creases temporarily after stress reversal, however. MASING and MAUKSCH [1923] tried to explain the effect on the basis of intergranular stresses (so-called Heyn stresses) set up in polycrystalline material during the test, as the result of different anisotropic deformations of neighbouring grains. WOOLLEY [1953] has re-estimated the magnitude of these Heyn stresses in actual cases, and found them far too small to explain the observed effects. SACHS and SHOJI [1927] were the first to prove the existence of a Bauschinger effect also in single crystals (of α-brass), and in view of this the explanation mentioned cannot be generally upheld. Bauschinger effects have been observed in single crystals of iron, zinc, cadmium and aluminium, and in many polycrystalline metals. An illustration of the effect is given in Fig. 1, which pertains to the axial straining of a pure aluminium crystal (BUCKLEY and ENTWISTLE [1956]). After *extension* to b, the specimen is unloaded to c. If the specimen is stressed again, in *tension*, the flow curve follows the fulldrawn line: plastic flow begins anew at b', differing only slightly from b (owing to elastic strains) and forms then a perfect continuation ($b' - d$) of the first part. However, if loaded in *compression*, plastic flow already starts at e and the flow curve is that indicated by the dotted line, which has a higher work-hardening coefficient. The critical flow stress for compression is thus strongly dimin-

ished; on reaching the flow stress for tension, σ_t, corresponding to the point b, the compression has already progressed to the point f. Buckley and Entwistle defined a quantity they call the Bauschinger strain,

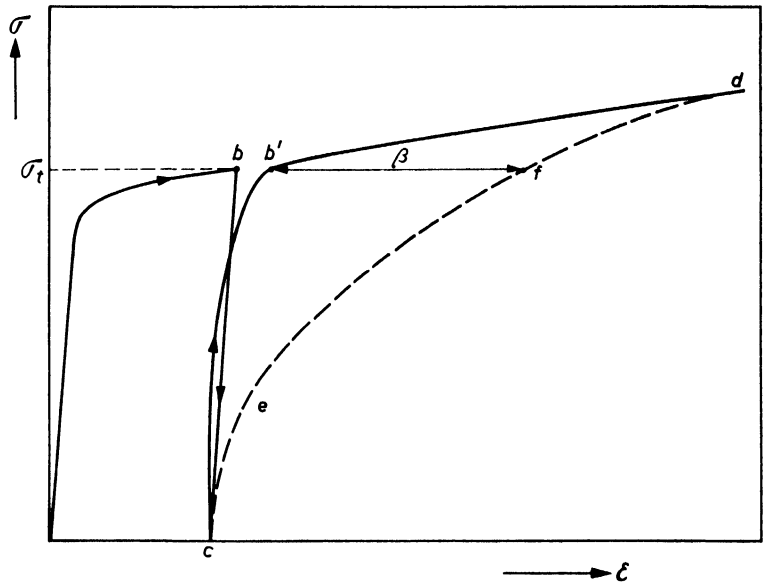


Fig. 1. Bauschinger effect in axially strained aluminium single crystals (after BUCKLEY and ENTWISTLE [1956]). The specimen is loaded in tension to b and then unloaded in c. Renewed loading in tension results in the full-drawn curve $cb'd$; the flow stress at b' differs only insignificantly from that at b. Renewed loading in *compression*, however, causes plastic flow already at e, and the flow curve follows the dotted line $cefd$. The flow stress is thus diminished but the work-hardening coefficient has temporarily increased. For the sake of presentation no distinction has been made between tensile and compressive stresses and strains in this diagram.

which is simply the difference between compressive and tensile strains at the stress level σ_t. They use this strain, denoted by β in the diagram, as a measure of the Bauschinger effect. Their results can be summarized as follows:

1) The Bauschinger strain of polycrystalline aluminium increases approximately proportionally to the pre-stress level σ_t.

2) In single crystals the Bauschinger strain is uniquely related to the total amount of work-hardening produced during the prestressing. Prestressing in the easy glide range results in relatively large Bauschinger strains which are proportional to the amount of work-hardening; prestressing in the second part of the stress-strain curve (see Chapter

VIII) produces but little additional Bauschinger strain (Fig. 2).

3) Surface polishing after prestressing does not effect the Bauschinger strain, and the Bauschinger effect is *not* correlated with the formation of slip markings on the surface of the crystals. During the

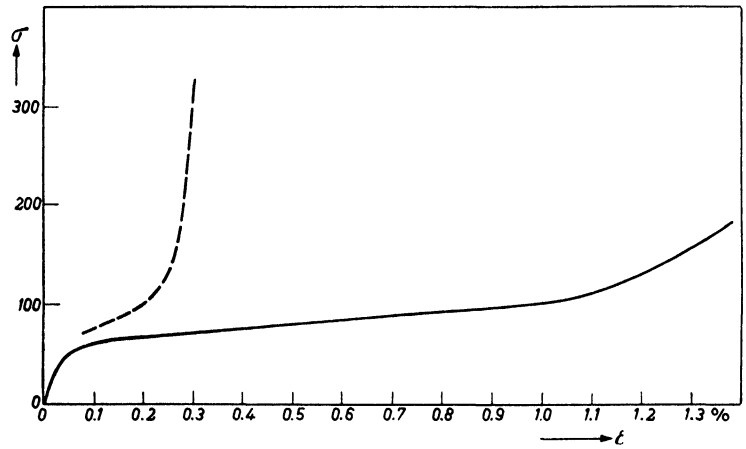


Fig. 2. Glide strain (drawn curve) and Bauschinger strain (dotted curve) as a function of the tensile stress (in g/mm²) for an aluminium single crystal. Both strains are approximately proportional to the flow stress (and thus to the work-hardening) in the easy glide region, which extends up to a stress level of about 100 g/mm² (1% glide strain). (After data by BUCKLEY and ENTWISTLE [1956].)

renewed straining *new* slip lines only appear after the stress level of approximately σ_t has been reached, and these do not differ from those which should have appeared after continued loading in the original sense.

4) The maximum Bauschinger strain observed amounts to about one tenth of the maximum prestrain; it is remarkably resistant against annealing treatment. Other experimenters have observed more or less the same characteristics (for instance THOMPSON *et al.* [1955]) and we shall not go into more detail here. A review of data can be found in a recent paper by THOMPSON and WADSWORTH [1958].

The ensemble of results of this and other work indicates clearly that the process responsible for the Bauschinger effect is the *reverse motion* of dislocations emitted by Frank-Read sources during the pretreatment. The easy glide region is characterized by the emission of a gradually increasing number of dislocation loops, which move a long distance and are finally stopped somewhere in the crystal. The con-

figuration thus obtained has an inherent stability, because on unloading and reloading no great difference in strength is observed (see below). On reversing the stress, however, a certain redistribution of the dislocations is apparently still possible and this results in the additional reverse deformation that is called the Bauschinger strain. Once multiple glide has started, the *newly* formed loops are practically immobilized from the start, owing to the formation of Cottrell-Lomer locks etc., and the Bauschinger strain does not appreciably increase further.

It is not completely clear why the stabilization occurs on unloading and why the Bauschinger effect is relatively so small in the easy glide region, as it is not obvious a priori why the dislocations emitted by a source should not nearly all return to their source on stress reversal. Probably some Lomer-Cottrell locking takes place already behind the loops during the early stages of strain, caused by incidentally activated dislocations on other slip systems.

EDWARDS and WASHBURN [1954] have investigated the Bauschinger effect in hexagonal zinc crystals deformed in shear at −196 °C. Again a strong Bauschinger strain was observed, similarly to that shown in Fig. 1. This is not surprising in view of the predominance of easy glide in such crystals. The Bauschinger strain was found to increase practically proportionally to the glide strain over a wide range, which is in agreement with this idea.

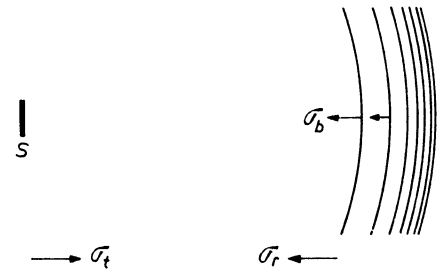


Fig. 3. Stresses acting on dislocation loops emitted by Frank-Read source S, after reversing the direction of the originally applied stress σ_t to σ_r.

The loops of a Frank-Read source, once having been emitted, will on the average, after removal of the applied stress σ_t, be acted upon by an oppositely directed stress σ_b exerted on them by the other dislocations in the piled-up group of which it forms a component. Only the "final" dislocations of such a group (Fig. 3) will move back under

this stress, but apparently the back stress is too small to produce an appreciable shrinking of these loops in the absence of an external stress; other strain-hardening effects, to be represented by a stress of order of magnitude σ_t prevent this. When, in addition, a reverse stress σ_r is applied which acts in the same sense as σ_b, plastic flow will start in the reverse direction at that value of σ_r such that

$$\sigma_r + \sigma_b \approx \sigma_t . \tag{1}$$

The Bauschinger strain contributed to by these loops is now in first approximation given by:

$$\Delta\varepsilon = \left(\frac{\partial\varepsilon}{\partial\sigma}\right)_{\sigma_t} \cdot \sigma_b = \frac{1}{h}\sigma_b , \tag{2}$$

where h is the coefficient of work-hardening, at the level σ_t. If the distribution function of dislocation segments acted upon by a back stress between σ_b and $\sigma_b + d\sigma_b$ is $N(\sigma_b)d\sigma_b$, the total Bauschinger strain is

$$\Delta\varepsilon_B = \frac{1}{h}\int N(\sigma_b)\sigma_b d\sigma_b . \tag{3}$$

Assuming σ_b to be independent of the strain, it is thus found that the Bauschinger strain in the easy glide region is proportional to the inverse strain-hardening coefficient times the total length of dislocations in the material.

In the easy glide region h is small and $\Delta\varepsilon_B$ is large, as observed; it further increases (approximately proportionally to the strain) as the dislocation density increases. In the second part of the flow curve $1/h$ becomes much smaller, and the new dislocations formed will contribute but slightly to the Bauschinger effect.

Indications exist that the presence of impurities in face-centred cubic crystals increases the ratio of proportionality between Bauschinger strain and integrated work-hardening during prestressing. This should be interpreted in terms of the formation of more and smaller clusters, which are more easily resolved on reversed stress. The same cause is operative in extending the range of easy glide in these materials.

"HAASEN-KELLY" EFFECT

Closely related to the Bauschinger effect in so far as it is also due to the stabilization of dislocation configurations on unloading, is the occurrence in various metals of a small *apparent yield point* on re-

loading in the *same* direction. The flow stress immediately after reloading is sometimes slightly higher than immediately before; the stress gradually decreases to the "normal" level on further straining. An extensive investigation of this effect has been carried out by HAASEN and KELLY [1957]. It has been shown to occur in single crystals of aluminium, nickel and copper, and the difference between observed yield stress and continuation of the flow stress was found to be proportional to the measure to which the loading stress had been relieved.

2. Size and Surface Effects

The influence of the *surface* of a metal crystal or a polycrystalline aggregate on its plastic properties has been studied extensively in the past, but the results of this work are not very clear, and a systematic theory of surface effects does not yet exist. A discussion of this intriguing subject was given by ANDRADE [1953], and HAASEN and LEIBFRIED [1955] devoted a section to it in their review paper. We shall not repeat this work here but only mention a few of the most important aspects, according to the second paper mentioned.

The critical shear stress and rate of work-hardening of metal crystals depend on the diameter of the test specimen; in zinc both quantities begin to increase when the diameter becomes smaller than a few tenths of a mm. The choice of the slip system in this metal depends on its orientation with regard to the smaller dimension of the sample (GILMAN [1953]). A theoretical explanation of this has only been attempted.

Andrade and coworkers have since long drawn attention to the large influence which *oxyde layers* or other *thin deposited layers* on the surface of a metal have on the mechanical properties. Such a layer gives rise to a very considerable increase of the critical shear stress, and its removal before or during the test (by polishing or etching) usually results in a strong increase of deformation. A remarkable "abnormal" after-effect in twisted and etched zinc wires has been shown to be caused by this phenomenon (BARRETT [1953a]). The explanation of the effect of oxyde layers etc. must undoubtedly be found in the obstruction of the paths of emerging dislocations, in such a way that pile-ups are formed immediately below the surface, which cause additional work-hardening and give rise to a spontaneous extra deformation on removal of the surface layer. The influence of point defects on surface effects of this sort has been recently considered by WILLIAMS and HAYFIELD [1957].

REHBINDER and coworkers [1941] found the creep rate of cadmium and gold to be increased by immersing the samples in paraffine oil. This *Rehbinder effect* must also be explained in terms of the decrease of the surface barrier for the emergence of dislocations, but in this case the decrease is due to the lowering of the surface tension of the metal with respect to its surface-active surroundings. Electrolytically polished wires do not show this effect.

3. Plastic Deformation of Polycrystalline Metals

3.1. MATHEMATICAL THEORY

Glide on a given slip system starts when the shear component of the applied stress in the direction of the Burgers vector reaches a certain critical value, the critical shear stress. When the applied stress is a uniaxial tension σ, this shear stress component is related to σ by the formula

$$\tau = \sigma \sin \zeta \cos \lambda , \tag{4}$$

where ζ is the angle between the tension axis and the normal to the glide plane, and λ the angle between this axis and the glide direction (Burgers vector). In a polycrystal this *orientation factor* varies from grain to grain, and to obtain the overall stress-strain relation for a polycrystal, some averaging treatment with respect to orientation is necessary over the flow curves of single crystals. The problem is further complicated by the simultaneous action of several distinct slip systems, with different orientation factors, in each crystal. This is a necessary condition, since in general glide on more than one independent slip system is required for each crystal to deform in coherence with its neighbouring grains and not to leave voids at its boundaries. The most general kind of deformation can be taken up by five independent glide systems (VON MISES [1928]; TAYLOR [1938]). This corresponds to the six components of the general strain tensor, reduced to five independent quantities because of the condition of invariance of volume.

In practice, less than five systems often suffice since *grainboundary slip* and *mechanical twinning* accomodate part of the forced deformation. We shall not consider these extra possibilities here; they play a very important role in metals with only a restricted number of slip systems, such as the hexagonal metals.

TAYLOR [1938] has studied the averaging problem, starting from the assumption that in each grain the choice between the possible slip

systems is made in such a way that the total amount of dissipated strain energy

$$\sum_{i=1}^{n} \tau_i \mathrm{d}\gamma_i ,$$

where τ_i and γ_i refer to the shear stress and increment of shear strain on the i^{th} system respectively, is a minimum. Under the additional assumption that τ is the same on all systems and equal to the resolved critical shear stress for activation of a Frank-Read source of average length, this condition reduces to $\Sigma\,|\mathrm{d}\gamma_i| =$ minimum. A formal proof of the validity of this assumption has been given by BISHOP and HILL [1951]. Further, the conditions of continuity of stress and strain on the boundaries between the grains must be satisfied, which conditions form a formidable set of requirements. For its solution only approximative methods come in question; Taylor derived the value $\bar{m} = 3.06$ as the *upper* limit for the ratio $\bar{m}$ between the flow stress of a polycrystal to the resolved critical shear stress on each glide system, or, in mathematical terms, for the physically averaged value of $m = (\sin\zeta \cos\lambda)^{-1}$ (this quantity varies between 2 and 3.7 in face-centred cubic crystals).

SACHS [1928] had already shown earlier that one should obtain $\bar{m} = 2.238$ (the weighted mean of all 12 single crystal m's) under the assumption that each grain deforms independently of its neighbours. This case, however, refers more exactly to a single crystal loaded in an arbitrary sense, than to a polycrystal. Nevertheless, the value 2.238 can be regarded as a *lower* limit to $\bar{m}$ for the latter. BISHOP and HILL [1951] have later reanalyzed Taylor's problem, and obtained essentially the same numerical results as this author. They stated that the principal conditions under which Taylor's approximation should be valid, are:

1) Only slip contributes to the deformation.
2) The critical shear stress and rate of work-hardening are the same on each active slip system; the latter remains constant during the deformation.
3) The condition of the least shear sum is not necessary but follows from two assumptions regarding the absence of correlation between strain, hardening, and orientation of a specific grain in the polycrystal.

Under these conditions the average value $\bar{m}$ should be independent of the amount of deformation.

The question now arises to *which single crystal* work-hardening curve should be applied to obtain the polycrystal curve, since we have seen in Chapter VI that the orientation of the crystal determines the shape of the first curve to a large extent. Taylor took as his basic stress-strain curve that of a crystal stressed along an axis oriented in the middle of the stereographic triangle. KOCKS [1958] advanced arguments to use the curve pertaining to a crystal stressed along ⟨ 100 ⟩ or ⟨ 111⟩, because such crystals deform in multiple slip, whereas the orientation used by Taylor is especially favourable to single slip, a phenomenon which does not occur in polycrystals.

Any experimental determination of the mean orientation factor depends on the comparison between the stress-strain curves of single and polycrystals of the material. Suppose one has for the relation between shear stress and shear strain on a given slip system

$$\tau = f(\gamma) \,. \tag{5}$$

Then, for a polycrystal consisting of a completely random aggregate of crystals, which moreover *remains* completely random during the deformation (no texture), the tensile stress $\sigma = \bar{m}\tau$, and the tensile strain

$$e = \frac{1}{\bar{m}} \gamma \,.$$

Thus

$$\sigma = \bar{m} f(\bar{m} e) \tag{6}$$

becomes the relation between tensile strain and stress of the aggregate. From such a comparison in aluminium TAYLOR and QUINNEY [1931] derived $\bar{m} = 3.1$. Kocks has extended this comparison, taking account of the division of the stress-strain curve into three parts, and could confirm the validity of (6), especially in part II, again with $\bar{m}$ near 3.1.

The stress-strain curve of polycrystals thus agrees remarkably well with that obtained by using Taylor's formula (6), when taking for $\bar{m}$ the value of about 3.1. Bishop and Hill have derived formulae connecting stress and strain theoretically for other deformations than simple extension. They have shown that Taylor's theory will remain valid even if inhomogeneous deformations are present. Several other conclusions of Taylor's theory have not been confirmed, however, such as the predicted occurrence of a certain deformation texture. Probably, other deformation modes play a deciding role here (see Chapter XX).



Fig. 4. Slip-line pattern in polycrystalline high-purity aluminium sample, deformed in tension by a few % (after BOAS and OGILVIE [1954]).

The problem of averaging over a polycrystalline aggregate arises in many other instances, e.g. in the computation of the elastic constants of a polycrystal, etc. It would lead us to far to discuss the particularities of several proposed averaging procedures here. A semi-empirical approach to the computation of the average elastic constants of a polycrystalline metal has been given by SMIT and VAN BUEREN [1954], and an exact treatment of the problem has been presented by KRÖNER [1958].

3.2. PHYSICAL ASPECTS OF POLYCRYSTALLINITY

From the physical point of view, the plastic deformation of polycrystals differs in several respects from that of single crystals. BOAS and OGILVIE [1954] and KOCKS [1958] have investigated by microscopic techniques plastically deformed polycrystalline high-purity aluminium and α-brass specimens. The observed slip-line patterns indicated that grains entirely surrounded by others deform *in*homogeneously in a tensile test. The active slip systems and the density of slip lines vary from one place to another in the same grain, the density being greatest near the boundaries. Occasionally the slip lines propagated from one grain into another (with a change of direction, of course, see also Chapter VI). The patterns on both sides of a boundary often showed great similarity (Fig. 4). The maximum number of independent sets of slip lines observed was six in brass; in aluminium this number was on the average three but sometimes increased to five. As in face-centred cubic crystals only four sets of lines are observable on one surface if the glide is octahedral, non-octahedral slip must have taken place in the latter case. The occurrence of non-octahedral slip in aluminium polycrystals seems indeed not uncommon, as shown by recent work of HARTMANN and MACHERAUCH [1958], who found not less than 40% of the slip lines belonging to non-octahedral systems.

Another well-known phenomenon, occurring especially clearly in polycrystals, is the easy formation of Lüders bands (see Chapter IX). This can be explained, apart from by the blocking of dislocations by impurities which gives rise to a yield point, as due to the impediment of the motion of dislocations by grain boundaries. An activated source can only send its dislocation loops as far as the nearest grain boundary; there the loops are usually stopped and form pile-ups (cf. FRIEDEL [1956]). The pile-ups produce large stresses in the neighbouring grain, and activate the sources there, helped in this perhaps by thermal

activation. Thus the deformation proceeds from grain to grain in jumps, that is, with a finite velocity. PETCH [1953] has observed that the yield stress of a polycrystal depends on the grainsize d according to:

$$\sigma_y = \sigma_i + kd^{-\frac{1}{2}} \tag{7}$$

and this can be easily explained by the pile-up concept.

FRIEDEL [1956] considers two stages of polycrystal work-hardening. *Intragranular* hardening can directly be compared with the second part of the stress-strain curve in single crystals. The work-hardening is largely determined by the elastic interaction between the dislocations in the *same* grain. This interaction will, however, now be particularly effective, since piled-up groups are formed against the boundaries which act as barriers. Mott's original theory of work-hardening (MOTT [1952, 1953]), for a discussion of which we refer the reader to the original papers, should now apply, and the stress-strain curve should be *parabolic*. In practice the stress-strain curve will probably be intermediate between the linear (stage II) and quadratic forms, and this has indeed been observed. This first stage is characterised by long straight slip lines extending fully through the grains (Fig. 5*a*).

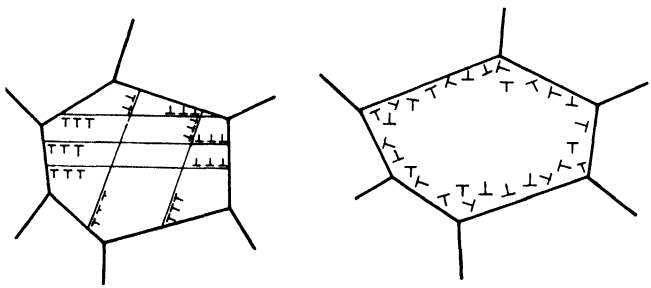


Fig. 5. Dislocation distribution in *intra*granular (*a*) and *inter*granular (*b*) hardening, according to FRIEDEL [1956] (schematic).

In metals with a sufficient number of possible glide systems the already mentioned *activation* of slip on other systems in neighbouring grains will gradually give rise to an overall dislocation distribution, in which the stress fields of the many dislocations on both sides of a boundary effectively cancel. In this second deformation stage, termed *intergranular* hardening, the effect of the pile-ups disappears (Fig. 5*b*) and the hardening decreases. Stress-strain curves consisting of an

initial strongly work-hardening parabolic part followed by less steep part, have indeed been observed by CRUSSARD and co-workers [1950, 1952, 1953]. The transition occurs together with the formation of many secondary slip lines in the boundary zones.

In crystals with only a few possible slip systems, the described "softening" process is not possible, and the range of normal plastic deformation is severely restricted by the occurrence of exceptionally severe intragranular work-hardening. This is indeed the case in hexagonal metals like zinc, which is only plastically deformable in its polycrystalline form at fairly elevated temperatures (above room temperature). At low temperatures polycrystalline zinc is virtually *brittle*, deformation is there possible only by the intermediate of twinning.

4. Deformation by Mechanical Twinning

We have already indicated in Part I that crystals can acquire a permanent deformation by *mechanical twinning*, a dislocation process that is to be distinguished from slip.

A twin in a crystal is a fault that can be obtained by mutually reflecting with regard to a given *translation plane*, the lattice structure on either side of a certain *boundary plane*, as has been indicated in Fig. 6. The crystallographic structure obtained by the reflection is the

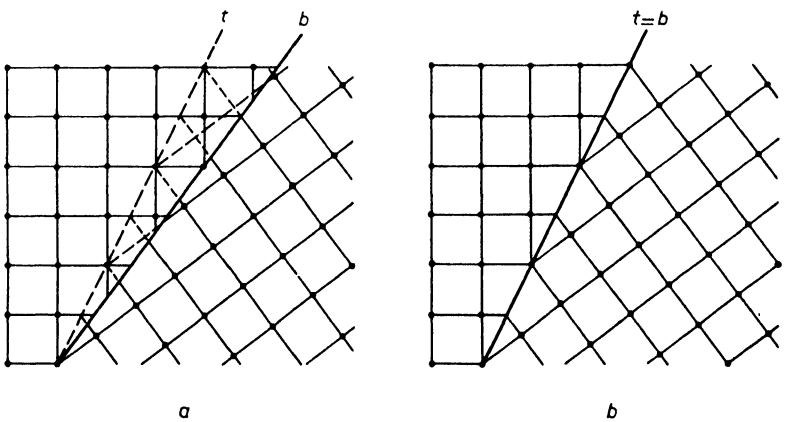


Fig 6. Twin in a square lattice. Translation plane t is in both diagrams a (112) plane. The boundary b in (a) is a noncoherent, in (b) a coherent twin boundary.

same as the original structure. This is a necessary consequence of the reflection; if it is not fulfilled, but a transformation occurs which is

(partly) a homogeneous shear without this property, one speaks of a *martensitic* transformation (see Chapter XXI).

A twin can be obtained by translating part of the crystal with respect to the other by a vector parallel to the translation plane, the magnitude of which *increases proportionally* to the distance of this plane (homogeneous shear). The amount of translation *per* lattice plane is called the twin vector or translation vector. The translation plane can be a (low-index) crystallographic plane that is the plane of symmetry between two twinned configurations, in which case one speaks of a twin of the first kind (Fig. 6). It can also be a more arbitrary (but well-defined) plane, in which case the twin is of the second kind. The *order* of a twin pertains to the number of reflections necessary to obtain it. We shall only regard first order twins.

The boundary between the twinned regions need not necessarily be the translation plane nor even a lattice plane. If a twin of the first kind is bounded by the lattice plane that is its own translation plane, one speaks of a *coherent* twin boundary (Fig. 6*b*). This case often occurs in practice, as such a boundary generally represents the least amount of interfacial energy (the lattice planes *fit* everywhere on both sides along the boundary).

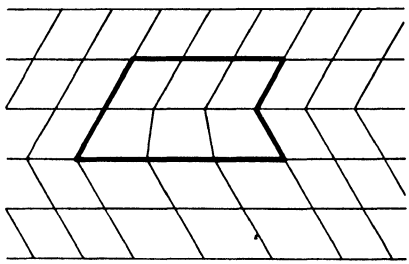


Fig. 7. Structure of a twin dislocation (after FRIEDEL [1956]).

A twin boundary (we shall in the following, for the sake of simplicity, consider a coherent boundary) need not be confined to one lattice plane. It can pass from one plane to a neighbouring one. In this case the translation plane is bounded by a *twinning dislocation* (Fig. 7). In view of the necessary geometrical relation between the Burgers vector of this dislocation and the primitive lattice vectors, it can easily be shown that a twinning dislocation of this kind is in fact a *partial* dislocation (the twinning translation is always smaller than the slip translation).

If Fig. 7 is regarded as referring to a face-centred cubic lattice, the

twinning dislocation depicted is in fact a Shockley half dislocation (Chapter III: the twinning replaces the cubic stacking by the hexagonal over a width of two octahedral planes). In correspondence herewith, a twin in the face-centred cubic matrix can be regarded as a continuous series of certain stacking faults on neighbouring lattice planes. It is therefore not surprising that the *generation* of a twin in the lattice is directly related to the motion of partial dislocations (COTTRELL and BILBY [1951]). To study the twinning on the octahedral plane in face-centred cubic crystals, we must consider the properties of Shockley partials. Let us consider a Shockley stacking fault terminating somewhere on such a partial dislocation with Burgers vector b (Fig. 8).

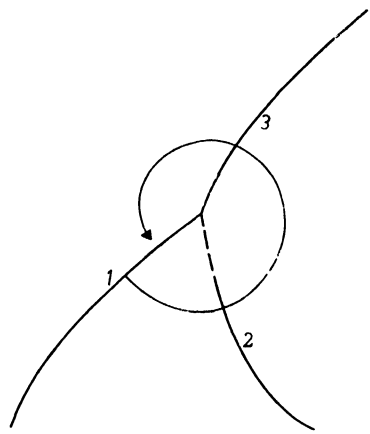


Fig. 8. Cottrell-Bilby pole mechanism for generating a twin from a rotating partial dislocation.

Where the partial (1) ends in a node with two other dislocations (2 and 3) the vector sum of the Burgers vectors of these three dislocations must be zero. In general the two dislocations 2 and 3 each have a Burgers vector with a component perpendicular to the plane of the stacking fault (the two components are of opposite sign of course). If now the partial dislocation 1 rotates around the node in the sense indicated by the curved arrow, it is displaced normal to the rotation plane by an amount equal to the sum of the normal component of the Burgers vector of dislocation 2 and the *negative* of the normal component of the Burgers vector of the dislocation 3 (this is because the signs of the Burgers vectors are defined according to the *same* sense of revolution of a Burgers circuit, section III–1). When this sum is equal

to a *lattice vector*, the partial dislocation sweeps successively over neighbouring lattice planes, and translates the atoms in these in such a way to form a translation twin at the cost of the original lattice. For this *pole* mechanism to be feasible, the three dislocations constituting the node must obey certain geometrical conditions, which indeed can be fulfilled in these metals. The mechanism is completely general and can give rise to coherent as well as to incoherent twins. It can be compared directly with the action of Frank-Read sources, which operate in one plane. In fact, BILBY [1954], H. SUZUKI [1954] and others indicated the possibility of Frank-Read sources operating in successive planes, by a mechanism similar to that discussed above. Such sources do not emit a succession of dislocation loops but a dislocation spiral. The only distinction from the twinning source is now that in the former case the dislocations are perfect; in the twinning case they are partial dislocations. SUZUKI and BARRETT [1958] have discussed a specific "twinning mill" that should be operative in Ag-Au alloys.

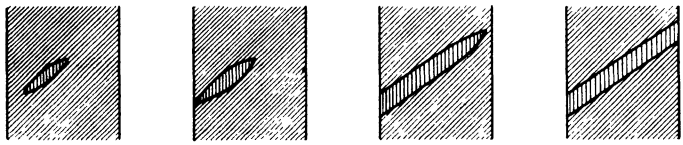


Fig. 9. Growth of a twin (schematic).

As for a Frank-Read source, there exists a certain critical shear stress for a twinning mill to operate. This has already been computed by Cottrell and Bilby; it depends, apart from on the line tension of the dislocation loop, also on the surface tension of the twin boundary and usually exceeds therefore the critical shear stress for slip. Therefore, when glide is *possible*, twinning will in general not occur at a large scale. Only in those cases where the deformation cannot be taken up by slip alone (polycrystals, hexagonal crystals arbitrarily deformed), twinning will take place, as has indeed been observed. The nucleation of twins by stress and the dynamics of twinning have been investigated by a number of authors (OROWAN [1954], BELL and CAHN [1957]). A well-defined critical resolved shear stress as exists for slip is absent for twinning, but a high stress level is necessary. Bell and Cahn concluded that, at least in zinc crystals, some slip is necessary to introduce the nucleation of a twin, presumably because dislocation pile-ups must form to act as stress concentration centres. Once formed, the nucleus

extends to a lens-shaped twin that straightens out to a twin lamella when it reaches the boundaries of the crystal (Fig. 9). The lens is necessarily bounded by a number of twinning dislocations in the form of expanding loops. The expansion also requires a certain minimum stress which is, however, much smaller than that needed to nucleate the twin. This situation is somewhat comparable to that of yielding in body-centred cubic alloys: once a nucleus is formed the stress drops to a considerable lower value. Avalanches (twinning bursts) are a common occurrence, and are sometimes accompanied by an audible clicking sound, the so-called twinning cry (e.g. in tin). The time of formation of a twin lamella is essentially governed by the velocity of the dislocations responsible for the twinning. One must actually assume that these *twinning* dislocations propagate with at least one tenth of the velocity of sound, in order to explain the observed rapidity of the twinning. This is in disaccord with the general impression about the velocity of *slip* dislocations; according to Eshelby's and Leibfried's computations, mentioned in Part I, the damping associated with ordinary dislocation motion would restrict their velocity to less than 0.1 of that of sound (see also section III–5). It is not impossible, however, that twinning dislocations, by their special geometry, are damped less than ordinary dislocations.

Twinning and the operation of Frank-Read sources are closely related to the dislocation mechanism of crystal growth. This will be discussed in Chapter XXIII.

For a description of the crystallographic and metallographic aspects of twinning, we refer to the books of HALL [1954] and BARRETT [1953b], in which a list of twinning elements in various metals can also be found, and to the extended review paper by CAHN [1954].

5. Deformation by Bending and Kinking

During a deformation test often stress concentrations occur near the grips or near other places. These can give rise to peculiar local formations of slip lines and shapings of the specimen, especially of hexagonal crystals where basal glide takes place almost exclusively.

When during or after a bending test of a metal showing single glide the temperature is high enough, polygonization will occur, resulting in a decrease of the dislocation density to the value given by formula XI–1, and to a distribution of these dislocations, all of the same sign,

as pictured in Fig. XI–3*b*. The dislocations become aligned in rows perpendicular to the glide planes, which gives rise to special deformation patterns (kinks). These have been observed among others by GILMAN [1955] in zinc crystals, deformed in bending along a basal direction (Figs. 10*a* and 11). Stress inhomogeneities in a *compression* test may also easily result in local bending, and thus in similar polygonization to socalled *bend planes*, as for instance depicted in Fig. 10*a*.

A situation such as that of Fig. 10*a* cannot form at low temperatures. Under the circumstances mentioned (compression with glide planes nearly parallel to the stress axis), the dislocations associate rather in the form of *condensed groups*, as illustrated in Fig. 10*b*. The result is the formation of a so-called *kink*, a region of high local curvature, which gives the deformed specimen a particular shape (Fig. 11). These kinks occur preponderantly near the grips and corners of the deformed specimens, in accord with the general trend discussed above.

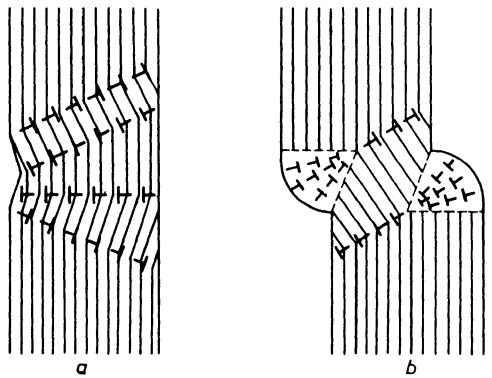


Fig. 10. *a*) Bend planes in compressed hexagonal crystal, as found near an inhomogeneity of the deformation at elevated temperatures; *b*) The configuration *a*) cannot form at low temperatures where dislocation climb is impossible, and a kink is formed instead.

Various other surface and glide-band features have been found in many hexagonal crystals, loaded in compression or otherwise. These have been named differently by different authors, but they can all be explained in terms of rotation and curvature of the glide planes, sometimes followed by polygonization, sometimes not.

There exists a great analogy between the kinks, formed in hexagonal metals, and the various types of deformation bands (usually also called kink bands) found on the surface of face-centred cubic metals such as

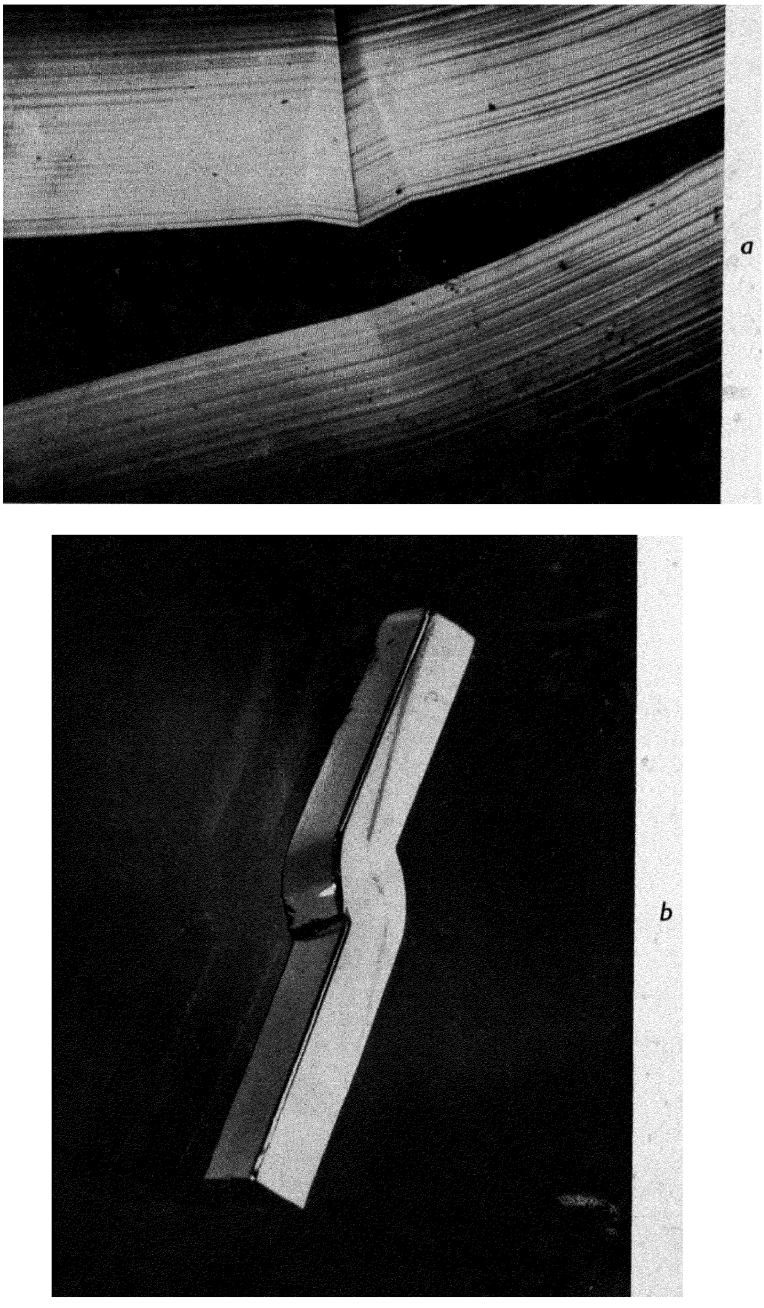


Fig. 11. *a*) Bend planes formed at elevated temperature in compressed zinc crystal; *b*) Kinks in monocrystalline cadmium bars, compressed along a basal plane direction (GILMAN [1950]).

aluminium, which have been discussed in Chapters VI, VII and VIII. As explained there, the latter are caused by the mutual hindrance, during easy glide, of the motion of dislocation of opposite sign on neighbouring glide planes (Fig. 12). Whether deformation bands in face-centred cubic metals are necessarily connected with local stress inhomogeneities, like kinks in zinc and cadmium, is not certain; the former can perhaps better be regarded as an essential result of the slip process in the cubic crystals, which leads to stress inhomogeneities on a microscopic scale. A theory of kinking has been given by FRANK and STROH [1952].

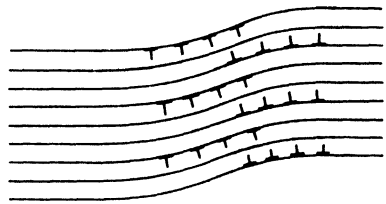


Fig. 12. Deformation band in face-centred cubic crystal.

In the face-centred cubic crystals the deformation bands contribute essentially to the work-hardening (see Chapters VI to VIII). The work-hardening of hexagonal crystals is probably not determined directly by the presence of kinks or bend planes. These should more be considered as a form of *macroscopic* inhomogeneities in the deformation pattern.

REFERENCES

ANDRADE, E. N. DA C. (1953); Properties of Metallic Surfaces (Inst. of Metals Report No. **13**, London).
BARRETT, C. S. (1953a); Acta Metall. **1**, 2.
BARRETT, C. S. (1953b); The structure of metals (McGraw-Hill, New York).
BELL, R. L. and R. W. CAHN (1957); Proc. Roy. Soc. **A239**, 494.
BILBY, B. A. (1954); Rep. Conf. Defects in Cryst. Solids, Bristol, p. 124 (The Physical Society, London).
BISHOP, J. F. W. and R. HILL (1951); Phil. Mag. **42**, 414 and 1298.
BOAS, W. and J. OGILVIE (1954); Acta Metall. **2**, 655.
BUCKLEY, S. N. and K. M. ENTWISTLE (1956); Acta Metall. **4**, 352.
CAHN, R. W. (1954); Advances in Physics **3**, 363.
COTTRELL, A. H. and B. A. BILBY (1951); Phil. Mag. **42**, 573.
CRUSSARD, C. (1953); Rev. Métall. **50**, 697.
CRUSSARD, C. and B. JAOUL (1950); Rev. Métall. **47**, 589.
CRUSSARD, C. and B. JAOUL (1952); C. R. Acad. Paris **234**, 700.

EDWARDS, E. H. and J. WASHBURN, (1954); J. Metals **6,** 1239.
FRANK, F. C. and A. N. STROH (1952); Proc. phys. Soc. **B65,** 811.
FRIEDEL, J. (1956); Les Dislocations (Gauthier-Villars, Paris).
GILMAN, J. J. (1953); J. Metals **5,** 1217.
GILMAN, J. J. (1955); Acta Metall. **3,** 277.
HAASEN, P. and G. LEIBFRIED (1955); Fortschritte der Physik **2,** 79.
HAASEN, P. and A. KELLY (1957); Acta Metall. **5,** 192.
HALL, E. O. (1954); Twinning (Butterworth Scient. Publ., London).
HART, E. W. (1955); Acta Metall. **3,** 143.
HARTMANN, R. J. and E. MACHERAUCH (1958); Z. Naturwiss. **45,** 540.
KOCKS, U. F. (1958); Acta Metall. **6,** 65.
KRÖNER, E. (1958); Z. Physik **151,** 504.
MASING, G. and W. MAUKSCH (1923); Wiss. Veroff. Siemens **4,** 74.
MISES, VON (1928); Z. Angew. Math. Mech.
MOTT, N. F. (1952); Phil. Mag. **43,** 1151.
MOTT, N. F. (1953); Phil. Mag. **44,** 742.
OROWAN, E. (1954); Dislocations in Metals (New York, AIME) p. 116.
PETCH, N. J. (1953); J. Iron Steel Inst. **173,** 25.
REHBINDER, P., V. I. LICHTMANN and V. M. MASLENIKOW (1941); C. R. Acad. Sci. U.R.S.S. **32,** 125.
SACHS, G. (1928); Z. Ver. Deutscher Ing. **72,** 734.
SACHS, G. and H. SHOJI (1927); Z. Physik **45,** 776.
SMIT, J. and H. G. VAN BUEREN (1954); Philips Research Reports **9,** 460.
SUZUKI, H. (1954); J. Phys. Soc. Japan **9,** 531.
SUZUKI, H. and C. S. BARRETT (1958); Acta Metall. **6,** 156.
TAYLOR, G. I. (1938); J. Inst. Metals **62,** 307. See also: Deformation and Flow of Solids (Springer, Berlin) p. **1.**
TAYLOR, G. I. and Y. H. QUINNEY (1931); Phil. Trans. Roy. Soc. **A230,** 323.
THOMPSON, N., C. K. COOGAN and J. G. RYDER (1955); J. Inst. Metals **84,** 73.
THOMPSON, N. and N. J. WADSWORTH (1958); Adv. in Physics **7,** 72.
WILMAN, H. (1951); Proc. Phys. Soc. **A64,** 329.
WILLIAMS, E. C. and P. C. S. HAYFIELD (1957); Vacancies and other Point Defects in Metals (The Institute of Metals, London).
WOOLLEY, R. W. (1953); Phil. Mag. **44,** 597.

CHAPTER XIV

NON-THERMAL GENERATION OF POINT DEFECTS

1. Generation of Defects by Plastic Flow

Plastic deformation results in the generation of vacancies, interstitial atoms and probably more complicated point defects also, as has already been explained in Chapter VII. There it was estimated that, in order of magnitude, the concentration of point defects after a plastic strain ε is given by

$$F \approx A\varepsilon^p , \tag{1}$$

where p is an exponent between 1 and 2 and A varies correspondingly between 10^{19} and 10^{21} cm^{-3}, depending on the model used. It is then assumed that all defects generated remain as such in the metal, an assumption which is probably true at low temperatures at which no diffusion can take place. After 10% deformation, the defect concentrations to be expected in a typical metal are of the order 10^{18}–10^{19} cm^{-3}.

SEEGER [1954] has pointed out that, according to the temperature of deformation, differences may exist in the mechanism of defect production, according to the relative importance of thermal activation. Seeger estimated the transition temperature above which this becomes of influence near 200 °K in copper.

The details of the mechanism(s) of generation of point defects during plastic flow are not yet completely clear. Undoubtedly a dependence on stress and strain rate exists, but even the exact strain dependence has never been carefully determined. This is due primarily to the difficulty of interpreting observable data such as resistivity, X-ray pattern and density, in terms of the concentration of point defects.

2. Generation of Point Defects by Irradiation

2.1. NATURE OF IRRADIATION

As has also already been mentioned before (Chapter II), irradiation with fast particles (high-energy neutrons, protons, electrons, etc.) or with energetic electromagnetic radiation (γ-rays) usually produces Frenkel defects in metallic crystals.

For each *displacement* a certain minimum amount of energy needs to be transferred from the impinging particle to an atom of the crystal. This *threshold energy* E_T has been estimated be several authors to be of the order of 25 eV in a large diversity of solid materials.

It is much larger than the energy of formation U_F of a single Frenkel defect (a few eV), because the potential barrier that must be overcome to create a separated interstitial-vacancy pair in a perfect lattice is much higher than the difference in free energy between an atom on a normal site and that on an interstitial site (Fig. 1). More or less detailed calculations have been carried out by DIENES [1952] and by HUNTINGTON [1957].

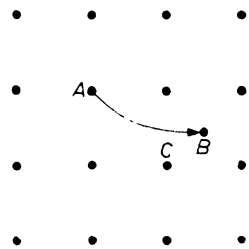


Fig. 1. The threshold energy for the production of a well-separated Frenkel defect by momentum transfer is essentially the energy difference between the situations in which the atom A is on the top of the potential barrier at C and on its own site A, respectively. The energy of formation governing the thermal equilibrium concentration of defects is the smaller difference in free energy between the positions B and A.

2.2. DISPLACEMENT THEORY

A number of reviews of the theory of the displacement of atoms by incident particles exist (KINCHIN and PEASE [1955a], SEITZ and KOEHLER [1956], DIENES and VINEYARD [1957]) and we can therefore be very brief. Only the effects of *elastic* collisions of the impinging particles with the atoms or ions of a crystal will be considered, and the influence of nuclear transformations will not be dealt with. Neither

shall we discuss changes introduced in the electronic structure of the solid.

2.2.1. *Neutrons*

Owing to their electrical neutrality, neutrons produce only disturbances in a crystal lattice when they make a head-on collision with a lattice atom (elastic collision). The latter (*primary knock-on*) then leaves its site, and subsequently produces many other displaced atoms as well as a *thermal* spike (local hot region) or a *displacement* spike (small region of completely disordered lattice). The energy transferred in an elastic collision by a (non-relativistic) neutron to a nucleus of mass number M is, according to (7), section II–4.1, approximately

$$T = \frac{4M}{(1+M)^2} E \sin^2(\zeta/2) \approx \frac{4E}{M} \sin^2(\zeta/2), \tag{2}$$

where E is the energy of the neutron and ζ the angle of deflection in the centre-of-mass system. Some values of the maximum transferable energy T_m are (for a 2 MeV neutron): H, 2.0 MeV; C, 0.56 MeV; Al, 0.28 MeV; Cu, 0.12 MeV; U, 0.033 MeV. The mean energy transferred is half this maximum energy. This results from the fact that the differential elastic cross-section σ (for transfer of energy between T to $T + dT$) is *in*dependent of T:

$$d\sigma(T) = \frac{\sigma_t(E)}{T_m} dT \tag{3}$$

(all recoil energies equally probable). $\sigma_t(E)$ is the *total* collision cross-section of a neutron with energy E which lies between 1 and 10 barns (increasing with increasing M; 1 barn $= 10^{-24}$ cm^2). This total collision cross-section for energy absorption is obviously related to the mean free path L of an incident neutron as:

$$L = \frac{1}{n_0 \sigma_t}, \tag{4}$$

where n_0 is the atomic density. L is generally found of the order of 1 cm for neutrons of about 1 MeV energy, so the chance that a neutron hits a nucleus is but small.

However, the primary knock-on itself carries so much energy and momentum that it produces a great number of *secondary* displacements

($T \gg E_T$). HARRISON and SEITZ [1955] computed the number of these approximately as:

$$\begin{aligned} s &= 0.561\,(1 + T/E_T)\,, \qquad T \gg E_T\,; \\ s &= 0\,, \qquad\qquad\qquad\quad T \leqslant 2E_T\,. \end{aligned} \tag{5}$$

The primary displaced atoms themselves are distributed in energy according to a distribution function that can be written as

$$n(T)\mathrm{d}T = \frac{a(E)}{(1 + T/E_T)^2}\,\mathrm{d}T\,. \tag{6}$$

From (6), (3) and (5) the average total number of displacements per primary knock-on can be computed as a function of the energy E of the neutron. This average number is found to be:

$$\bar{\nu}(E) = \int_{E_T}^{T_m} \sigma(T, E) n(T, E) s(T)\,\frac{\mathrm{d}T}{T_m}\,.$$

Evaluation of this integral leads to the approximate formula:

$$\bar{\nu}(E) \approx \frac{E}{4T_m}\,\{0.108 - 0.561 \log R\}\,(1 + R)^{-1}\,, \tag{7}$$

where

$$R = \frac{(1 + M)^2}{4M}\,\frac{E_T}{E}\,. \tag{8}$$

This amounts to about 400 in the case of Cu bombarded by 1 MeV neutrons. The process of primary and secondary displacements has been depicted in Fig. 2*a*. The *cascade* ends in atoms having received less than $2\,E_T$ energy, which cannot produce any more displacements. These are spaced only one or two atomic distances apart.

Let $Q(E)\mathrm{d}E$ be the partial flux (number of neutrons incident on unit area per unit interval of time) of neutrons of energy between E and $E + \mathrm{d}E$. Then after irradiation during a time t the number of atoms displaced per cm³ is

$$n = n_0 t \int_0^\infty Q(E) W(E)\mathrm{d}E\,, \tag{9}$$

where $W(E)\mathrm{d}E$ is the total probability that a *given* atom be displaced. For a typical case of irradiation (in the Oak Ridge reactor; flux 7.4×10^{11} neutrons/cm²/sec, of which 23% have energies larger than

25 keV) Seitz and Koehler computed $n = 6.3 \times 10^{18}$ cm^{-3}/week for copper (E_T was taken to be 25 eV).

GIBSON *et al.* [1960], by assuming some approximate model for the interatomic forces in metallic copper, solved the many-body collision problem numerically on an electronic computer. Fig. 2*b* gives an example of the results. Evidence for focusing collisions, displacement spikes, etc. (see below) has been clearly obtained in this way.

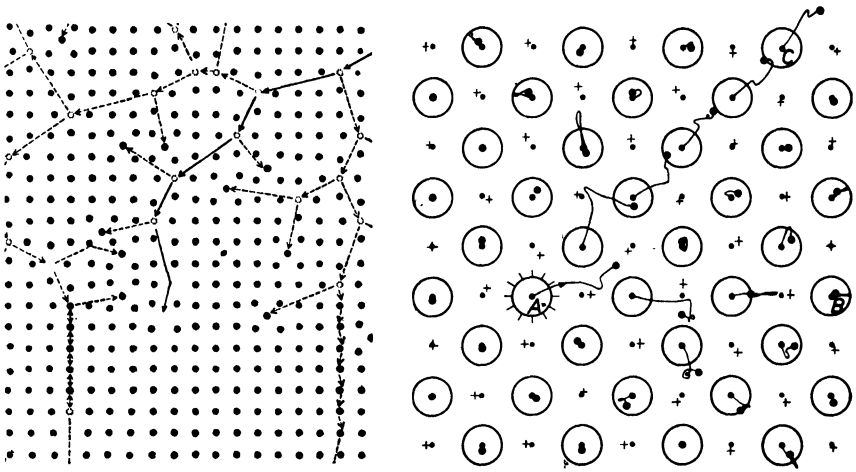


Fig. 2. *a*) Damage caused by a fast neutron incident upon a metal crystal (schematical). A *primary knock-on* (upper right corner) describes the path indicated by the full-drawn line, and causes in its way *secondary* and higher order *displacements*. Veritable *cascades* of *vacancies* (dotted circles) are formed, most branches ending finally in *interstitial atoms*. Focusing and defocusing collisions are indicated, whereas some atoms have received a large amount of extra vibrational energy (thermal spikes). (After BILLINGTON [1959]). *b*) Atomic orbits produced by incident particle transferring an energy of 40 eV to the atom at A in a direction 22½° above the horizontal axis. Large circles give the initial positions of the atoms in the plane of drawing, small dots of those in the plane below. Large dots give the final positions of the former, crosses those of the latter. A vacancy is left at A, an interstitial will probably be formed at C. (After GIBSON *et al.* [1960]).

2.2.2. *Protons, deuterons and other charged nucleons*

The passage of *charged* particles through matter poses a different problem, as the interaction between them and the lattice ions is primarily of an electric nature, viz. Coulomb interaction, and elastic collisions of the nature discussed above are unimportant. Moreover, the charged particles can spend part of their energy in excitation of the electrons bound to the lattice atoms or of the conduction electrons in

a metal. The theory of the Coulomb interaction between charged incident particles and nuclei has been given by N. BOHR [1913, 1948]. It is represented in detail in the paper by SEITZ and KOEHLER [1956] and we shall again quote only the most important results. The energy that is transferred by an incident particle of energy E in an elastic collision is, of course, again given by

$$T = \frac{4M_1M_2}{(M_1 + M_2)^2} E \sin^2 \tfrac{1}{2}\zeta , \tag{10}$$

which is the same formula as (2), only there M_1, the mass of the incident particle, was set equal to unity and M_2, the mass of the lattice nucleon, was denoted simply by M. The effect of the charge enters in the scattering cross-section σ. For non-relativistic charged particles in first approximation the well-known Rutherford formula holds for the differential cross-section:

$$\mathrm{d}\sigma = \frac{\pi a^2}{4} \frac{T_\mathrm{m}}{T^2} \mathrm{d}T ; \tag{11}$$

where T_m is again the maximum value of T, and

$$a = \frac{M_1 + M_2}{M_2} \frac{e_1e_2}{E^2} = \frac{Z_1Z_2e^2}{E^2} \frac{M_1 + M_2}{M_2} ; \tag{12}$$

$e_1 = Z_1e$ and $e_2 = Z_2e$ denote the charges of incident particle and lattice ion respectively. Screening of the Coulomb field of the ion in the solid material modifies this expression. Taking account of the screening by using an approximate formula for the electrostatic potential of the ion:

$$V = \frac{e_1e_2}{r} \exp(-r/b) , \tag{13}$$

(b can be evaluated by the Thomas-Fermi method), formula (11) transforms into:

$$\mathrm{d}\sigma = \frac{\pi a^2}{4} \frac{T_\mathrm{m}}{T^2} \left[1 + \frac{\lambda}{4\pi b} \frac{T}{T_\mathrm{m}}\right]^{-2} \mathrm{d}T \tag{14}$$

for $a \gg b$ (*weak* screening), where λ is the de Broglie wavelength of the incident particle in the centre-of-mass system. For *heavy* screening the Rutherford method must be replaced by an appropriate quantum mechanical approximation, such as the Born approximation; the scattering is isotropic when $\lambda/2\pi$ becomes much larger than the screening constant b.

Rutherford scattering obviously differs from elastic (hard-sphere) scattering in that, whereas in the latter case all energy transfers below T_m are equally probable, now a preference exists for *small* energy transfers. Collisions in which the energy transferred is of the order of $E_T = 25$ eV or larger, thus which can produce displacements, can all be described in practice by the classical Rutherford formula (14). The total cross-section for a collision in which the energy transferred lies between E_T and T_m is then simply

$$\sigma_t(T_m, E_T) = \frac{\pi a^2}{4}\left(\frac{T_m}{E_T} - 1\right). \tag{15}$$

With (12), we have for the total *Coulomb* collision cross-section for *displacement by non-relativistic particles*

$$\sigma_{displ} = \pi \frac{M_1}{M_2} \frac{Z_1^2 Z_2^2 e^4}{E\, E_T}. \tag{16}$$

This is usually about 10^{-16} cm^2 or less.

As M_2 for an electron is many times smaller than for a nucleus, the cross-section for absorption collisions with *electrons*, if present in the material, is of the order of 1000 times larger than for nuclei. The mean energy transmitted from the particle to the atom or electron is small:

$$\bar{T} = E_T \log \frac{T_m}{E_T},$$

which is usually only a few hundreds of electronvolts, in strong contrast to the neutron case. However, the cross-section is much larger. In connection with the foregoing, it is clear that a fast charged particle loses much more energy to electrons than to atoms (see below). The low value of the mean transmitted energy results further in a much smaller value for the number of secondary displacements, caused by a primary knock-on, than in the case of neutrons. This number amounts only to 5 or 6 (instead of 400) for a typical case of (10 MeV) charged particle irradiation. The mean free path for Coulomb collisions is again of the order of $(n_0\sigma)^{-1}$ which is a few atomic distances only, about 10^{-7} times that of neutrons, and the number of direct hits is correspondingly much higher.

Below an energy of about 20 keV in materials with $Z = 10$, or about 50 keV in Cu or Al (and still higher values in materials of larger atomic

number) the moving particle no longer suffers Rutherford collisions with the ions, but isotropic elastic "hard sphere" collisions as long as $E \gg E_T$. The cross-section then increases, and the kinetic energy of the moving ion becomes randomly partioned among the incident and knocked-on particles. This applies also to the action of the (charged) primary displaced atoms in making secondary displacements. Formula (5) was derived in this way.

Charged particles incident on a solid with a very high (*relativistic*) energy (0.1 to 1 GeV) behave much like electrons of a few MeV. We will discuss this *relativistic Coulomb scattering* in section 2.2.3 dealing with electron collisions.

The rate of energy dissipation per unit path length in displacement collisions follows from the above considerations as:

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\mathrm{displ}} = -n_0 \int_{E_T}^{T_m} T \mathrm{d}\sigma =$$

$$= -\pi n_0 \frac{M_1 Z_1^2 Z_2^2 e^4}{M_2 E} \log \frac{T_m}{E_T} \tag{18}$$

in the case of weak screening, where (14) holds. If collisions are also taken into account in which no displacement occurs but energy is only dissipated in the form of heat, E_T must be replaced in this formula by a smaller limit, of the order of 0.1 eV.

As has already been mentioned, fast charged particles passing through matter lose a very appreciable part of their kinetic energy in *exciting electrons*.

Electron excitation (a quantum mechanical process) is rather difficult to treat exactly. In insulators, it will presumably be of little importance, when the band gap is much wider than $\varepsilon = (m/M)E$, where m is the electron mass and M the mass of the incident particle; ε is the energy of an electron having the same *velocity* as that of the particle. [One should compare velocities rather than energies, as here quantum-mechanical transitions are concerned, characterized by *frequencies* (SEITZ and KOEHLER [1956]).] This formula leads to the simple rule that electron excitation is unimportant whenever the energy of the particle expressed in keV is less than its atomic weight (DIENES and VINEYARD [1957]).

From this it follows that *primary displaced atoms* cannot produce electron excitation in insulators or semiconductors, since ε for these

atoms is of the order of 0.1 eV at the utmost, and bandgap energies are much larger, from 0.5 to 2 eV. Thus in these solids the only process in which energy is lost by the displaced atoms is by elastic collisions and formula (5) should be valid. In metals the situation is less simple. However, it can be shown that the length over which a primary displaced atom of mass M_2 must move in order to lose an appreciable fraction of its energy by electron excitation is of the order of M_2/m atomic diameters. The corresponding length for Coulomb collisions is only a few tens of atomic spacings. Also in metals, therefore, the displaced atoms do not produce appreciable electron excitation.

For fast particles the rate of energy dissipation due to electron excitation is given by

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{elect}} = -2\pi n_0 \frac{M_1 Z_1^2 e^4}{mE} Z_2' \log \frac{2\varepsilon}{E}\,; \tag{19}$$

for slow particles it is negligible (see above). Here Z_2' is the number of bound electrons per atom which can be excited, that is, of which the mean excitation energy $\bar{E}$ does not exceed ε. For metals, equation (19) can thus be written as

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{elect}} = -2\pi \frac{Z_1^2 e^4}{\varepsilon} z \log \frac{\varepsilon}{E_F}, \tag{20}$$

where z is the number of conduction electrons per atom, and E_F is the Fermi energy.

Of course, ε and E in formulae (19) and (20) refer to the instantaneous values after the distance x has been travelled by the incident particle, and not to the values at the surface of incidence.

In all solids, the rate of energy dissipation by electron excitation of fast charged particles in the MeV range is computed from formulae (19) and (20) to be about 1000 times as large as that due to Coulomb collisions. Owing to this, the range of fast (MeV) charged particles in solids is only a few microns, in contrast to pile neutrons which have a range of the order of centimeters.

Fission fragments, as occurring in fissionable material such as uranium-base alloys during neutron irradiation, form a special kind of fast charged particles. They usually transport huge amounts of energy (100 MeV or more) and are heavy. Accordingly, they are extremely efficient in making lattice defects (up to 25000 Frenkel pairs per fission particle) and other disturbances. Moreover, they are usually unstable

and consequently β- and γ-rays as well as helium atoms are generated within the material. The damage produced by these particles is therefore highly complex (*fission spike*). Its occurrence is of paramount technical importance however, as we shall discuss briefly in Chapter XV.

2.2.3. *Electron irradiation*

Electrons of energy beneath the relativistic range do not possess sufficient momentum to displace atoms in the usual solid. However, electrons with energies of the order of 1 MeV energy can do so, at any rate in head-on collisions. The behaviour of fast electrons can be described by extending the discussion of the previous section to include relativistic effects in the passage of charged particles through matter.

The kinetic energy E of an incident particle (electron) will be expressed in terms of its rest energy mc^2:

$$\varepsilon_e = E/mc^2 \tag{21}$$

(ε_e is thus different from ε defined in the previous section). The momentum p_e of an electron with velocity v is

$$p_e = \frac{mv}{\sqrt{1 - v^2/c^2}} \tag{22}$$

and the momentum-energy relation is, of course

$$p_e{}^2c^2 = E(E + 2\,mc^2) = (mc^2)^2\varepsilon_e(\varepsilon_e + 2)\,. \tag{23}$$

As the displaced nucleon is always very much heavier than the electron, the momentum transferred in a head-on collision to the nucleus is at its maximum equal to $2p_e$, giving rise to a maximum transferred energy of

$$T_m = \frac{4p_e{}^2}{2M_2} = \frac{2E}{M_2c^2}(E + 2mc^2) = 561\,\frac{\varepsilon_e(\varepsilon_e + 2)}{M_2}\ \text{eV}\,, \tag{24}$$

where M_2 is, as usual, expressed in atomic units. A generalization of (14) for the differential cross-section follows by applying the Dirac equations to the problem instead of the ordinary Schrödinger equation. The result is:

$$\mathrm{d}\sigma = \frac{\pi a^2}{4}\,\frac{T_m}{T^2}\,f(v/c,\ T/T_m)\,\mathrm{d}T\,, \tag{25}$$

where f is a rather intricate function. SEITZ and KOEHLER [1956] give a table, computed by K. Fuchs (unpublished) which facilitates the computation of f. If T_m is but slightly larger than E_T, the integrated cross-section for displacement is

$$\sigma_{displ} = Z_2^2 \left(\frac{e^2}{mc^2}\right)^2 \left(\frac{T_m}{E_T} - 1\right) \left(\frac{1 - v^2/c^2}{v^2/c^2}\right), \tag{26}$$

which formula takes the place of formula (16). For very large values of $E(\gg mc^2)$ the cross-section becomes a constant, independent of the electron energy:

$$\sigma_{displ} = \frac{\pi}{4} \frac{a^2}{1 - v^2/c^2} \frac{T_m}{E_T} \qquad (E \gg mc^2), \tag{27}$$

and the mean energy transferred to the atoms is then

$$\bar{T} = E_T \left\{\log\left(\frac{T_m}{E_T}\right) - 1 + \frac{\pi Z_2}{137}\right\}, \tag{28}$$

which can directly be compared to (17). Also in this case the displaced atoms possess but relatively little energy and thus perform only elastic collisions. No secondary displacements are in general produced. The effect of screening is relatively unimportant. The range of electrons in most solids is rather small; a semi-empiral formula valid for copper is:

$$R \approx 400\, E(\text{MeV})^{1.26} \text{ microns}. \tag{29}$$

Heavy particles of mass M_1 with relativistic velocities can be treated in the same way as electrons. The maximum transferred energy is now M_1/m times larger than in the former case, and thus very large compared to E_T. Therefore, relativistic heavy particles behave very much as electrons with $E \gg mc^2$. Protons of 100 MeV and electrons of 2 MeV, to give but an example, behave comparably, although the range of the former is much greater. The discussion regarding the total number of displacements made by a given flux of impinging particles, outlined at the end of section 2.2.1., can be applied without difficulties to the case of charged particles. This holds in particular for the formulae (6), (7) and (8). Of course, the appropriate cross-sections for displacements must be used, in connection with the much smaller numbers of secondary displacements.

Conclusion: During electron irradiation the transferred energy is in all practical cases only slightly larger than the threshold energy E_T and

only *one* interstitial-vacancy pair per incident electron of about 1 MeV is produced. Heavy particles, on the contrary, usually effect quite a number of primary and secondary displacements per particle and neutral particles (neutrons) are especially productive: the total number of displacements amounts on the *average* to a few thousands per incident neutron of 1 MeV in 1 cm aluminium. Charged particles, losing their energy by electron excitation, produce 10 to 100 times fewer displacements than neutrons. Heavy charged particles (nucleons) have a very short range in solids. They thus create veritable "bunches" of lattice defects, in contrast with electrons, which give a far more homogeneous defect distribution.

2.2.4. *Irradiation with photons*

Photons (X-rays, γ-rays) do not interact directly with atomic nuclei to produce displacements, but they can transfer momentum to electrons by the Compton effect. The amount transferred is of the order of $h\nu/c$, the momentum of the photon itself. With this, an electron can in turn produce one displacement. We shall not discuss the mechanism of Compton scattering here; it is only important for γ-rays with energy above 0.1 MeV. Since the scattering cross-section is of the order of the square of the electronic radius times the factor $(h\nu/mc^2)^2$, where $mc^2 \approx 0.5$ MeV, the rate of energy dissipation of γ-rays by Compton scattering is usually small and a *very homogeneous* defect distribution can be obtained in this way.

X-rays have a different influence; they can excite bound electrons, or conduction electrons in metals, and produce ionization (*photoelectric effect*). The same is true for low-energy γ-rays. The liberated electrons again possess sufficient energy to produce incidental displacements. Neither the effects of X-ray bombardement nor of low-energy γ-irradiation are of much importance in metals, the cross-section for defect formation (about 10^{-26} cm^2) being too small to produce a sufficient number of displacements, even inside reactors with high fluxes up to 10^{16} photons/cm^2/sec. To the processes occurring in X-ray and γ-ray irradiated homopolar crystals and alkali halides we shall return in Chapters XXV and XXX. It should be mentioned at this point, however, that in irradiated ionic crystals an additional displacement mechanism may be operative (VARLEY [1954]), which is a direct result of photoionization. Negative ions becoming deprived of more than their number of valence electrons become positively charged. They then

find themselves at a very unstable position in the ionic lattice, and may easily be pushed into an interstitial position. Negative-ion vacancy-interstitial pairs can be produced in this way (but *not* positive-ion vacancy-interstitial pairs). Even X-rays should be able to produce displacements in this manner.

A third mechanism by which photons can make fast electrons which then can produce displacements is by *pair production* in the field of a nucleus. For this, the γ-rays must be very energetic (the mere energy of formation of a positon-electron pair is 1 MeV), and an appreciable cross-section is only to be expected at photon energies above 5 MeV.

2.2.5. *Replacement collisions and crowdions*

Apart from producing complete Frenkel defects (vacancy-interstitial pairs), fast moving atoms can also, in principle, *interchange* with normal lattice atoms (KINCHIN and PEASE [1955b]). The incoming atom then *replaces* the struck atom, and the latter becomes interstitial (in which state it may eventually cause other replacements or even displacements). No vacancy is left behind (compare interstitialcy diffusion, Chapter II). Of course, the incoming atom may well be of a different species from the replaced atom. In this way a very efficient way of *disordering* an ordered alloy can be visualized. It was indeed to explain the unusually high observed rates of disordering in neutron-irradiated alloys (for instance MnNi, ARONIN [1954]) that the concept of replacement collisions was introduced.

The mechanism will be important only if the energy E_R necessary to replace an atom is less than E_T, the threshold energy for displacement. Kinchin and Pease investigated this process theoretically, and DIENES and VINEYARD [1957] summarized their results in the following formula for the ratio r between the number of atoms replaced to the number of atoms displaced during fast neutron bombardment:

$$r = 1.614 \{\log (E_T/E_R) + 1\} . \tag{30}$$

No direct calculations of E_R have yet been made. Taking $E_R = \frac{1}{2}E_T$, the ratio becomes 2.7, sufficiently high to explain the observed disordering rates.

The large momentum given to a displaced atom may in some materials result in the formation of *crowdions* (Chapter II) rather than of interstitial atoms. This has some bearing on the annealing behaviour

of irradiated materials. Transport of momentum, resulting in displacement at several atomic distances from the site of transfer, may occur also ("focusing collisions", LEIBFRIED [1959]).

2.2.6. *Thermal and displacement spikes*

Not all of the kinetic energy of the incident particles or primary knock-ons is dissipated as potential energy of displaced atoms; a large part of it is taken up as kinetic energy or excitation energy and finally dissipated as heat. Along part of the path of the primary knock-on, there occurs therefore a region of particularly enhanced lattice vibrations, which has been called a *temperature spike* (Seitz and Koehler).

Consider a quantity of thermal energy Q released at a certain point (with temperature T_0) at $t = 0$. Let $\varkappa$ be the thermal conductivity, c the specific heat per unit volume (= specific heat per gram times the density) and d the density. Then, at a radial distance r, the temperature at time t will be given by

$$T(r, t) = T_0 + \frac{Q\sqrt{cd}}{(2\sqrt{\pi})^3} \frac{1}{(\varkappa t)^{3/2}} \exp\left(-\frac{cd}{4\varkappa t} r^2\right). \tag{31}$$

If a single atom receives from the incident particle a vibrational energy $Q < E_T$, formula (31) describes the diffusion of heat into the medium surrounding this atom. When the atom is in addition displaced, the spherical source model must be replaced by a somewhat more complicated model, more resembling a cylindrical source.

Seitz and Koehler concluded that within a time t_n of the order of 10^{-11} sec, the temperature in a spherical spike produced by the dissipation of 300 eV (which corresponds to a typical average pile-neutron displacement) will have become equal to the melting temperature of the metal over a region of radius r_d of the order of 10 atomic radii (1000 atoms). This spherical region, which is thus to be considered as molten for a very short time, is called a *displacement spike*. If the energy transmitted is less than E_T, only about one atom is "heated to the melting temperature", and one speaks of a *thermal spike*. A number of such spikes, separated from each other by cool regions, will occur along the path of a primary displaced atom.

Seitz and Koehler considered the possibility of diffusion of the atoms in a displacement spike, which can, for instance, result in local disordering of an originally ordered alloy. Of great importance is also the

thermal generation of vacancy-interstitial pairs and even dislocations in displacement spikes. The application of equilibrium formulae like (II–3) to the region of a displacement spike is of course not permissible; the region is far too small for that. Moreover, it is in a state of "turmoil" and "interstitials" are immediately adjacent to "vacancies". The problem is to find out how many of these *close pairs* will be retained after the temperature has again decreased (after a few times 10^{-11} sec). This depends primarily on the diffusivity of these defects. To obtain the integrated number of jumps n_j exerted by a defect in a spike, the simple formula

$$n_j = \nu_0 \int_0^\infty \mathrm{e}^{-Q/kT(t)}\,\mathrm{d}t\,, \tag{32}$$

can be applied, where Q is the activation energy for a jump (thus for diffusion) of the defect and $T(t)$ the time dependent temperature in the spike. Inserting for $T(t)$ the expression (31), for the typical case cited above $n_j \approx 6$ if $Q = 0.1$ eV; $n_j \approx 1$ if $Q = 1$ eV.

In the latter case only a certain (small) chance exists that the constituents of a close pair survive separately. We shall return to the question of the magnitude of Q in Chapter XV; both values quoted may actually apply in noble metals. Larger values of n_j could only be obtained if the diffusivity of *matter* in a particular material were much larger than that of *heat*, instead of being of about the same magnitude.

COTTRELL [1956] takes the view that the point defects present near to the boundary of a displacement spike can somehow escape recombination. This means that of a spike consisting of n defects (≈ 1000 in copper), about $n^{\frac{1}{3}} \approx 100$ point defects could yet be retained. The question whether in the temperature spike more defects can be formed and retained than corresponds to the initial number of displacements remains yet open.

Dislocations can be formed in a displacement spike by reason of the stresses occurring in it, as was first pointed out by BRINKMAN [1954]. The stresses around a displacement spike can be considerable owing to the high local temperature. The linear expansion f of a spherical zone when it melts is of order of 1% for most metals. The radial stresses are then of the order (cf. IV–33 and –36):

$$\sigma_r = 2Gfr_d^3/r^3\,, \tag{33}$$

which is at the boundary comparable to the theoretical critical shear

stress of a metal crystal, viz. $\approx G/30$. The (shear) stresses remain present during part of the cooling period of the spike, because to balance the decrease of f with decreasing temperature, the radius r_d of the spike increases.

To release these stresses, dislocation loops could be of some help. Their dimension can be estimated by equating the stress that tends to contract a loop of radius R into a point, viz. Gb/R, to the stress σ_r. One then finds $R = 5$ or 6 atomic radii in a typical displacement spike in copper. A dislocation loop of this size represents an energy of 30 to 100 eV, that is, it takes up a considerable part of the energy content of the spike. The persistence of such small dislocation loops is still largely a matter of conjecture. One way in which a stable pattern can possibly be formed is by intersecting and interlocking of different dislocation loops in the same spike by glide motion.

BRINKMAN [1954] has pointed out that when the average distance between displacements (i.e. secondary displacements or primary displacements at the end of the path of a primary knock-on) becomes of the same order as the interatomic distance, the individual displacement spikes will overlap, and a much larger "*Brinkman spike*" would result (about 100 atomic distances in diameter). It would first contain a large superconcentration of vacancies, whereas the interstitial atoms would have been "expelled" to the boundaries. Such a region, being inherently unstable, would easily recrystallize. The existence of Brinkman spikes is highly speculative. Some experimental evidence on the existence of locally molten regions in n-irradiated germanium has been obtained recently by GONSER and OKKERSE [1957].

Thermal spikes resulting from electron excitation along the path of the incident particle are of little importance, as it can be shown that the temperature in these spikes is relatively little enhanced over that of the surroundings, by a few hundreds of degrees centigrade at most. This together with the short duration of the spike, means that no significant atomic disorder can occur in such spikes.

In a recent paper SEEGER [1958] has reanalyzed the course of events leading to a Brinkman spike. His conclusion was that, instead of a zone in which a large number of vacancies and interstitials were present together, a relatively empty region, out of which the interstitials had completely disappeared in the form of *dynamic crowdions*, should be formed, at least in face-centred cubic metals. This is because during the displacement process the knocked-on atoms receive so

much kinetic energy that they should be able to move as crowdions over a distance of the order of 100 atomic distances, in specific directions (⟨110⟩) of course (estimate by Seeger). The crowdions should then finally come to rest at dislocations, pinning the latter and causing an increase of hardness, whereas the vacancies coalesce into a void, which is a relatively stable defect configuration and can only be repaired by the aid of self-diffusion. These voids also contribute considerably to the hardness of the material.

Seeger's model has been proposed in first instance to account for several features observed in *radiation hardening*, and we shall come back to it in Chapter XV.

The effect of neutron and electron irradiation on lithium fluoride crystals has been studied by GILMAN and JOHNSTON [1958]. These authors investigated the etched surface of neutron-irradiated crystals for specific markings and found several, which they interpreted as caused by the presence of vacancy clusters, constituted by fluorine vacancies. Similar markings were observed after electron bombardment. The clusters increased in size and decreased in number by annealing. Above 400 °C their structure radically altered: the clusters transformed into *dislocation loops*. This transformation is what one must expect according to the theoretical work of Kuhlmann-Wilsdorf, to be explained below.

Above 600 °C another transformation of the etch markings occurred; the dislocation loops dissolved and rectangular cavities appeared, with sides directed along the cube axes. More will be said about this in the next chapter.

2.2.7. *Comparison of various irradiation procedures for the production of radiation damage*

In the following Table 1 some of the most important distinctions between the amount and nature of the structural imperfection produced by irradiations of different types have been compiled.

To obtain a well-defined homogeneous defect distribution, bombardment of thin slices by electrons or γ-rays is to be preferred, although the efficiency of these methods, especially of the second, is rather low. A much higher production rate of defects is obtained by pile irradiation. However, then structural disorder in the form of spikes is also produced, together usually with a fair amount of radioactivity, which makes subsequent handling more difficult. This holds even stronger if

the material is fissionable. Ionizing irradiation below 0.1 MeV cannot generate lattice defects, it only causes colour centres and analogous electronic defects.

TABLE 1

Irradiation effects

Radiation	*Energy*	*Range*	*Electronic defects*	*Displacements*	*Spikes*	*Transmutations*
u.v. radiation	10–10^3 eV	variable	centres	none	none	none
X-ray radiation	10^3–10^5 eV	cm to m	centres	none	none	none
γ-ray radiation	10^5–10^8 eV	cm to m	centres	few	none	none
neutrons (thermal)	0.01–0.1 eV	variable	none	none	none	yes
neutrons (fast)	10^4–10^7 eV	cm	none	many	yes	few
charged nucleons	10^4–10^9 eV	μ to mm	excitation	yes	yes	few
electrons	$\geqslant 10^6$ eV	0.1 to 1 mm	excitation	few	none	none
fission fragments	$\approx 10^8$ eV	1 to 10μ	excitation	very many	yes	yes

3. Formation of Defects During Quenching

Since in most materials at elevated temperatures point defects exist in thermal equilibrium in appreciable concentrations (see Chapter II), large superconcentrations of defects can be *frozen in* by rapidly cooling such a material from a high to a low temperature, where no diffusion is possible. Vacancies can be quenched-in in copper, silver or gold, and both kinds of point defects in the alkali metals, in this way. A determination of the concentrations of the defects after the quench, e.g. by measuring the electrical resistivity, yields information on the energy of formation of the defects. The efficiency of the quench in retaining all defects depends on the quenching rate and on the diffusivity of the defects over the complete range of temperature. The fastest cooling rate so far obtained (by cooling a wire in a blast of cold helium, KAUFFMAN and KOEHLER [1955]), was a few times 10^4 °C/sec. It can easily be computed that during the quenching of a gold wire from 900 to 0 °C at this rate, the vacancies obtained in it can still make about 10^7 jumps (using the value for the diffusivity of vacancies mentioned in Chapter II). This means that when arbitrarily distributed traps are present to an atomic density of not more than 10^{14} cm^{-3} (10^{-6}%), a considerable fraction of the vacancies will reach these traps during the quench. Dislocations may easily provide such traps. It has been realized

later that elaborate measures to produce rapid cooling are not needed, since also during simple water or even air cooling, the first and most important 100° are passed in about 10^{-2} sec. What happens later is of much less importance.

A further complication inherent in quenching experiments is due to the possibility of mutual encounter of the defects during the quench, which may result in annihilation, or in the formation of associated defects which diffuse even more rapidly (vacancy pairs in the noble metals, for instance) or in the stable, immobile clusters which easily escape detection. In that case it becomes virtually impossible to retain all imperfections, however rapid the cooling may be.

A third disturbing factor is the simultaneous occurrence of some plastic deformation, caused by the thermal stresses set up in the material during the necessarily inhomogeneous cooling. The outer layers of the specimen always cool more rapidly than the inner parts, and, owing to differences in the thermal expansion, an inhomogeneous temperature gradient is always accompanied by stresses (see Chapter IV). If these stresses become so large that plastic flow can take place, part of them will be released at the cost of the formation of a number of dislocations in the material. As these can in turn act as traps (or sources) of point-defects, this invalidates to a certain extent the conclusions to be drawn as to the equilibrium concentrations before the quench.

From the theory of thermal stresses simple approximate formulae can be derived, which state the conditions to be obeyed either by the quenching speed v_q or by the radius R of the quenched wire, so that no plastic deformation by thermal stresses shall occur.

$$-\frac{\mathrm{d}T}{\mathrm{d}t} = v_q < v_{cr} = 5(1-\nu)\,\frac{\delta\sigma_{cr}}{\alpha G R^2}\,; \tag{34}$$

or

$$R < R_{cr} = \sqrt{5(1-\nu)\,\frac{\delta\sigma_{cr}}{\alpha G v_q}}\,. \tag{35}$$

Table 2 lists values of v_{cr} and R_{cr} for various materials. In these formulae, δ stands for the thermal diffusivity (heat conductivity/specific heat per unit volume), α is the coefficient of thermal expansion, ν is Poisson's ratio, and σ_{cr} and G are the critical shear stress and shear modulus of the material.

TABLE 2

Critical quenching speed and critical radius of wires

Material	*Heat diffusivity* δ (cm²/sec)	*Lin. expansion coeff.* α (°C)$^{-1}$	v_{cr} *at* $R = 0.1$ mm (°C/sec)	v_{min} (°C/sec)	R_{cr} *at* $v_q = 10^4$ °C/sec (mm)
Cu	1.00	1.8×10^{-5}	1.7×10^4	$\approx 10^4$	0.13
Au	1.21	1.5	2.4	$\approx 10^4$	0.16
Ag	1.63	2.0	2.4	$\approx 10^4$	0.16
Al	0.92	2.8	1.0	$\approx 10^5$	0.10
Ni	0.14	1.6	0.8		0.09
Pt	0.23	1.0	2.1		0.15
Fe	0.10	1.5	0.6		0.08
Mo	0.52	0.5	1.95		0.31
W	0.62	0.5	1.1		0.33
Ge	0.28	0.7	12 ($R = 1$ cm)	≈ 1	0.33
Si	0.42	0.3	42 (,,)	≈ 1	0.65

It should be realized that a lower limit to the quenching speed is set by the freezing-in condition. Inserting appropriate values for the expected trap density and defect diffusivity, these lower limits (estimated) have been entered in the column headed v_{min} of the table. Usually, only an uncomfortably narrow range exists between the two limits for v_q.

It will be clear that the results of quenching experiments must always be considered with much reserve, especially so if the samples were not in the form of very thin wires.

SEITZ, KOEHLER and BAUERLE [1957] have discussed theoretically the behaviour of vacancies in a quenched gold wire under the assumptions: 1) all defects originally present in thermal equilibrium are conserved, either as such or as pairs or triplets; there are no deep traps; 2) only a small fraction of the vacancies coalesce (rapid quench); 3) there is no long-range interaction between dislocations and defects; 4) near a dislocation thermal equilibrium is continuously established. They solved the equilibrium equations between the concentrations of the various defects by approximative methods and found the following values for the retained fraction of single vacancies after quenching (Table 3), for various values of the association energy U.

Appreciable changes in this fraction occur by association if U is larger than 0.1 eV. Seitz *et al.* concluded from this and the results of their experiments that the association energy in gold can not exceed the very small value 0.2 eV.

SCHOTTKY [1960] has reconsidered this problem, and has analyzed the average diffusion coefficient of single, di-, and trivacancies left over after quenching in the light of the latest knowledge about the mobilities of these defects. He obtained limits between which the various association energies must range, in order that the observations of Bauerle and Koehler can be explained.

TABLE 3

Fraction of vacancies which are single before and after quenching to 40 °C of gold wires (after SEITZ, KOEHLER and BAUERLE [1957])

Quenching temperature and association energy	Z_0 *(before quenching)*	Z *(after quenching)*	
		at $v_q = 3 \times 10^4$ °C/sec	*at* $v_q = 6 \times 10^4$ °C/sec
$T_q = 600$ °C $U = 0.1$ eV	0.9994	0.9878	0.9878
0.2 ,,	0.9957	0.8544	
0.3 ,,	0.9839	0.1816	
$T_q = 700$ °C $U = 0.1$,,	0.9962	0.9551	
0.2 ,,	0.9877	0.5542	
$T_q = 800$ °C $U = 0.1$,,	0.9900	0.8774	0.8860
0.2 ,,	0.9711	0.1494	
$T_q = 900$ °C $U = 0.1$,,	0.9777	0.7435	0.7581
$T_q = 1000$ °C $U = 0.1$,,	0.9569	0.5751	

A large number of quenching experiments have recently been carried out (see section XV–1.4) which indicate strongly that the interpretation of Seitz *et al.* of their own observations is not universally applicable. In particular, the results obtained by PANSERI and FEDERIGHI [1958] and by WINTERBERGER [1958] on quenched aluminium indicate the occurrence of certain processes by which vacancies interact with dislocations and vice versa, which were not taken into account in Seitz's theory. A full description of all theoretical considerations would carry us too far, we only mention at this point some interesting ideas advanced by MADDIN, KUHLMANN-WILSDORF and KIMURA [1958] (based on their experiments on copper) and by Panseri and Federighi that bear directly on this subject. According to the temperature of quenching, the history of the quenched-in vacancies is different. Quenching from a high temperature gives the vacancies the chance to make a large number of jumps; it will result in the formation of the most stable configuration: of vacancy *clusters*, which nucleate prefer-

entially at dislocations. These can only be dissolved when self-diffusion takes place. Quenching from a lower temperature leads predominantly to the formation by association of many divacancies that can migrate to dislocations. This resembles most the situation to which the theory of Seitz *et al.* applies. Quenching from a still lower temperature, where the concentration of vacancies is very small, makes the formation of divancies also improbable, and we are left with slowly migrating single vacancies only.

In accordance with this picture, after quenching from a very high temperature stable voids are formed and little resistance increase is observed, whereas subsequent annealing only results in a small effect. The hardness is increased at all temperatures up to the temperature where self-diffusion can take place. Quenching from a moderate temperature results in a complex annealing behaviour in which vacancies and divacancies both take part: the hardness is affected only near room temperature and below in Cu and Al. After quenching from a low temperature the annealing behaviour is again simple, but characterized by a relatively high activation energy, since the easy moving divacancies are absent. Little hardness increase is to be expected.

Predicted by KUHLMANN-WILSDORF [1957] and HIRSCH [1958] and beautifully confirmed by the latter is the process of the *conversion* of vacancy conglomerates into mobile dislocation loops and other structures that can act as dislocation sources. The vacancy conglomerates, in the form of disks, can for instance have been formed after a quenching treatment, along the lines discussed above, but they may also have been generated during plastic flow. It is assumed that the excess vacancies agglomerate in the form of disks of about 100 atomic distances diameter. Calculations by HOWE and ELBAUM [1961] show that energy is gained in transforming such a disk, which in the f.c.c. lattice will be situated in an octahedral plane, into a loop of sessile dislocation of the prismatic type. (This dislocation was mentioned in Chapter III and we refer for a more detailed consideration of its properties to READ's book [1953].) The transformation is simply equivalent to a collapse of the lattice across the vacancy disk. The loop being sessile, it can not move by glide nor act as a source (its Burgers vector $\frac{1}{3}a\langle 111\rangle$ lies outside the plane of the loop). However, a partial dislocation of the Shockley type which lies in the same octahedral plane can easily combine with the loop dislocation according to the dislocation reaction:

$$\tfrac{1}{6}a\langle\bar{2}11\rangle + \tfrac{1}{3}a\langle 111\rangle \rightarrow \tfrac{1}{2}a\langle 011\rangle .$$

Then, the loop dislocation is converted into a normal glissile dislocation of the type occurring in face-centred cubic crystals. This dislocation has been called an *R-dislocation* by Kuhlmann-Wilsdorf; it can act as a source of new dislocations as not all parts of the loops are equally mobile.

This is due to the fact that the dislocation loop is closed; not all of its parts can lie in suitable glide planes, and therefore, in a field of applied stress, the dislocation is still partly immobile. The mobile portions are then anchored by the immobile parts, which makes the loop an excellent source.

Electronmicroscopical observations of thin foils by Hirsch *et al.* have shown the formation of R-dislocations in quenched aluminium indeed to take place to a large extent. In gold the situation is slightly more complex; the difference in stacking fault energy of these two metals here plays a certain role.

About the effect of R-dislocations on hardness and resistivity we refer to the next chapter.

REFERENCES

ARONIN, L. R. (1954); J. appl. Phys. **25**, 344.

BILLINGTON, J. (1959); Sci. American **201**, 200.

BOHR, N. (1913); Phil. Mag. **25**, 10.

BOHR, N. (1948); Kgl. Danske Vidensk. Selskab., Biol. Medd. **18**.

BRINKMAN, J. A. (1954); J. appl. Physics **25**, 961.

COTTRELL, A. H. (1956); Brit. J. appl. Physics, Metallurgical Reviews Suppl. **5**, 43.

DIENES, G. J. (1952); J. appl. Phys. **23**, 1194.

DIENES, G. J. and G. H. VINEYARD (1957); Radiation Effects in Solids (Interscience, New York).

GIBSON, J. B., A. N. GOLAND, M. MILGRAM and G. H. VINEYARD (1960); Phys. Rev. **120**, 1229.

GILMAN, J. J. and W. G. JOHNSTON (1958); J. appl. Phys. **29**, 877.

GONSER, U. and B. OKKERSE (1957); Phys. Rev. **105**, 757.

HARRISON, W. and F. SEITZ (1955); Phys. Rev. **98**, 1530.

HIRSCH, P. B., J. SILCOX, R. E. SMALLMAN and K. H. WESTMACOTT (1958); Phil. Mag. **3**, 897.

HOWE, S. and C. ELBAUM (1961); Techn. Rep. NONR-562 (27), Metals Research Laboratory, Brown University.

HUNTINGTON, (1957); unpublished.

KAUFFMAN, J. W. and J. S. KOEHLER (1955); Phys. Rev. **97**, 255.

KINCHIN, G. H. and R. S. PEASE (1955a); Rep. Progress in Physics **18**, 1.

KINCHIN, G. H. and R. S. PEASE (1955b); J. Nuclear Energy **1**, 200.

KUHLMANN-WILSDORF, D. (1956); unpublished.
LEIBFRIED, G. (1959); J. appl. Phys. **30,** 1388.
MADDIN, R., D. KUHLMANN-WILSDORF and F. KIMURA (1958); Z. Metallk. **49,** 584; (1959) Phil. Mag. **4.**
PANSERI, C. and T. FEDERIGHI (1958); Phil. Mag. **3,** 1223.
READ, W. T. (1953); Dislocations and the Mechanical Properties of Crystals (McGraw Hill, New York).
SCHOTTKY, G. (1960); Z. Physik **160,** 16.
SEEGER, A. (1954); Z. Naturf. **9a,** 870.
SEEGER, A. (1958); Proc. Second Geneva Conference, Paper 998.
SEITZ, F. and J. S. KOEHLER (1956); Solid State Physics **2,** 305.
SEITZ, F., J. S. KOEHLER and J. E. BAUERLE (1957); Phys. Rev. **107,** 1499.
VARLEY, J. H. O. (1954); Nature **174,** 856.
WINTERBERGER, T. (1958); Louvain Conference on Low-Temperature Physics.

CHAPTER XV

PHYSICAL OBSERVATION OF DEFECTS IN METALS

1. Electrical Resistivity

1.1. GENERAL

As has been mentioned in Chapter V, dislocations as well as point defects cause scattering of the conduction electrons in metals, and thus are responsible for an increase of the electrical resistivity over the residual resistivity and the thermal resistivity. One usually assumes the approximate validity of *Matthiessen's rule*. This includes the simple additivity of the extra resistivity due to lattice imperfections and the residual resistivity of the metal, as well as the temperature independence of the former. Many indications exist to show that the theoretical assumptions on which this rule could be based are not fulfilled, and also in practice it is found that the resistance due to lattice defects cannot be regarded as independent of temperature (BROSS [1959], MAGNUSSON and PALMER [1958]). The situation is not yet very clear at this moment, and we shall refrain from a further discussion (see also GERRITSEN [1956]).

The ultimate purpose of a determination of the resistivity of imperfect metals is to separate the influences of the different types of lattice imperfections. In as far as this can be done, new information can be obtained on the scattering properties of the defects and thus indirectly on their structure and on the electronic structure of the metal, and on the concentration of the imperfections and their dependence on the nature and intensity of the previous treatment. However, this separation is very difficult. To attain it, measurements on quenched, irradiated and cold-worked metals and metals treated in more than one of these ways must be used in combination. Relevant evidence can further be obtained by observing the *recovery* of the electrical conductivity on annealing. The resistivity due to defects then disappears gradually or in stages owing to the diffusion of the various

defects out of the material or to sinks. It is of great importance to introduce the imperfections into the metal, and to measure their resistivity, at such low temperatures (preferably at the temperature of liquid nitrogen, hydrogen or helium) that no uncontrolled diffusion can occur and all the imperfections are retained in the lattice. Subsequent careful warming up then releases them successively, leading in principle to the separation wanted.

We shall limit our discussion only to those cases where these experimental conditions have been fulfilled. Further, we shall predominantly consider pure metals, as the corresponding effects in alloys are even more complicated, owing to the ordering and disordering processes that are induced by the various treatments. Some attention to this point will be given in Chapter XXI.

Until now, even in pure metals, the actual experimental situation is not yet completely clear. Since a number of reviews concerning this subject have appeared not long ago (BROOM [1954], GLEN [1955], VAN BUEREN [1957], COTTRELL [1958]) we shall here restrict ourselves to the "case-history" of the noble metals and note only the most important researches.

1.2. RESISTIVITY OF COLD-WORKED COPPER AND OTHER NOBLE METALS

Measurements up to 1953 have been adequately reviewed by BROOM [1954]. In Fig. 1 the most important results pertaining to polycrystalline copper, and obtained up to 1956, have been represented.

For purposes of reference and comparison, VAN BUEREN [1956] has tried to express the observational results in the simple powerlaw XIV–(1), containing only two parameters.

This proved to be possible surprisingly well for not too heavily deformed metals. Table 1 lists some values of α and p derived from the literature (for pure polycrystalline metals only).

BLEWITT, COLTMAN and REDMAN [1954] have measured the resistivity-strain curve of a copper single crystal at helium temperature. It could be represented by the empirical formula

$$\Delta\varrho = 0.01\varepsilon^2 . \tag{1}$$

Whereas the polycrystalline metals can obviously be well characterized by a value of p between 1.4 and 1.5, and a value of A of about 0.1

$\mu\Omega$ cm, the single crystal data clearly indicate a quadratic dependence on strain and a much smaller coefficient A.

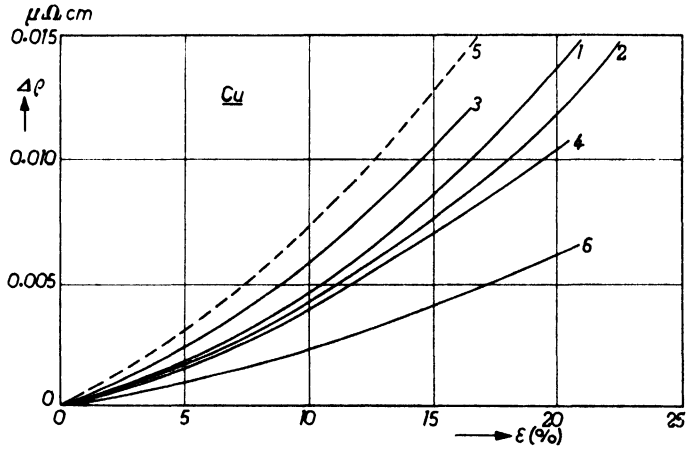


Fig. 1. Observed resistivity-strain relations of extended polycrystalline copper wires. Plotted is the absolute amount of additional resistivity against the mean shear strain, measured at 78 °K or less. (1) Observations by MOLENAAR and AARTS [1950]; (2) by DRUYVESTEYN and MANINTVELD [1951, 1952]; (3) by BERGHOUT [1957]; (4) by AARTS and JARVIS [1954]; (5) by PRY and HENNIG [1954], extrapolated; (6) by JONGENBURGER and the present author, see VAN BUEREN [1957].

TABLE 1

Parameters A and p (formula XIV–(1)), of the electrical resistivity-strain relation in pure polycrystalline noble metals deformed at low temperature

Author		*Metal and temperature of deformation*	A ($\mu\Omega$ cm)	p
MANINTVELD	[1954]	OFHC Cu, 78 °K	0.14	1.51
BERGHOUT	[1957]	id.	0.16	1.43
MOLENAAR-AARTS	[1950]	id.	0.09	1.3
AARTS-JARVIS	[1954]	id.	0.10	1.4
v. BUEREN	[1957]	id.	0.07	1.49
v. BUEREN	[1957]	OFHC Cu, 20 °K	0.07	1.7
v. BUEREN	[1957]	99.998% Cu, 78 °K	0.09[5]	1.3[5]
v. BUEREN	[1957]	99.998% Cu, 20 °K	0.05	1.48
MANINTVELD	[1954]	99.98% Ag, 78 °K	0.14	1.53
AARTS-JARVIS	[1954]	99.99% Ag, 78 °K	0.10	1.4
v. BUEREN	[1957]	99.99% Ag, 78 °K	0.05[5]	1.15
v. BUEREN	[1957]	99.99% Ag, 20 °K	0.04[5]	1.23
MANINTVELD	[1954]	99.98% Au, 78 °K	0.13	1.52
AARTS-JARVIS	[1954]	99.99% Au, 78 °K	0.07	1.4

In some alloys a *decrease* of resistivity with cold-work has been detected (KROPSCHOTT *et al.* [1959]). This may be connected to other peculiarities of the resistivity of these alloys, however (resistance minimum at He temperatures, and to disordering phenomena).

These observational data can now be compared to those expected theoretically, viz. by combining the theoretical results on the scattering cross-sections of imperfections (section V–1) with the theoretically expected concentrations (Chapter VII, Table 1, etc.). The observed $\varepsilon^{3/2}$ or ε^2 dependence can be explained equally well in all three models discussed in Chapter VII, and a decision which model deserves preference is therefore not possible in this stage. Assuming point defects to be formed by the moving-jog mechanism and the dislocations by Frank-Read sources, to obtain numerical agreement one should take the density of sources equal to about 10^{10} cm^{-3}, which seems a reasonable value.

Other observations, to be mentioned in the following sections, point to the conclusion that in the metals under discussion dislocations and point defects contribute about equally to the resistivity. The scattering cross-section of dislocations has been discussed in Chapter V. To explain the above-mentioned equality, either this cross-section is too small by a factor of about ten, or the actual vacancy concentration is much lower than it should be according to theory. Since Jongenburgers' formula (section V–1) for the scattering cross-section of vacancies is probably accurate to within a factor of two or three and interstitials scatter at least as much in copper (see Chapter V) it is only by assuming that the *efficiency* of vacancy formation by moving intersecting dislocations is considerably smaller than unity (about a few per cent in the average case), that we can interpret the observed resistivity changes in terms of point defects formed by the interaction of dislocations.

To the present author (VAN BUEREN [1955]) the latter solution to the dilemma appeared unsatisfactory for a number of reasons, to be discussed below, and he concluded that the theoretical dislocation cross-section for conduction electrons in metals is wrong by a factor of about ten. A new calculation by SEEGER [1960] has confirmed this.

To determine the dislocation concentration separately, in principle the *anisotropy* of the resistance of a copper single crystal deformed in easy glide could be measured. One would expect this to be present if dislocations were responsible for an appreciable part of the additional

resistance, because such imperfections scatter anisotropically (section V–1.2). However, no measurable anisotropy above 1% has ever been observed (BROOM and CLOTHIER [1952], SOSIN and KOEHLER [1956]). This is not very surprising, since in the modes of deformation used in these experiments dislocations of either type (edge, screw or mixed) participate, and the resultant anisotropy may very well have been practically zero. Care should be taken in experiments of this kind to obtain a distribution of dislocations which enables one to detect the anisotropy in full measure.

1.3. RESISTIVITY INCREASE BY IRRADIATION

A survey of the experimental determinations up to 1955 of resistivity changes in *irradiated* metals is contained in the review paper by GLEN [1955], and some new data are contained in the Institute of Metals Seminar [1958]; the rest is spread through the literature. Limiting

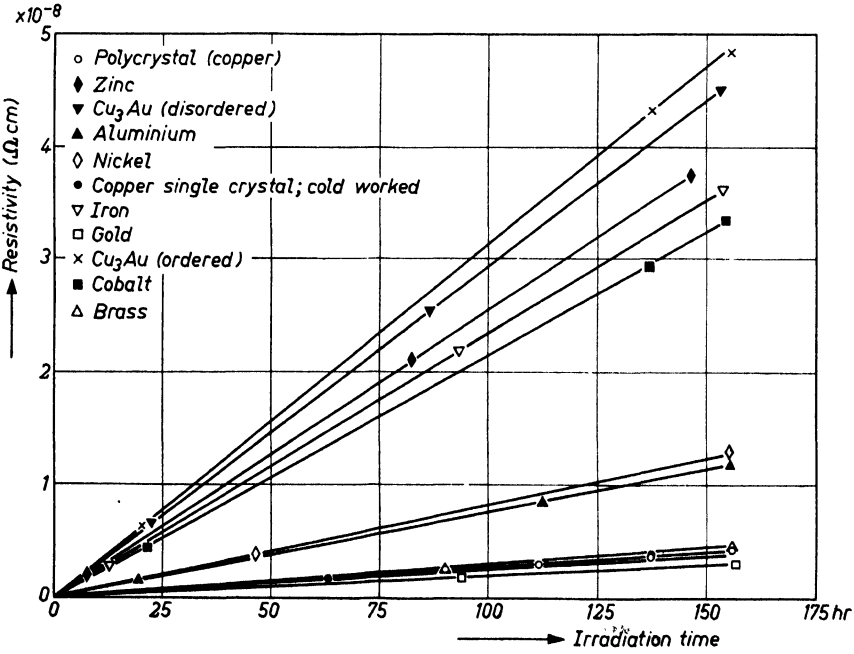


Fig. 2. The effect of neutron bombardment on the resistivity of several metals. The bombardment took place in the Oak Ridge ORNL graphite reactor at 18°K. (After BLEWITT *et al.* [1957].)

again our discussion to pure noble metals treated at very low temperatures, it is found that, apart from the occurrence of socalled "radiation annealing" at high doses, the additional resistivity is always strictly proportional to the total flux of incident particles received. Most authors give only relative resistivity changes. Those published until 1955 have been tabulated in Table 2 (mostly after Glen) and some results obtained by BLEWITT and coworkers [1957] have been compiled in Fig. 2. The origin of the large differences between various metals is not clear.

Only a few of the determinations yield reliable absolute values for the number of point defects formed during irradiation, because most of them were carried out at too high a temperature to prevent recovering taking place. COOPER *et al.* [1954, 1955] observed a resistance increase of 0.23 $\mu\Omega$cm in copper bombarded (in a cyclotron) at 78°K with 1.1×10^{17} *deuterons* of 12 MeV per cm^2. BLEWITT and coworkers [1954] found after *pile* irradiation of the same metal at 17°K (total integrated flux about 10^{17} neutrons/cm^2) an increase of only 0.0043 $\mu\Omega$cm. EGGEN and LAUBENSTEIN [1953] bombarded copper with 4.5×10^{18} cm^{-2} *electrons* of 0.81 MeV below 150°C and found a resistivity increase of 0.0044 $\mu\Omega$cm. In all three cases the observed resistivity increase is considerably smaller than that expected theoretically from the scattering cross-sections of point defects and the theoretical concentrations enumerated in the previous chapter. The discrepancy amounts to a factor 6.3 in the deuteron case, it lies between 10 and 12 in the neutron case, and it is 2.5 after electron irradiation (when using the average value of 3.7 $\mu\Omega$cm as the resistivity contributed by 1% Frenkel pairs in copper). Later observations also have led to similar discrepancies by a factor 3 to 20. Another estimate of the defect resistivity by JONGENBURGER [1957] as well as computations by OVERHAUSER and GORMAN [1956] yielded larger values for this quantity, between 5 and 10 $\mu\Omega$cm/% defects; then the discrepancy decreases, but nevertheless it still amounts to a factor 5 for neutrons. Evidently either the Seitz-Koehler displacement theory gives too large densities (in all three cases), or the scattering cross-sections of point defects are (also) false, or some recovery (by recombination) takes place even at helium temperatures. Of these explanations, only the first way out of the discrepancy seems acceptable at this moment.

SEEGER [1958b] has pointed out that, of the Frenkel defects formed during irradiation, an appreciable fraction may be mechanically un-

TABLE 2

Relative resistivity increase of irradiated pure metals at very low temperatures

Author		*Irradiation (particles and energy)*	*Temperature (°K) of irradiation*	*Metal*	*Integrated flux* (cm^{-2})	*Resistivity increase (at temperature of irradiation if not otherwise stated)*
MARTIN *et al.*	[1951]	α 33 MeV	120	Al, Cu		large
PEARLSTEIN *et al.*	[1955]	p 130 MeV	300	W	10^{16}	0.49%
PEARLSTEIN *et al.*	[1955]	p 260 MeV	300	W	10^{16}	0.32%
PEARLSTEIN *et al.*	[1955]	p 410 MeV	300	W	10^{16}	0.71%
RANDOLPH	[1952]	d 10 MeV	170	Mo		300% (at 1.2 °K)
COOPER *et al.*	[1954]	d 12 MeV	12	Cu, Ag, Au	1.1×10^{17}	13–15%
OVERHAUSER	[1953]	d 12 MeV	93	Cu	1.1×10^{17}	
DUGDALE	[1954]	n pile	300	W	10^{19}	10%
DUGDALE	[1954]	n pile	300	Mo	10^{19}	7%
DUGDALE	[1954]	n pile	300	Pt	10^{19}	1% (some annealing certainly present)
MCREYNOLDS *et al.*	[1955]	n pile	78–120	Cu	1.1×10^{19}	20% (at 78 °K)
MCREYNOLDS *et al.*	[1955]	n pile	78–120	Al	1.1×10^{19}	33% (at 78 °K)
EGGEN and LAUBENSTEIN	[1953]	e $\approx$ 1 MeV	93	Cu		
MEECHAN and BRINKMAN	[1959]	e 1.25 MeV	80	Cu	3×10^{18}	31% (at 40 °K)
WALKER *et al.*	[1959]	e 1.35 MeV	< 10	Cu	10^{17}	$8 \times 10^{-10} \Omega$cm
CORBETT *et al.*	[1957]	e 0.7–1.35 MeV	12	Cu (zone refined)	$1–2 \times 10^{16}$	$10^{-10} \Omega$cm

stable, for instance because the distance to another defect is only very small. These would then recombine even at helium temperature. The increase of the discrepancy so far between experiment and theory when heavier incident particles are concerned is attributed by this author to the formation of clustered defects and spikes in the latter case.

1.4. RESISTIVITY CHANGE DUE TO QUENCHING

KAUFFMAN and KOEHLER [1952] quenched gold wires in a helium blast and observed the resistivity to increase as a function of the quenching temperature. Assuming that all equilibrium defects are quenched-in, one can thus find their energy of formation. These experiments were afterwards repeated, first by the same authors (KAUFFMAN and KOEHLER [1955]), and later by BAUERLE [1957], who used water as a cooling medium. LAZAREV and OVCHARENKO [1955] performed the same experiment, but quenched at an even slower rate. BRADSHAW and PEARSON [1956, 1957] have quenched gold, platinum and aluminium wires in a similar manner and PANSERI and FEDERIGHI [1958] made a very detailed investigation of quenched aluminium. AIROLDI *et al.* [1959] have studied copper in this way. These and some other data are contained in Table 3.

Usually the energy of formation of the quenched-in point defects (vacancies in all metals mentioned) is indeed obtained in this way fairly accurately. However, in other respects the method presents certain difficulties. To one of these, that has a bearing also on the complex recovery behaviour, we shall devote a few words.

JONGENBURGER [1957] has compared some of the earlier results mutually and with a second class of indirect experimental determinations of the resistivity of defects. These are based on the observation of deviations, at temperatures near the melting point, of the resistivity, the thermal dilatation and other physical properties from their theoretical values at these temperatures, which can be predicted for a pure metal from the conventional Grüneisen theory. These deviations are then assumed to arise from the influence of point defects in thermal equilibrium. While for particulars we refer to Jongenburger's paper, the results can be summarized as follows.

From the observed absolute values of the resistivity, the equilibrium concentrations of vacancies at the melting point of copper and gold can be deduced by extrapolation. They are found to be substantially

TABLE 3

Resistivity increase of metals caused by quenching

Author		*Metal*	*Quenched*		*Energy of formation of defects* (eV)	*Energy of migration* (eV)
			from T_q(°C)	*to* T_f(°C)		
AIROLDI, BACHELLA and GERMAGNOLI	[1959]	copper	600–800	20	1.0	1.3
KAUFFMAN and KOEHLER	[1955]	gold	600–900	−190	1.28–1.1	0.68
BAUERLE, SEITZ and KOEHLER	[1957]	,,	500–900	20	1.02	0.66
LAZAREV and OVCHARENKO	[1955]	,,	500–900	20	0.79	0.52
BRADSHAW and PEARSON	[1957a]	,,	900	2	0.95	0.68
LAZAREV and OVCHARENKO	[1954]	platinum	900	20	1.2	1.1
BRADSHAW and PEARSON	[1956]	,,	1200–1400	2	1.4	1.1
BACHELLA, GERMAGNOLI and GRANATA	[1959]	,,	900–1400	20	1.20	1.13–1.48
BRADSHAW and PEARSON	[1957b]	aluminium	360–530	2	0.76	0.44
DESORBO and TURNBULL	[1959]	,,	≈ 500	20	0.79	0.52
WINTENBERGER	[1957]	,,	≈ 500	20		0.58 and 0.37
PANSERI and FEDERIGHI	[1958]	,,	≈ 420	2	0.76	
BRINKMAN, DIXON and MEECHAN	[1953]	Cu_3Au	450	20		1.2
BRADSHAW and PEARSON	[1957c]	Pt–Rh alloy	1200–1400	2	as Pt	as Pt
BRADSHAW and PEARSON	[1957c]	Pt–Au alloy	1200	2	as Pt	as Pt
WECHSLER	[1957]	Au–Cd	450	40	0.38	0.58

greater than those derived from the results of quenching experiments (the discrepancy amounted to factors between 10 and 20). Also the energy of formation of the defects came out differently in both kinds of experiments, a smaller value being derived from the high-temperature observations (0.8 eV in gold, compare with Table 3). The correspondence between the two experimental methods is thus very bad. There are reasons, indicated already in section XIV–2.3, to believe that what is measured after quenching experiments is not the original vacancy concentration, but the concentration of a mixture of vacancies, vacancy pairs, etc., and that part of these defects have already disappeared from the material during the quench. Against the high-temperature method can be brought forward that the comparison between experimental and theoretical temperature variations of physical quantities is not very reliable, since neither is known very accurately at high temperatures.

It is especially the study of the *recovery* of the quenched-in resistance that leads to interesting conclusions; we refer to section 5 for a discussion.

1.5. CHANGE OF MAGNETORESISTIVITY DURING COLD-WORK

VAN BUEREN [1956] and JONGENBURGER [1960] have observed the resistivity variation of cold-worked polycrystalline copper, silver, gold and aluminium wires in a strong transverse magnetic field. By plotting the relative resistance increase due to the magnetic field versus the field strength divided by the zero-field resistivity, they were able to show that impurities and other point defects did not affect the appearance of such a "*Kohler diagram*", but that dislocations produced significant changes in the position of the Kohler curve. The *sign* of this change was found to be different in different metals and dependent on the temperature. However, if the temperature of measurement was sufficiently low (with respect to the Debye temperature of the material), the shift was always in the direction of *increasing* magneto-resistance. The shift of the Kohler curve increased with the amount of cold-work, to reach a saturation after about 50% strain. In strong magnetic fields the effect disappeared. Fig. 3 shows some of the results.

The effect of dislocations is obviously to produce an increase of the "reduced magnetoresistivity", at least at low temperatures. This effect is presumably due partly to the anisotropic scattering by dislocations, and partly to distortion of the energy-momentum surfaces of the

conduction electrons, caused by the long-range stresses around these imperfections. VAN BUEREN [1957] has shown the first cause to give in first approximation an increase of the reduced magnetoresistivity

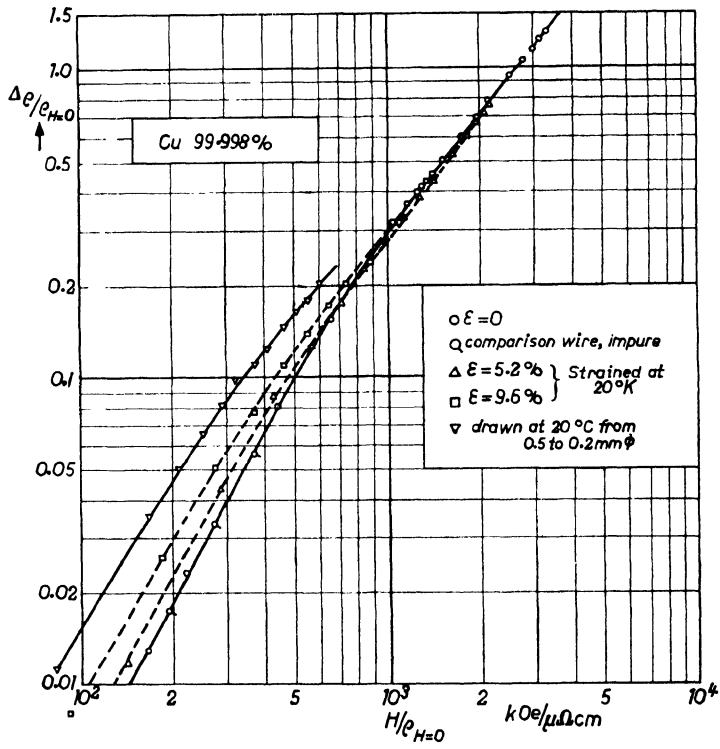


Fig. 3. Relative magnetoresistivity $\Delta\varrho/\varrho(0)$ of polycrystalline copper at 20 °K plotted logarithmically as a function of the reduced field strength $H/\varrho(0)$. Plastic strain results in an observable increase of the magnetoresistance, impurities have no effect.

of the observed order of magnitude. (Dislocation scattering was in this approximation characterized by an anisotropic relaxation time.) As to the second factor, no computations do yet exist. At first sight it could equally well have a positive or negative influence.

2. Thermal Lattice Conductivity

2.1. THEORETICAL

To describe theoretically the conduction of heat in metals and other solids is much more difficult than the conduction of electricity. There-

fore, the effect of lattice imperfections on the former has not been studied as extensively as that on the latter. Nevertheless, in recent years it has been shown that by measuring thermal conductivities in suitably chosen materials at very low temperatures (between 1 and 100 °K), useful information about the concentration of imperfections can be obtained, and that the *separation* of the effects of dislocations, stacking faults and point defects seems well feasible by this method.

This very important end can only be attained when one succeeds in removing the electronic contribution to the thermal conductivity as much as possible, since the influence of imperfections makes itself known especially in the *lattice* component of the conductivity. The scattering of phonons (lattice waves) by dislocations and other defects is relatively much more important than that of electrons. It is this that makes the study of the thermal conductivity such a useful tool, but it also determines its limitations. The method can be used only when electrons do but little contribute to the conduction, that is in *insulators*, in *metallic alloys* with a large residual resistance (thus not in pure metals) or in semiconductors and in *superconductors*.

KLEMENS [1956] has discussed the theoretical scattering cross-section for phonons of the various possible crystal imperfections. This scattering was described in terms of a relaxation time τ, which determines the rate of return to equilibrium of the disturbed phonon distribution function in momentum space. If the disordered region around the imperfection is spherical and extends only over a few atomic distances, and thus is small compared to the phonon wavelength $hc/k\theta_D$, where θ_D is the Debye temperature and c the velocity of sound, ordinary Rayleigh scattering applies, for which τ is inversely proportional to the fourth power of the frequency ν. Dislocations cannot be approximated simply by narrow cylindrical disordered regions, as their long range strain field is of the order of the wave length in width. Instead of a proportionality of τ to ν^{-3} as one would expect for Rayleigh scattering, one has because of this strain field rather $\tau \sim \nu^{-1}$. Planar imperfections like grainboundaries and stacking faults are described by a frequency independent relaxation time.

The temperature dependence of the heat conductivity $\varkappa$ is directly related to the frequency dependence of the relaxation time of the scattering process(es), as one has

$$\varkappa = \alpha \int_0^\infty Sc^2\tau \frac{\mathrm{d}k}{\mathrm{d}\nu} \mathrm{d}\nu, \tag{2}$$

where α is a numerical factor ($\frac{1}{3}$ in isotropic crystals), S is the specific heat carried by a lattice wave of momentum vector k, c is the group velocity of these waves and ν their frequency. All quantities depend on frequency as well as on temperature. The first variable is integrated out and one easily finds, if $\tau \sim \nu^{-n}$:

$$\varkappa \sim T^{3-n} . \tag{3}$$

We thus obtain the following scheme for the temperature variation of the thermal conductivity (Table 4):

TABLE 4

Temperature dependence of the thermal conductivity decrease due to imperfections

Scattering by	*Conductivity*
Perfect crystal (insulator) $T < \theta_D$	$\varkappa \sim T^3$ until maximum
,, ,, ,, $T > \theta_D$	$\varkappa \sim T^{-1}$
,, ,, (metal) $T < \theta_D$	$\varkappa \sim T^2$ until maximum
,, ,, ,, $T > \theta_D$	$\varkappa \sim$ constant
Point defects	$\varkappa \sim T^{-1}$
Dislocations	$\varkappa \sim T^2$
Stacking faults	$\varkappa \sim T^3$
Grain boundaries and surface	$\varkappa \sim T^3$

By studying the *temperature variation* of the lattice thermal conductivity of imperfect crystals, information can thus be obtained as to the nature and number of the imperfections present. In Table 4 also the temperature dependence of $\varkappa$ in perfect crystals has been included. At very low temperatures this is governed in insulators by the T-dependence of the specific heat, and varies consequently with T^3. Also scattering by the boundaries of the sample may be of some importance; it leads to a similar T^3-dependence. The conductivity reaches a maximum at higher temperatures and then falls of inversely proportionally to T. In metals phonon-electron scattering is usually the most important process, it leads to a T^2-dependence at low temperatures. Of course, the electronic contribution to the thermal conductivity, which we shall not discuss here, determines also the conductivity and disturbs the above-mentioned dependence at low and high temperatures.

In the superconducting state the electronic contribution to the thermal conductivity is of course strongly reduced, and also the scattering

of phonons by electrons becomes far less important. Therefore, the thermal conductivity may as well increase or decrease. However, what is important for our purpose, it is only the lattice conduction that remains. In an ideal superconducting metal crystal one should not expect any contribution by the electrons at all and a T^3-proportionality of $\varkappa$ should prevail.

Experiments often do not confirm the theoretically derived temperature dependences of $\varkappa$ very well. This is because a number of other processes contribute to the phonon transport and scattering, which we shall not go into here; we refer to KLEMENS's papers [1956, 1958] for a discussion.

2.2. EXPERIMENTAL

Several authors have measured the temperature variation of the thermal conductivity of metal alloys in the liquid helium range. An example is given in Fig. 4 for α-brass (KEMP *et al.* [1957]). Below 20 °K $\varkappa$ was found to be proportional to T^2 in well-annealed samples. Plastic straining (drawing at room temperature), resulted in the same T^2-dependence, but the conductivity was appreciably *decreased*. This decrease was attributed to *dislocations*. KLEMENS [1958] has calculated the scattering cross-section of dislocations for phonons and CARRUTHERS [1959] has corrected this (see Chapter V): for copper, $\varkappa_{\text{disl}} = 3 \times 10^6 \, T^2/N$, where N is the density of dislocations. From this, one can estimate the density of dislocations responsible for the observed decrease to be about 10^9 cm^{-2} in brass with 2% Zn and 2×10^{10} cm^{-2} in brass with 30% Zn. These seem not unreasonable values.

At higher temperatures the $1/T$ dependence to be expected for point defect scattering becomes apparent. This also increases with plastic deformation. ROSENBERG *et al.* [1959] have recently investigated quantitatively the dependence of the dislocation density in brasses upon the zinc content and the amount of plastic deformation. They found the number of dislocations in the first stages of strain to increase about proportionally to the $\frac{3}{4}$ power of the strain, in accordance with van Bueren's formula VI–(8) (derived on the assumption that the creation of point defects is the path-limiting factor for the expansion of dislocation loops). The absolute density of dislocations was found to be more or less of the expected order of magnitude. With increasing zinc content the density of dislocations attains its saturation value after a smaller amount of strain and some evidence for the presence of

stacking faults was found in the $\varkappa$-T curves. This is in accord with the ideas of Seeger as expanded in Chapter VIII; more zinc results in a lower stacking fault energy and a larger width of the extended dislocations.

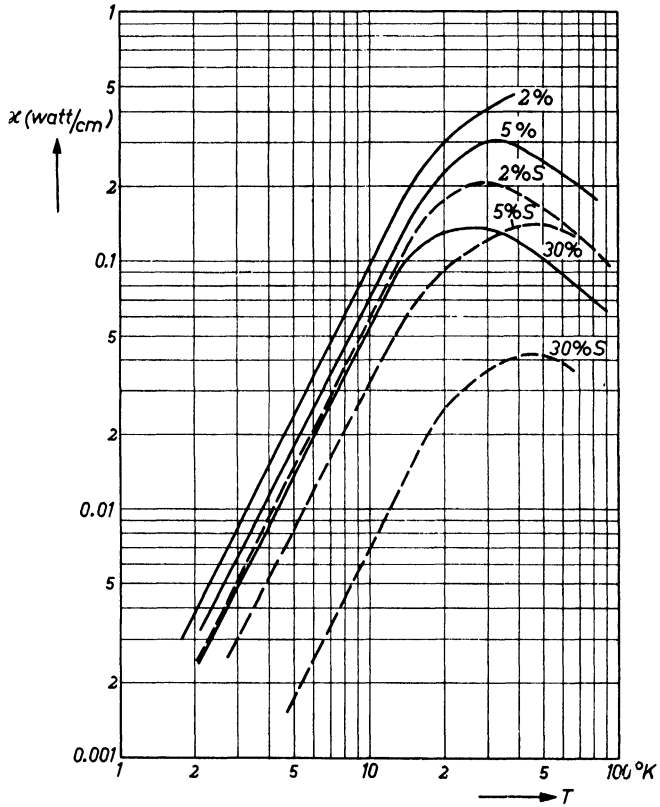


Fig. 4. Lattice thermal conductivity of α-brass (zinc content indicated), annealed and cold-drawn (S) as a function of temperature. (After KEMP *et al.* [1957].)

MENDELSSOHN [1958] has investigated superconducting lead (face-centred cubic) and niobium (body-centred cubic) and found a large decrease of the lattice conductivity on deformation. Again good agreement with van Bueren's formula was obtained in lead, but not in niobium, where the dislocation density deduced was much greater than the theoretical value. The scattering cross-section is very uncertain, here, of course.

KEMP *et al.* [1958] studied the thermal conductivity of deformed brass after annealing together with the *release of stored energy, recovery* of *electrical resistance* and of *density*, to be discussed in the following sections. They investigated only the recovery above room temperature and found a large energy release during annealing at 250 °C but small influence only on the heat conductivity. On the contrary, at 400 °C the lattice conductivity changed appreciably indicating a decrease of the density of dislocations and possibly other defects. This is in good accord with the proposed recovery behaviour to be discussed in section 5.

The investigation of lattice imperfections by thermal conductivity measurements is still in the preliminary stage. It might appear to be a powerful method indeed in the near future.

3. Energy Stored During Cold Work, Irradiation and Quenching

During cold work, energy is expended by the applied stresses and used to deform the material. By far the greater part of this energy is sooner or later transferred into heat, and lost. Only a few percents is *stored* in the form of lattice distortions and imperfections. As discussed in Chapter VII, in metals practically all the strain energy has been used to form lattice defects, or stored in their interaction. Knowing the energies of formation of these imperfections, from calorimetric measurements of the amount of stored energy (CLAREBROUGH *et al.* [1952, 1955, 1956]; see also BOAS [1957]) some information can be inferred about their concentrations. Again, as with the resistivity experiments, these measurements must be carried out at low temperatures in materials deformed also at low temperatures, in order to prevent the annealing out of the point defects. A complete review of experimental data up to 1958 has been given by TITCHENER and BEVER [1958].

Not only after cold-work, but also after irradiation energy is stored in the material in the form of lattice defects. The fraction of the incident energy that is stored is even smaller than that after cold-work, viz. only about 0.01%, since practically all the energy is dissipated as kinetic energy of the lattice atoms (heat). However, whereas during the deformation of metals the total amount of energy expended per unit volume (stress times strain) is only of the order of 1–100 cal/cm^3, the total kinetic energy released into a metal (irradiated in a pile

by 10^{19} fast neutrons/cm^2) is much more, about 10^6 cal/cm^3. The absolute amount of stored energy is therefore also very much larger in the latter case; it can amount to a few kilocalories/mole. Although quenching also causes energy to be stored in the form of lattice imperfections, experimental determination of this small quantity is not very well practicable.

Measurements of the stored energy get additional importance when they are combined with resistivity determinations on the same sample (BOAS [1957]). One can thus arrive at an important quantity, namely the "resistivity per cal/gram stored" (OVERHAUSER [1953]). As SEEGER [1955] has shown, this quantity allows us to draw interesting conclusions regarding the type of defect responsible for the observed effects. Measurements of this kind are usually also combined with studies of the recovery behaviour of the material; we shall discuss them in section 5. We refer to the review paper cited above for more details.

4. Changes of Volume and Lattice Parameter

Irradiation, cold work and quenching treatments result in slight changes of the lattice parameter and the volume of the material, again because imperfections have been created in the lattice. The linear dimensions usually change but by a few hundredths of a per cent, and the measurement is difficult. The results become more valuable when microscopical and dilatometric studies are combined with X-ray lattice parameter determinations.

Interpreted in terms of Frenkel defects, each of which contributes ≈ 1 atomic volume (see below) to the volume expansion, the concentration of defects formed after irradiation is invariably found 10 to 20 times lower than the theoretical concentration according to SEITZ and KOEHLER [1956]. This is again a similar discrepancy as was mentioned earlier in section 1.4. SIMMONS and BALLUFFI [1958] and VOOK and WERT [1958] investigated deuteron-irradiated copper in a dilatometer and by X-rays. They observed an equality between the *relative linear expansion* and the *change of the lattice parameter*. A striking parallelity between the *recovery* of these changes and the recovery of resistivity was further observed, which will be commented upon in the next section. TUCKER and SAMPSON [1954] pointed out that, as far as the lattice parameter is concerned, in metals interstitials must have much more effect than vacancies, and concentrations as low as 0.01% should

be detectable by X-rays. From the theory discussed in Chapter II it follows that the numerical interpretation of volume and lattice parameter variations in terms of concentrations of point defects is rather uncertain, as the contribution of the *relaxation* of the lattice around the defect can only be estimated with difficulty. TEWORDT [1958] and Seeger and Mann (unpubl.) recently obtained 1.3 at. vol. per % Frenkel defects as the most probable value for the lattice expansion in copper.

The effect of dislocations on the volume of a crystal has been formed the subject of much controversy. Not only the strain field but also the dislocation core contributes to this effect. Using the methods of non-linear elasticity theory, SEEGER and HAASEN [1958] estimated the dilatation by a dislocation of arbitrary type to be about 1 atomic volume per lattice distance (see also Chapter XVI).

5. Thermal Recovery of Influence of Lattice Imperfections

5.1. SURVEY

A large and still growing literature exists on the thermal annealing of the physical properties of metals subjected to one of the treatments mentioned above. For particular details the reader is referred

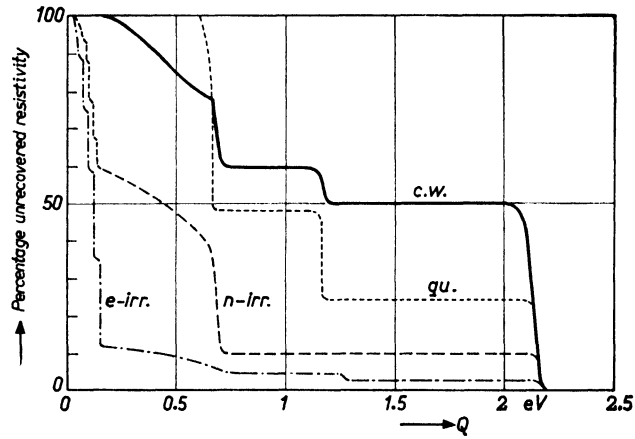


Fig. 5. Schematic representation of the recovery of electrical resistivity of noble metals cold-worked, irradiated or quenched. The five recovering steps I, ... V are characterized by more or less well defined energies of activation Q. Between the stages also often some gradual recovery is observed (garbage annealing).

to the review papers by BROOM [1954], GLEN [1955], VAN BUEREN [1957], etc. It has been possible to classify roughly the recovery be-

TABLE 5

Recovery of physical properties of copper and gold

	Step I *percentage recovery*	*II* *percentage recovery*	*III* *percentage recovery*	*IV* *percentage recovery*	*V* *percentage recovery*
Recovery of electrical conductivity after cold-work	0	20% gradual	20%	10%	50%
Recovery of electrical conductivity after n- or d-irradiation after electron irradiation	40% 90%	20% gradual 5%	30% small	0 2%	10% 3%
Recovery of electrical conduction after quenching	not investigated	0	50%	30%	20%
Recovery of magnetoresistivity after cold-work	not investigated	0 (+ 30%; silver only)	0	0	100%
Release of stored energy after cold-work	0	≈ 40%	≈ 30%	10% (nickel)	10%
Release of stored energy after irradiation	25%	25%	30%	10%	10%
Recovery of lattice parameter after d-irradiation	60%	0–10%	20%	0	10%
Recovery of volume expansion after d-irradiation	60%	0–10%	20%	0	10%
Activation energy in Cu (approximately)	0.1–0.2 eV	0.2–0.6 eV	0.7 eV	1.2 eV	2.1 eV
Temperature region	30–40 °K	90–200 °K	210–320 °K	≈ 400 °K	> 500 °K
Kinetics	2nd order or complex, detailed structure		2nd order + dislocation effects	Diffusion	1st order

haviour of copper and gold, the metals which have been studied most extensively, according to the scheme presented in Table 5 and Fig. 5.

Copper and gold show, as far as can be judged from the latest determinations, quite comparable annealing behaviour even as regards the domains of temperature in which the various recovery stages occur (except stage IV). Therefore, observations on both metals have been used together in the compilation of this table and figure. Occasionally, data on other metals have been included.

As is shown in the table, the resistivity of cold-worked metals recovers in four more or less well-defined stages, which we have numbered II, III, IV, V respectively. In irradiated metals step IV seems to be weak or even absent. However, a strong very low temperature recovery stage, stage I, is invariably found after fast-particle irradiation, which is absent in the cold-worked material. In electron-irradiated copper a very large part of the additional resistivity anneals out below 100 °K, thus in step I, and this step shows a structure in which no less than five substages are present (CORBETT and WALKER [1958]). In quenched metals only steps III, IV and V have been observed so far.

The magnetoresistance of cold-worked metals anneals out completely during the fifth step whereas steps II to IV are absent; step I has not been investigated, but in view of the absence of this step in the recovery of the ordinary resistivity, one can be reasonably certain that it does not occur either in the recovery of magnetoresistivity. During step II in one case (silver) a remarkable increase of the magnetoresistivity has been observed (JONGENBURGER [1960]). The release of stored energy has not been followed as accurately as the recovery of the conductivity. BLEWITT [1957] has shown step I to be practically *absent* in the stored energy release measurements in irradiated metals. Step IV has been observed in nickel (BOAS [1957]). The fifth stage is strongly present in cold-worked metals, and possibly also in irradiated material. According to SIMMONS and BALLUFFI [1958] there covery of the lattice parameter of d-irradiated copper runs exactly parallel to that of the electrical conductivity. Also the recovery of dilatometric changes follows this pattern (VOOK and WERT [1958]).

With the various steps can be associated *activation energies*, except for the "garbage annealing" of stage II. The determination of these quantities from the temperature and time dependence of the relevant recovery process should be carried out with much consideration. Obvious complications are that the kinetics of the process is of a high

order, or that a *range* of activation energies occurs. BRINKMAN *et al.* [1953] have outlined a reliable analytic method.

The region of temperature in which the recovery stage takes place within a measurable time interval has also been mentioned in Table 5.

5.2. INTERPRETATION OF RECOVERY IN TERMS OF DIFFUSION PROCESSES

An unambiguous interpretation of all this complex recovery behaviour has not yet been given. That it is connected with diffusion processes of point defects, leading to their partial or complete elimination, is obvious. The various diffusion processes that one can envisage in metals are the following:

1. *Close (interstitial-vacancy) pair recombination*

This is presumably especially important in irradiated metals, where such close pairs are abundant. In cold-worked metals interstitials and vacancies also exist together, but they are distributed much more at random and not necessarily in equal numbers. In quenched (noble) metals there are no interstitials at all. It seems therefore reasonable to expect a recovery stage due to this mechanism primarily in irradiated metals. The activation energy for the process would be the migration energy of the fastest moving component or even less, because the near presence of the other component might well reduce the height of the potential barrier for the migration of the first one. According to Huntington's calculations (see Chapter V) an interstitial in the noble metals would be characterized by a migration energy of only a few tenths of an eV. In view of this and of the occurrence in irradiated metals only, it seems appropriate to attribute part or all of the substeps of the first recovery stage to close-pair annihilation. Recent measurements of the release, during step I, of the amount of stored energy in electron-irradiated copper and gold showed this quantity to be in fair agreement with this supposition (MEECHAN and SOSIN [1959]), concluding herewith an argument of long standing whether an "energy paradox" exists during stage I or not. From the observed decrease of resistivity the number of annihilated defects can be inferred, and from this again the theoretically expected value of the energy release: this was found in early experiments to be about 25 times higher than the observed amount. Although the discrepancy has now been removed, its previous existence, and also a revisement of Huntington's calculation by SEEGER [1956], resulting in a much higher theoretical migration energy

for interstitials, have led to several other explanations for stage I. One of them (BLEWITT [1957]) is in terms of the motion of *crowdions* rather than of normal interstitials. These might well require only a very small activation energy to move in their length direction. The strong effect of impurities also found by Blewitt *et al.* is in favour of this hypothesis. A second explanation (KINCHIN and PEASE [1955]) makes use of a concept originally due to Huntington and Seitz (Chapter XIV), of the specific kind of interstitial migration called *interstitialcy* migration. Here the moving interstitial atom changes place with a normal atom, pushing the latter into an interstitial position. This might require even less energy than the normal interstitial motion. CORBETT and WALKER [1958] have recently investigated the first recovery stage in electron-irradiated zone-refined copper very carefully, and found it to exist of five substages (numbered I-A to E). In order to explain this, the mutual dependence between the relative importance of the stages and the conditions of purity, radiation dosis and of annealing, and finally the complicated annealing kinetics of these five substages, they advanced a model theory in which the migration of two kinds of interstitial atoms, leading to close-pair recombination and elimination at dislocations play a part. It may well be that one of these interstitial types is the crowdion.

A final possible explanation of stage I recovery might be found in attributing it to local rearrangement processes occurring in displacement spikes. As the first step also occurs, and even extremely strongly, in electron-irradiated metals, this seems improbable. A way out of this difficulty could be found in assuming that after electron irradiation the number of close pairs is exceptionally large and compensates for the lack of displacement spikes (SEEGER [1958b]). No completely satisfactory explanation for stage I recovery can be offered at this time. It is probably different according to the type of irradiation used.

2. *Diffusion of interstitial atoms* (not resulting in close-pair annihilation)

This process could lead, for instance, to absorption of interstitials by dislocations, thereby causing the latter to climb.

The process would be expected to occur especially intensively in cold-worked metals, probably also in irradiated metals, if enough dislocations are present, but not in quenched copper and gold. The energy of activation would lie either in the range from 0.1 to 0.4 eV (HUNTINGTON [1953]) or be about 0.6 eV (SEEGER [1955]). In view also of the

absence of step II in quenched metals, the present author is inclined to attribute *this* stage to interstitial (or possibly also crowdion or interstitialcy) migration. Other authors have assigned this process to either step I or III, however. The arguments advanced in the interpretation of step I make the first alternative somewhat improbable. In view of the observed wide range of activation energies observed in step II after n- or d-irradiation, possibly a number of processes operate together. Interstitial-vacancy pair recombination of successively wider spaced pairs may be responsible for part of this recovery process (SEEGER [1955]) or, what seems more probable, the *release* of interstitials trapped at an impurity (HASIGUTI [1960]).

3. *Diffusion of vacancies and divacancies*

Vacancy diffusion must be expected to take place after all three kinds of treatment mentioned, for in each of them vacancies are formed in large numbers. However, which particular recovery stage reflects the diffusion of vacancies has become a matter of serious disagreement between various investigators. Until recently most of them based their conclusion on the theoretically calculated energy of migration of vacancies in copper by HUNTINGTON and SEITZ [1941] which lies between 0.5 and 1 eV. They regard the *third step*, which is indeed clearly present after all three treatments, as due to vacancy migration. This computation is very crude, however, and has been replaced by a number of more accurate studies, which all lead to an energy of migration of at least 1 eV in copper, and 0.8 eV in gold (SEEGER and STEHLE [1956]).

The results of various quenching experiments by Koehler and coworkers (BAUERLE *et al.* [1957]) again appear at first sight to be in agreement with the association of stage III to vacancy diffusion. To give an example: Bauerle's experiments lead to an energy of formation of vacancies (viz. 1.02 eV in Au), which, together with the known energy of activation for selfdiffusion (1.81 eV in Au and 2.05 eV in Cu) points indeed to a migration energy of about 0.6 eV. This latter energy was actually observed as the activation energy for the recovery of the quenched-in resistivity of gold.

On the other hand, the variation of the coefficient of dilatation and of the specific conductivity at high temperatures (MEECHAN and EGGLESTON [1954], JONGENBURGER [1957]) leads to an energy of forma-

tion of vacancies in copper and gold which is somewhat smaller than 1 eV. The same quantity has also been recently recalculated theoretically by FUMI [1955] and others, who all found the small value 0.7 to 0.8 eV. Assuming *this* to be true, the energy of migration of vacancies in copper should be 1.1 eV, in gold about 1 eV, and *stage IV* could well be interpreted as vacancy diffusion. MEECHAN and BRINKMAN [1959] have indeed found a distinct stage IV recovery in electron-irradiated metals, although by many other investigators the step has *not* been detected in bombarded materials. The fourth step is certainly present in cold-worked metals (JONGENBURGER [1960]) although it is again rather shallow. It has clearly been shown to be present in quenched metals. Stage IV seems thus to be present after all three treatments, but it is always much less conspicuous than the third stage.

Finally, the quenching experiments carried out by Germagnoli and his group seem to indicate much higher recovery energies than earlier observations.

A few more observations that shed some light on this controversy are the following. Studies on the ordering kinetics of proton-irradiated Cu_3Au (in which about the same recovery behaviour is observed) by BRINKMAN, DIXON and MEECHAN [1953] have shown that the stage III recovery is associated with the migration of a lattice defect that has no effect on the degree of order. On the other hand, during step IV a defect migrates that can produce ordering. Since presumably only Cu atoms can occupy interstitial positions (the gold atoms being much too large), it follows theoretically that pure interstitial migration will not affect ordering. Vacancies can jump between Cu and Au sites and affect the ordering considerably. Therefore the latter must be regarded as the diffusing agent in step IV.

ROSWELL and NOWICK [1957] have investigated the effects of quenching on the dislocation internal friction of gold (see Chapter XVII). This is suppressed by quenching as had also been found by LEVY and METZGER [1955] in aluminium; the damping returns after annealing at the fourth step (activation energy between 1 and 1.4 eV). As the decrease of damping is certainly due to vacancies pinning the dislocations, this observation thus strongly points to vacancy diffusion as the cause of the fourth step, although the authors themselves attributed it to "dislocation diffusion" of the defects (cf. Chapter XVIII). Similar experiences on the return of damping of neutron-irradiated metals have resulted in the same conclusion, viz. that recovery only takes place at

temperatures above the room temperature, that is, not in stage III but in stage IV.

If indeed the fourth step should be associated with vacancy diffusion, the question remains how to explain the occurrence of step III after quenching and other characteristics of quenching experiments, and the shallowness of stage IV in irradiated samples. We have already pointed out that the inference drawn from quenching experiments must be considered with some reserve. As discussed previously, the vacancies present as such at the quenching temperature recombine for a large part during the quench to divacancies and these diffuse much more easily than single vacancies. It might therefore very well be that the observed stage III recovery in quenched materials is due to *divacancy* diffusion. As the activation energy for divacancy diffusion would theoretically be expected to amount to about half of that of vacancies (BARTLETT and DIENES [1953]), the observed value of 0.6 eV could in principle be understood in this way. Seeger, however, on the basis of his calculations of the migration energy of interstitials, definitely assigned this process to stage III, whereas HASIGUTI [1960] explained it as due to the release of interstitials from deep impurity traps—shallow traps having already been vacated during stage II—and subsequent elimination at dislocations. It will be clear that there are many difficulties associated with the explanation of this at first sight so simple matter of the recovery of the defect resistivity!

The other school of thought, according to which the third step is vacancy migration, cannot explain the fourth step at all. They consider this recovery step as spurious, and regard the recovery of additional resistivity above room temperature as characterized by a continuous range of activation energies. The problem can only be solved ultimately by a careful study of the *kinetics* of the various recovery stages, which is now being carried out in several places. A discussion of this aspect, however interesting, would carry us to far. A very extensive discussion on the relative merits of the various recovery models has been published recently (BRINKMAN *et al.* [1959]). It contains the proceedings of a meeting especially arranged for this topic. LOMER and COTTRELL [1955] have proposed still another interpretation of the stages III and IV, namely in terms of the diffusion of vacancies to traps (III), followed by the release of the vacancies from the traps (c.q. the migration of the traps) and the subsequent migration to dislocations (IV). We refer to the original paper for a discussion.

The fifth recovery step gives no difficulties to explain. Its activation energy and other characteristics point unambiguously to the already described "mechanical" recovery as the basic process (Chapter XI) (eventually also recrystallization). This is primarily associated with self-diffusion and dislocation climb in the metal. All remaining additional resistivity and stored energy is released in this step.

In Table 6 some of the proposed interpretations of the recovery phenomenon have been compiled.

TABLE 6

Diffusion processes underlying recovery stages

Model	*Stage I*	*II*	*III*	*IV*	*V*
A (this book)	close-pair annihilation	interstitial migration, release from shallow traps	divacancy migration	vacancy migration	self-diffusion
B (BLEWITT, KOEHLER)	crowdion-diffusion	interstitial-vacancy recombination	vacancy migration	simultaneous occurrence of several processes	
C (SEEGER)	close-pair annihilation, rearrangement in spikes, etc. (according to treatment)	1. "wide" pair recombination 2. divacancy migration	interstitial migration	vacancy migration	self-diffusion
D (CORBETT and WALKER	close-pair annihilation	interstit. migration	vacancy migration		self-diffusion
E (HASIGUTI)	close-pair annihilation	release of int. from shallow traps	release of int. from deep traps	vacancy migration	self diffusion

6. Effects of Fast-Particle Bombardment and of Quenching on the Mechanical Properties of Metals

Fast-particle irradiation, as well as quenching produces several typical changes of the properties of metals which, although they are due

to the production of lattice defects, do not occur in this form after cold-work. We mention first the high rate of *disordering* of ordered alloys by neutron irradiation. This effect does not occur during electron irradiation, and has consequently been attributed by many authors to the influence of thermal and displacement spikes, instead of to the influence of interstitialcy migration as discussed in section 4. Conversely, a disordered alloy in the course of ordering may actually show *accelerated ordering* during irradiation, presumably owing to the enhanced defect concentration. Such an acceleration has already been observed to take place under weak γ-irradiation in Cu_3Au. A well studied alloy in this respect is α-brass (30% Zn) (DAMASK [1956, 1957]) which shows equilibrium *short-range* order to a degree dependent on the temperature (ROSENBLATT *et al.* [1955]). By irradiation with neutrons slightly above room temperature the electrical resistivity at low temperatures is *decreased*, presumably because the degree of order is enhanced. This enhancement is caused by the creation of vacancies which wander through the lattice and enable Cu and Zn atoms to acquire a configuration of lower free energy, that is, of a higher degree of short-range order – each Cu-atom tending to surround itself with Zn neighbours only and vice versa.

By comparing the influence of electron and neutron irradiation, DAMASK [1957] concluded that a flux of 2.6×10^{14} cm^{-2} sec^{-1} of 2 MeV incident electrons was equivalent to a flux of about 10^{12} cm^{-2} sec^{-1} of fast pile neutrons. Inserting the known displacement cross-sections into the Seitz-Koehler displacement theory (Chapter XIV) one finds that, according to theory, the electrons should be about 14 times more efficient than the neutrons. Damask concluded from this that the majority of the defects caused by electron irradiation occur in close vacancy-interstitial pairs that recombine rapidly already during the treatment and are lost to the ordering process. This is in agreement with the very pronounced occurrence of stage I annealing after electron irradiation, as discussed in the previous section.

It should be noted here that an increase of short-range order need not necessarily cause a decrease of resistivity such as in α CuZn; in Cu_3Au the resistivity *in*creases with increasing short-range order.

A curious phenomenon observed in some anisotropic and fissionable metals or alloys is *radiation growth*. Uranium and some of its alloys show the phenomenon conspicuously; the various grains deform anisotropically during irradiation, and when a texture is present this results

in macroscopic growth (for instance of the fuel elements in a reactor; see HOLDEN [1958]). In fine-grained random polycrystals the overall effect is zero. A theory of radiation growth in uranium has been developed by PUGH [1955] in terms of the anisotropic diffusion of the interstitials and vacancies created during irradiation. The two kinds of defects would move preferentially in *different* directions in the highly anisotropic uranium lattice, for a reason which is not clear. The result is a "corrugated layer" structure of the bombarded U-lattice.

Radiation growth in uranium only occurs at moderate temperatures, where the defects cannot anneal out. Above 450 °C a new phenomenon takes place, that of *swelling*, which is due to the formation of *gaspockets* filled with *helium* or *krypton*, disintegration products which are practically insoluble in the metal. When the temperature is high enough for diffusion to occur, these gases collect in voids. Especially where dislocations or grainboundaries are present in sufficient concentration to provide the vacancies needed for the voids, large quantities of helium can be stored, as has been illustrated in a very elegant manner by COTTRELL *et al.* [1958].

In fissionable materials, owing to the high energy of the fission particles, the number of displacements is extremely high (for instance 25000 per split U-nucleus): 0.1% of the atoms are knocked from their places after only 3 months irradiation in a reactor of common design; therefore, radiation growth is technically important. The amount of energy stored in the form of defects is very large, and the annealing out of it can give rise of temperature, leading in turn to accelerated annealing. (A similar, even more extreme, situation exists as is well known in graphite: Wigner-effect.)

The local stresses around gas bubbles or heavily damaged sites (spikes) may surpass the flow stress. Then the material begins to flow spontaneously, that is, it behaves like a Newtonian viscous fluid (ROBERTS and COTTRELL [1956]). For instance, a bar of uranium irradiated with neutrons begins to creep under its own weight at a rate of about 10^{-11} sec^{-1} (depending on the geometry of the sample). Characteristic for this kind of creep is that it is not temperature activated, and occurs even at low temperatures ($\leqslant$ 100 °C).

The specific (mechanical) properties of metals, such as creep rate, critical shear stress, work-hardening characteristics, hardness and brittleness all change under the influence of irradiation. Especially reactor irradiation or bombardment with heavy nucleons is very effec-

tive in this. Roughly speaking, all changes are in the direction of increasing hardness. The creep rate under a given stress usually decreases, although the very first observations (on α-irradiated cadmium) showed it to increase and incidentally similar reports have been published. Owing to the small penetration depth of α-particles this concerns only a surface effect, however.

The effect of irradiation on technical materials is especially marked in zirconium (a much used reactor material) where a decrease of the creep rate by a factor of the order 10 was observed after an integrated flux of only 10^{18} neutrons/cm^2. The increase of the flow stress, yield stress, fracture strength, hardness, brittleness etc. of many metals by irradiation are usually denoted by the term *radiation hardening*. Irradiation with a total flux of 10^{18} fast neutrons per cm^2 at very low temperatures increases the critical shear stress of copper single crystals by a factor of about 10 (BLEWITT and COLTMAN [1951]). The critical shear stress moreover becomes strongly temperature dependent. The effect is largest in well-annealed single crystals, and small in already cold-worked metals. The increase usually reaches a saturation after an integrated flux of about 10^{20} neutrons per cm^2. Not only the critical shear stress, but also the shape of the stress-strain curve alters; the work-hardening coefficient of irradiated copper crystals decreases somewhat by irradiation; the easy glide region extends further, slip lines become more clustered (coarse) and better visible, especially in the early portions of the stress-strain curve (Fig. 6). In polycrystalline molybdenum the ductile-to-brittle transition (see Chapter XXII) was raised from $-136°$ to -73 °C after about 10^{19} pile neutrons/cm^2 (MAKIN and GILLIES [1957]). A similar, but smaller rise was observed in deuteron-irradiated mild steel, zirconium, tungsten etc. The hardness of such metals also increased considerably. For a critical review of these and other data and references we refer to the books by DIENES and VINEYARD [1957] and by HOLDEN [1958], and to a recent paper by SEEGER [1958b].

Many of the mechanical changes indicate a progressive *anchoring* of the dislocations during the irradiation. Sometimes annealing is needed to produce this anchoring in full effect (e.g. in irradiated molybdenum heating to 200 °C increases the yield drop (MAKIN and GILLIES [1957]). But also the *mobility* of the dislocations is decreased, especially so in the face-centred cubic metals. In aluminium, radiation hardening already develops to its full extent at helium temperatures, and anneal-

ing between −80 and −20 °C eliminates most of it. In copper, annealing below room temperature (steps I to III) has no or little effect, the radiation hardening disappears only in the Vth stage.

The theory of radiation hardening is intimately tied up with the theory of the nature of the imperfections caused in a crystal by irradiation. The anchoring of dislocations could well be due to point defects (vacancies, interstitials, or may be crowdions), which cause jogs or other irregularities at the dislocations. The decrease of the dislocation mobility observed in face-centred cubic metals cannot be understood in this way, however. An explanation in terms of a special kind of "precipitation" hardening caused by the existence of defect clusters seems more appropriate. SEEGER [1958b] regards as the cause of the hardness the formation of empty zones at the end of the path of incident neutrons ("depletion zones") that can be formed (already at very low temperatures) by the vacancies left over after the interstitials have been dispersed as *dynamic* crowdions (see the preceding chapter, section 2.2.6). These clusters are stable against annealing up to the temperature where self-diffusion begins to take place, and the hardness is therefore also temperature independent below this temperature range. Indirect evidence for the presence of defect clusters after irradiation or quenching can be obtained by small-angle X-ray scattering.

A different approach to the subject has been given by COULOMB and FRIEDEL [1957], who argued that conglomerates of several thousands of vacancies could perhaps form *along the dislocations* (see also Chapter III); these cavities do not produce a temperature independent type of hardening, however, but result in a flow stress inversely proportional to T.

Various interesting features of radiation hardening and its recovery behaviour have been observed by BLEWITT and his coworkers [1957], and a review of data concerning the mechanical properties of various types of irradiated solids has been prepared by LEVI [1959]. We refer to the original papers for particulars.

A type of hardening resembling very much that due to radiation has been observed after *quenching* and subsequent medium temperature annealing (LEVY and METZGER [1955], MADDIN and COTTRELL [1955], ROSWELL and NOWICK [1957], KUHLMANN-WILSDORF and MADDIN [1960]). It is certain that vacancies or their agglomeration products cause the hardening, either by forming jogs or by a cavity mechanism. Again yielding, coarse slip lines, etc. have been observed. (It has already

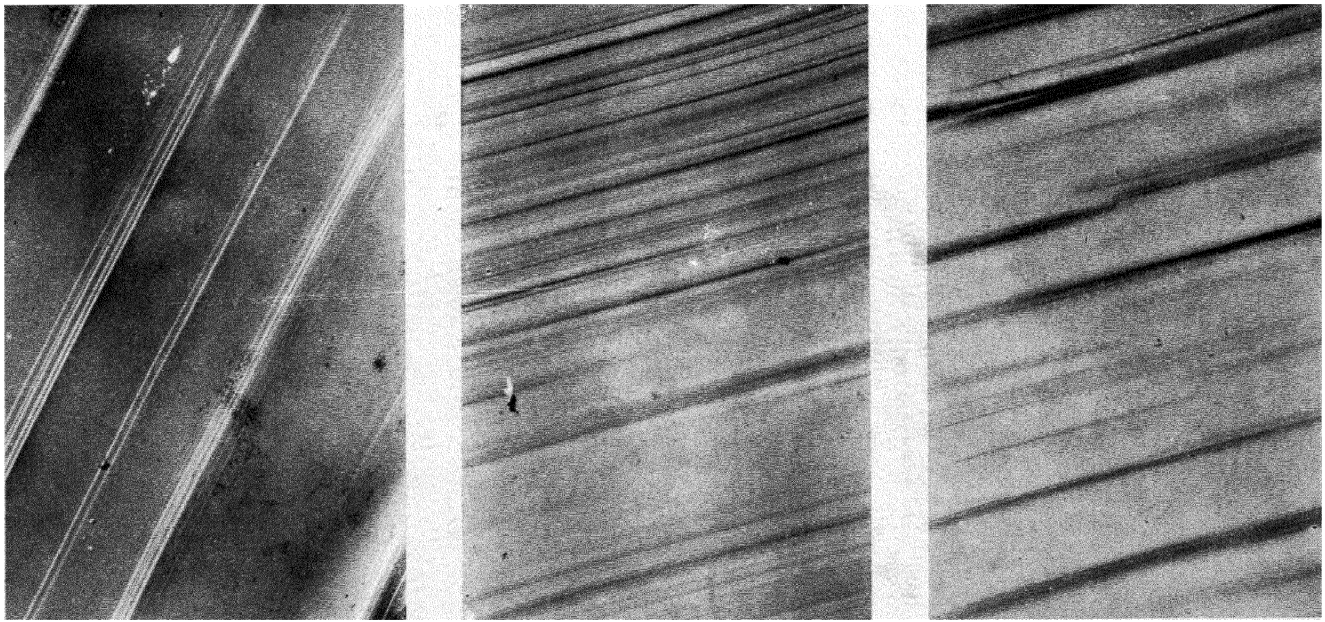


Fig. 6. Surface pattern (5000×) of neutron-irradiated and deformed Cu crystals (total dose 6×10^{17} neutrons/cm^2). These photographs should be compared to Figs. VI–5 to 8. *a*) In the easy glide range (0.5% strain) great differences between irradiated and non-irradiated crystals can be detected: the slip is much less homogeneous in the first case. *b*) and *c*) As the deformation proceeds these differences gradually become less apparent. I am grateful to Dr S. Mader for allowing me to use these photographs.

been mentioned that by irradiation as well as by quenching also the dislocation *damping* (Chapter XVII) can be fully suppressed.) According to the temperature from which was quenched, the hardening is different. The explanation follows the lines discussed in section XIV-3.

The role of dislocations as sources and sinks for vacancies has been considered in great detail in the last mentioned report. Especially the "new" dislocations, that is those that have moved not too long ago during or since annealing treatments, can easily interact with vacancies, viz. by *climbing*. The interaction can lead, when *deformation* follows the irradiation or quenching, to the collection of very many vacancies by a slow moving dislocation; collapse takes place ultimately and prismatic loops are formed which partially merge with the glide dislocations, and thereby cause so-called *tangled* structures, typical for this sort of treatment.

Irradiation has a pronounced effect on diffusion and phase-transformation phenomena, owing to the creation of a non-equilibrium number of lattice defects. We shall discuss some aspects of this very interesting and as yet only superficially explored field in Chapters XVIII and XXI.

REFERENCES

AARTS, W. H. and R. K. JARVIS (1954); Acta Metall. **2,** 87.

AIROLDI, G., G. L. BACHELLA and E. GERMAGNOLI (1959); Phys. Rev. Lett. **2,** 145.

BACHELLA, G. L., E. GERMAGNOLI and S. GRANATA (1959); J. appl. Phys. **30,** 748.

BARTLETT, J. H. and G. J. DIENES (1953); Phys. Rev. **89,** 848.

BAUERLE, J. E., F. SEITZ and J. S. KOEHLER (1957); Phys. Rev. **107,** 1493 and 1499.

BERGHOUT, C. (1957); Thesis Delft Technical University; (1958) Acta Metall. **6,** 613.

BLATT, F. J. (1955); Phys. Rev. **99,** 1708.

BLEWITT, T. H. and R. R. COLTMAN (1951); Phys. Rev. **82,** 769.

BLEWITT, T. H., R. R. COLTMAN, P. K. HOLMES and T. S. NOGGLE (1957); Dislocations and the Mechanical Properties of Crystals (Wiley, New York) p. **603.**

BLEWITT, T. H., R. R. COLTMAN and J. K. REDMAN (1954); Report Conference on Defects in Solids, Bristol, p. **369** (The Physical Society, London).

BOAS, W. (1957); Dislocations and the Mechanical Properties of Crystals (Wiley, New York) p. **333.**

BRADSHAW, F. J. and S. PEARSON (1956); Phil. Mag. **1,** 812 and 880; (1957a) Phil. Mag. **2,** 379; (1957b) Phil. Mag. **2,** 570; (1957c) Phil. Mag. **2,** 1387.

BRINKMAN, J. A., C. F. DIXON and C. J. MEECHAN (1953); Acta Metall. **2,** 38.

BRINKMAN, J., A. SOSIN and C. J. MEECHAN (1960); Atomics International Discussion Meeting, Report on Radiation Damage.

BROOM, T. (1954); Advances in Physics **3,** 26.

BROOM, T. and W. K. CLOTHIER (1952); Austr. J. Sci. Res. **A5,** 128.

BROSS, H. (1959); Z. Naturf. **14a,** 560.

BUEREN, H. G. VAN (1953); Acta Metall. **1,** 464 and 607.

BUEREN, H. G. VAN (1955); Acta Metall. **3**, 519.
BUEREN, H. G. VAN (1956); Colloquium on Deformation and Flow of Solids, Madrid, p. 79 (Springer, Berlin).
BUEREN, H. G. VAN (1957); Philips Res. Rep. **12**, 1 and 90.
CARRUTHERS, P. (1959); Phys. Rev. **114**, 995.
CLAREBROUGH, L. M., M. E. HARGREAVES, D. MICHELL and G. W. WEST (1952); Proc. Roy. Soc. **A215**, 507; (1955) Proc. Roy. Soc. **A232**, 252; (1956) Phil. Mag. **1**, 528.
COOPER, H. G., J. S. KOEHLER and J. W. MARX (1954); Phys. Rev. **94**, 496.
COOPER, H. G., J. S. KOEHLER and J. W. MARX (1955); Phys. Rev. **97**, 599.
CORBETT, J. M., J. M. DENNEY, M. D. FISKE and R. M. WALKER (1957); Phys. Rev. **108**, 954.
CORBETT, J. M., R. SMITH and R. M. WALKER (1959); Phys. Rev. **114**, 1450.
CORBETT, J. M. and R. M. WALKER (1958); Phys. Rev. **110**, 767.
COTTRELL, A. H. (1956); Brit. J. applied Physics, Suppl. **5**, 43.
COTTRELL, A. H. (1957); Dislocations and the Mechanical Properties of Crystals (Wiley, New York) p. 577.
COTTRELL, A. H. (1958); Symp. on Vacancies and other Point Defects (Inst. Metals, London) p. 1.
COTTRELL, A. H., G. B. REDDING and R. S. BARNES (1958); Phil. Mag. **3**, 97.
COULOMB, P. and J. FRIEDEL (1957); Dislocations and the Mechanical Properties of Crystals (Wiley, New York) p. 555.
DAMASK, A. C. (1956); J. applied Phys. **27**, 610; (1957) J. Phys. Chem. Solids **4**, 177.
DIENES, G. J. and G. H. VINEYARD (1957); Radiation Effects in Solids (Interscience, New York).
DRUYVESTEYN, M. J. and J. A. MANINTVELD (1951); Nature **168**, 868.
DRUYVESTEYN, M. J. and J. A. MANINTVELD (1952); Nature **169**, 632.
DUGDALE, R. A. (1954); Report Conference on Defects in Solids, Bristol, p. 246 (London, The Physical Society).
EGGEN, D. T. and M. J. LAUBENSTEIN (1953); Phys. Rev. **91**, 238.
FUMI, F. G. (1955); Phil. Mag. **46**, 1007.
GERRITSEN, A. N. (1956); Encyclopedia of Physics Vol. XIX, p. 137.
GLEN, J. W. (1955); Advances in Physics **4**, 381.
HASIGUTI, R. R. (1960); J. Phys. Soc. Japan **15**, 1807.
HOLDEN, A. N. (1958); Physical Metallurgy of Uranium (Addison-Wesley Publ., New York).
HUNTINGTON, H. B. (1941); Phys. Rev. **61**, 335.
HUNTINGTON, H. B. (1953); Phys. Rev. **91**, 1092.
HUNTINGTON, H. B. and F. SEITZ (1941); Phys. Rev. **61**, 326.
Institute of Metals Seminar on Vacancies and Other Point Defects (London 1958).
JONGENBURGER, P. (1957); Phys. Rev. **106**, 66.
JONGENBURGER, P. (1957a); Conf. Phys. Basses Temp. Paris (C.I.F.).
JONGENBURGER, P. (1960); to be published.
KAUFFMAN, J. W. and J. S. KOEHLER (1952); Phys. Rev. **88**, 149.
KAUFFMAN, J. W. and J. S. KOEHLER (1955); Phys. Rev. **97**, 255.
KEMP, W. R. G., P. G. KLEMENS and R. J. TAINSH (1958); Physica **24**, Supplement 5, 170.
KEMP, W. R. G., P. G. KLEMENS, R. J. TAINSH and G. K. WHITE (1957); Acta Metall. **5**, 303.
KINCHIN, G. H. and R. S. PEASE (1955); J. Nuclear Energy **1**, 200.
KLEMENS, P. G. (1956); Encyclopedia of Physics, Part XIV, page 198 (Springer, Berlin).
KLEMENS, P. G. (1958); Solid State Physics **7**, 1.
KROPSCHOTT, R., H. M. GARBER and F. J. BLATT (1959); Phys. Rev. Lett. **2**, 91.

KUHLMANN-WILSDORF, D. and R. MADDIN (1960); ASM Seminar on Strengthening Mechanisms in Solids, Philadelphia.
LAZAREV, B. G. and O. N. OVCHARENKO (1955); Doklady Akad. Nauk SSSR **100,** 875.
LEVI, F. A. (1959); Suppl. al Nuovo Cimento **12.**
LEVY, M. and M. METZGER (1955); Phil. Mag. **46,** 1021.
LOMER, W. M. and A. H. COTTRELL (1955); Phil. Mag. **46,** 711.
MADDIN, R. and A. H. COTTRELL (1955); Phil. Mag. **46,** 735.
MADDIN, R., D. KUHLMANN-WILSDORF and H. KIMURA (1958); Acta Metall. **7,** 145.
MAGNUSSON, G. D. and W. PALMER (1958); Phys. Rev. **109,** 1990.
MAKIN, M. J. and F. GILLIES (1957); J. Inst. Metals **86,** 10.
MANINTVELD, J. (1954); Thesis Delft Technical University.
MARTIN, A. B., S. B. AUSTERMAN, R. R. EGGLESTON, J. F. McGEE and M. TARPINIAN (1951); Phys. Rev. **81,** 664.
McREYNOLDS, A. W., W. AUGUSTYNIAK, M. McKEOWN and D. B. ROSENBLATT (1955); Phys. Rev. **98,** 418.
MEECHAN, C. J. and J. BRINKMAN (1960); Atomics International Report on Rad. Dam.
MEECHAN, C. J. and R. R. EGGLESTON (1954); Acta Metall. **2,** 680.
MEECHAN, C. J. and A. SOSIN (1959); Phys. Rev. **113,** 422.
MENDELSSOHN, K. (1958); Physica **24,** Supplement 5, 43.
MOLENAAR, J. and W. H. AARTS (1950); Nature **166,** 690.
OVERHAUSER, A. W. (1953); Phys. Rev. **90, 393; 92,** 530.
OVERHAUSER, A. W. and R. L. GORMAN (1956); Phys. Rev. **102,** 676.
PANSERI, C. and T. FEDERIGHI (1958); Phil. Mag. **3,** 1223.
PEARLSTEIN, E. A., H. INGHAM and R. SMOLUCHOWSKI (1955); Phys. Rev. **98,** 1530.
PRY, R. H. and R. W. HENNIG (1954); Acta Metall. **2,** 318.
PUGH, S. F. (1955); Proc. Int. Conf. Geneva I, Vol. **7,** p. 441.
RANDOLPH, B. W. (1952); Phys. Rev. **85,** 710.
ROBERTS, A. C. and A. H. COTTRELL (1956); Phil. Mag. **1,** 711.
ROSENBERG, H. M. and J. N. LOMER (1959); Phil. Mag. **4,** 467.
ROSENBLATT, F., R. SMOLUCHOWSKI and G. J. DIENES (1955); J. applied Physics **26,** 1044.
ROSWELL, A. E. and A. S. NOWICK (1957); Acta Metall. **5,** 228.
SEEGER, A. (1955); Encyclopedia of Physics, Part VII–1, p. 383 (Springer, Berlin).
SEEGER, A. (1956); unpublished.
SEEGER, A. (1958a); Conference on Radiation Damage, Aachen (unpublished).
SEEGER, A. (1958b); Proc. Int. Conf. Geneva II, paper 961.
SEEGER, A. (1960); in the press.
SEEGER, A. and P. HAASEN (1958); Phil. Mag. **3,** 470.
SEEGER, A. and H. STEHLE (1956); Z. Physik **146,** 217.
SEEGER, A. and E. MANN (1960); J. Phys. Chem. Solids **12,** 326.
SEITZ, F. and J. S. KOEHLER (1956); Solid State Physics **2,** 307 (Academic Press, New York).
SIMMONS, R. O. and R. W. BALLUFFI (1958); Phys. Rev. **109,** 335.
SORBO, W. DE and D. TURNBULL (1959); Acta Metall. **7,** 83.
SOSIN, A. and J. S. KOEHLER (1956); Phys. Rev. **101,** 972.
SOSIN, A. and L. TEWORDT, (1958); Phys. Rev. **109,** 612.
TEWORDT, L. (1958); Phys. Rev. **109,** 61.
TITCHENER, A. L. and M. B. BEVER (1958); Progress in Metal Physics **7,** 247.
TUCKER, C. W. and J. B. SAMPSON (1954); Acta Metall. **2,** 433.
VOOK, R. W. and C. A. WERT (1958); Phys. Rev. **109,** 1529.
WALKER, R. M. and J. M. CORBETT (1959); Atomics International Report, to be published.
WECHSLER, M. S. (1957); Acta Metall. **5,** 150.
WINTENBERGER, M. (1957); C. R. Acad. Paris **244,** 2800.

CHAPTER XVI

INFLUENCE OF LATTICE IMPERFECTIONS ON PHYSICAL PROPERTIES

NUCLEAR SPIN RESONANCE;
X-RAY AND NEUTRON DIFFRACTION, MAGNETIC PROPERTIES

Apart from the electrical resistivity, many other physical properties are markedly influenced by lattice imperfections. These properties have been called *structure sensitive* properties. In this chapter we shall only deal with such changes as are of interest for the study of the imperfections themselves. We shall treat briefly the nuclear spin resonance of imperfect crystals and the X-ray and neutron diffractive properties, not only of metals but also of other crystalline solids where they are of interest. Finally we shall mention some aspects of the influence of dislocations on ferromagnetism.

1. Nuclear Spin Resonance in Imperfect Crystals

Evidence on the presence and distribution of lattice imperfections, especially of dislocations, can be obtained by the use of the nuclear spin resonance method as this has been developed during the last 15 years. In this method (Fig. 1) a sample of the material is placed in a radiofrequency (rf) magnetic field (frequency ω) which is applied at right angles to a strong magnetic field H. A nucleus with spin quantum number I has $2I + 1$ distinct spin states, which coincide in field zero; they are split up in the static field in as many equidistant energy levels (Fig. 2*a*). These are occupied in equilibrium according to a Boltzmann distribution. The radiofrequency field can disturb this equilibrium by inducing *transitions* between the various levels, namely between those that differ by ± 1 in m, where m is the magnetic quantum number denoting the orientation of the spin with respect to the field H. This restriction arises from the operation of one of the well-known

selection rules. Transitions occur only if the frequency of the radio-frequency field coincides more or less with the energy difference (divided

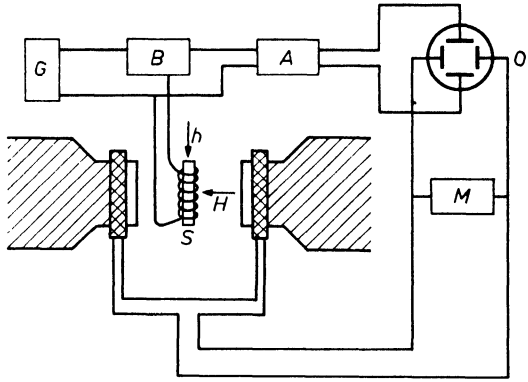


Fig. 1. Schematic diagram of the nuclear resonance method (after BLOEMBERGEN *et al.* [1948]). *G* rf signal generator, *S* sample mounted in coil in temperature bath. The rf magnetic field h is vertical in this set up. *B* bridge, *A* rf amplifier *O* oscilloscope. The horizontal static magnetic field H is "slowly" modulated by an extra set of windings on the permanent magnet. The current through these is varied by the modulator *M*, which is connected to the time base of the oscilloscope.

by Planck's constant) of the spin levels between which the transition has taken place, thus if

$$\omega \approx \omega_0 = \frac{\Delta E}{\hbar} = \frac{1}{\hbar}\frac{\mu H}{I}, \tag{1}$$

where μ is the magnetic moment of the nucleus. The quantity μH is simply the energy difference between two levels with $|\Delta m| = 1$ in a field H. This situation is called *resonance*, and the maximum absorption for $\omega = \omega_0$ is called the resonance absorption.

That it is indeed absorption that takes place and not emission, is clear when it is realized that before switching on the rf field the lower levels were (in equilibrium) more densely occupied than the upper levels. The application of an rf field of the right frequency results in a population of the upper levels that is enhanced over the equilibrium value. The shape of the absorption line (energy absorbed as a function of ω) is determined primarily by the *damping* caused by magnetic interaction between the nuclei mutually and by the interaction of the spin system with the crystal lattice. The latter interaction is usually

described by a relaxation time called T_1, viz. the time it takes for a given disturbance of the population of the spin levels to decay to the thermal equilibrium population.

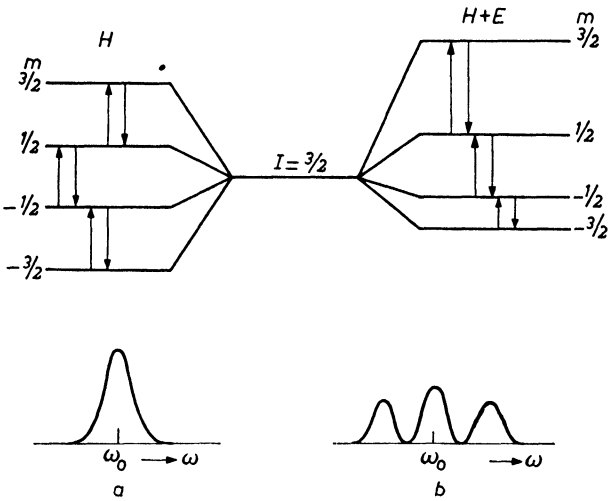


Fig. 2. *a*) Splitting of nuclear spin levels of a nucleus with spin $I = {}^3/_2$ into equidistant sublevels $-I \leqslant m \leqslant I$, in a static magnetic field H. A transverse rf field can induce transitions $\Delta m = \pm 1$, all of which correspond to the same absorption or emission line at $\omega = \omega_0$. *b*) If an electric field gradient acts on a quadrupole moment (for $I > \frac{1}{2}$), the line splitting becomes non-equidistant, each line being shifted proportional to m^2. Several absorption lines are now observed (3 in the example, one of which lies again at $\omega = \omega_0$, as the levels $m = \frac{1}{2}$ and $m = -\frac{1}{2}$ have been shifted by the same amount).

In some crystals a static magnetic field is not needed to split up the (degenerate) spin levels of a given nucleus. Such a splitting is already present naturally, owing to an *electric quadrupole interaction* that exists between some of the composing nuclei and the intracrystalline electric field (which is for instance due to ionic charges). Such an interaction can occur only if the nucleus to be studied has a spin quantum number greater than $\frac{1}{2}$ (otherwise its electric quadrupole moment is zero), and if a sufficient strong *electric field gradient* is present at the site of the nucleus. In perfect cubic crystals quadrupole splitting can thus never occur, because no gradient exists in the lattice points.

When a moderately strong quadrupole interaction is present and a magnetic field is applied along the z-axis, the $(2I + 1)$ spin levels are no longer equally spaced in energy by the amount (1), but their posi-

tions are now given by:

$$E_m = m\frac{\mu H}{I} - \frac{eQ}{4I}\{3m^2 - I(I+1)\}\frac{\partial F}{\partial z}, \qquad (2)$$

where $\partial F/\partial z$ is the gradient of the electrostatic field at the site of the nucleus and Q is the electric quadrupole moment of that nucleus (Fig. 2*b*). The spacing between the m^{th} and $(m-1)^{\text{th}}$ level is no longer $\hbar\omega_0$, but, when the axial electric field inserts an angle θ with the z-axis:

$$\Delta E(m, m-1) = \hbar\omega_0 + (2m-1)(3\cos^2\theta - 1)\frac{3eQ}{8I(2I-1)}\frac{\partial F}{\partial z}. \qquad (3)$$

There are now $2I$ different line spacings instead of only one, and the resonance line is split up into as many components. In the particular case that $I = {}^3/_2$, the $m = \frac{1}{2} \rightarrow -\frac{1}{2}$ transition keeps its original frequency ω_0, but two *satellite* lines appear at both sides. It is to be repeated that such a quadrupole splitting occurs only in a sufficiently unsymmetrical crystal electric field.

The effect of lattice imperfections is now that, even in highly symmetric crystals, they cause local departures from symmetry and thereby finite $\partial F/\partial z$ values. Nuclear resonance absorption lines in cubic crystals which theoretically have to be single and sharp, are influenced by the presence of lattice imperfections in that these produce local field gradients and cause slight but measurable "satellite broadening". Especially dislocations are effective in this, as they have long range strains around them and therefore influence the electric field gradient at many more nuclei than for instance point defects can do.

The effect of dislocations on nuclear spin resonance has been demonstrated convincingly by WATKINS and POUND [1953], who found the resonance lines of Br^{79} and Br^{81} (both with spin ${}^3/_2$) in KBr and I^{127} (spin ${}^5/_2$) in KI crystals to be considerably *weaker* (about 40%) than was expected from the known spin values. Watkins explained this by assuming that the ever present dislocations had caused a quadrupole splitting of the line. The satellites were smeared out so much (since various values of θ occur at different sites in the crystal) that they were no longer observable, and only the undisplaced $\frac{1}{2} \rightarrow -\frac{1}{2}$ transition could still be measured. This alone, however, represents only 30–40% of the total absorbed energy.

Plastic compression of a KBr crystal by 20% resulted in a measure-

able broadening of the resonance line, and finally even in an unsymmetrical line shape; all in accordance with the theory that predicts larger quadrupole effects when the dislocation density increases. Numerical agreement between theory and experiment could only be obtained by assuming that the actual field gradient at the site of the bromine nuclei is about 10 times greater than that which is computed from the displacements of the neighbouring ions alone. This can be explained only by also taking into account the effect of the lattice strains on the *electronic wave functions* of the nucleus. The known large quadrupole interaction between halogen ions in covalent bonds is an indication of the possibility that the electron cloud around a halogen nucleus can indeed deform easily under the influence of an inhomogeneous electric field, and at the site of the nucleus, the field gradient may therefore well be enhanced by a factor of the order of 10. Assuming this to be the case, the dislocation density responsible for the lattice strain can now be estimated. From the observed broadening Watkins deduced a density of 10^{10} cm^{-2} in the deformed crystals, and 6×10^8 cm^{-2} in the as-received crystals. The latter value seems fairly high. It was confirmed, however, by OTSUKA and KAWAMURA [1957], who also pointed out that from a close analysis of the asymmetrically broadened lines it can even be found whether edge or screw dislocations are present preponderantly.

In metals quadrupole effects have also been observed. We mention the work of BLOEMBERGEN and ROWLAND [1953] (see also BLOEMBERGEN [1954]) on copper and brass. In the latter alloy no copper (spin $^3/_2$) resonance could be observed at all; the absorbed intensity in the relevant frequency region was more than 100 times smaller than in pure copper. The absence was explained in terms of the non-cubic symmetry of the surroundings of a copper nucleus, owing to the presence of zinc atoms. Again the satellite lines arising from the field gradient are smeared out by reason of the range of different orientation angles; in this case even the $\frac{1}{2} \rightarrow -\frac{1}{2}$ transition is broadened so much as to render it non-observable. Why this should be so is not clear, however. Even if the Zn-atoms are considered as doubly charged (which is certainly not the case), the field gradient at a copper nucleus comes out 10 times smaller than necessary to account for the vanishing of the absorption. A nuclear enhancement factor of about 50 must be assumed to explain the observations. This should again be due to the polarizability of the electronic wave functions.

Cold-worked copper filings showed 50–60% less resonance absorption than well annealed pure copper. Using the same large enhancement factor, the dislocation density is found to be about 10^{11} cm^{-2} in the cold-worked material and less than 10^{7} cm^{-2} in well-annealed copper which shows no broadening at all. In aluminium no decrease in absorbed intensity after cold-work has been observed. No explanation for this has been advanced.

Although as yet no satisfactory agreement exists between experiment and theory, the study of quadrupole effects on nuclear spin resonance can in principle give interesting new information on lattice strains and dislocations. Those substances should preferentially be studied in which the electronic shells of the resonating nuclei are not too sensitive to deformation, so that the satellite lines do not disappear entirely when dislocations are present.

There exists another, less direct influence of lattice defects on the resonance of nuclear spins. It is associated with point defects and has nothing to do with quadrupole effects, but depends on simple magnetic dipole interaction only. Vacancies present in a metal enhance the rate of diffusion, and the nuclei (which are responsible for the resonance absorption) jump more frequently in a metal that contains many vacancies than in metal with few vacancies. The internal magnetic field acting on a nucleus will thus fluctuate with a mean frequency which is approximately given by

$$\nu_{\text{diff}} = \nu_0 n_0 \exp\left(-U_m/kT\right), \tag{4}$$

where U_m is the activation energy for migration and n_0 the concentration of vacancies. When this frequency exceeds the line width in the perfect lattice, the line will become *narrower*, as BLOEMBERGEN *et al.* [1948] have pointed out. This effect is called *motional narrowing* and is essentially the result of averaging over orientation factors such as occurring in (3). We refer to the paper cited for a theoretical discussion. It will be clear that by studying the line width as a function of temperature, the activation energy for vacancy diffusion can now be determined, or rather, if the vacancy concentration is in thermal equilibrium, the activation energy for self-diffusion. It is even possible to distinguish between activation energies of different isotopes in this way.

Measurements have been reported by SLICHTER and others [1956]

for the alkali metals and aluminium, and for molecular solids by ANDREW [1953].

A third effect on nuclear resonance is produced in non-metals by *centres* containing an electron, which then form *paramagnetic* units. These can, by absorption of a nuclear spin quantum, and subsequently losing their energy to the lattice, enhance the spin-lattice relaxation, because the electronic spin-lattice relaxation time is usually much smaller than the nuclear spin-lattice relaxation time T_1. This again results in line broadening.

Not only do paramagnetic centres affect the *nuclear* spin resonance of certain ionic nuclei, they are themselves responsible for *electron spin resonance* (*paramagnetic* resonance), albeit at much higher frequencies. This effect will be discussed in more detail in Chapters XXV to XXVII, dealing with non-metallic solids.

2. X-ray Diffraction by Lattice Defects

2.1. MOSAIC STRUCTURE

We have already mentioned previously the particular X-ray scattering characteristics of single dislocations. It is not possible in practice to deduce from these theoretically the X-ray scattering properties of a crystal containing a distribution of dislocations; this problem is too complex for a systematic theoretical attack. The other way has thus to be followed, i.c. the dislocation distribution in well-annealed and cold-worked crystals must be inferred as far as possible from the observed differences in X-ray scattering properties between them. For this purpose a number of techniques have been developed, which have been recently reviewed by HIRSCH [1957].

Dislocations do not only affect the crystalline diffraction by their own scattering, but they cause local *misorientations* of parts of the lattice, particularly so when they are present in ordered arrays, and such misorientations can be detected quite clearly in ordinary diffraction diagrams (compare e.g. Chapter XI for a particular case). In well-annealed crystals the dislocations still present usually give rise to the so-called *mosaic structure*. Point defects, on the other hand, have far less conspicuous effects on the X-ray diagram, since they have only short range stresses around them. They can, however, show up in a careful study of the lattice parameter.

The intensity of the monochromatic X-ray diffractions in a *large*

perfect crystal should, according to the *dynamical* theory of X-ray diffraction, differ from zero only in narrow angular regions (of the order of a few seconds of arc in width) centred around the predicted Bragg reflection angles. Moreover, the absolute amount of diffracted intensity should be directly proportional to the so-called structure factor F, a quantity that takes account of the crystal structure and of the spatial electronic distribution at the lattice nodes. It has a specific value for each reflection in a crystal. However, it was observed already in the very beginning of X-ray investigations that actually the reflected intensities were up to 100 times stronger than expected, and spread out over a range of minutes of arc instead of seconds. The reflected intensities were further found to be proportional to F^2 rather than to F, except for the strongest reflections (see Chapter V).

Theoretically, an F^2-proportionality must be expected when reflection occurs from a *very small* crystal only, that contains too few atoms for multiple scattering to play a role of importance. It was Darwin's proposal (DARWIN [1922]) that the discrepancy between the observed and theoretical diffraction intensities for a large crystal is due to the splitting up, as it were, of this crystal in a large number of small, slightly misoriented "blocks". According to Darwin, this would be the case even in crystals that seemed fully perfect in all other investigations. Each such small block would scatter according to the F^2-law, and, as the various diffractions do not interfere, also the crystal as a whole would thus scatter proportionally to F^2. Only the very strong reflections, whose width is large enough to comprise part of the range of misorientations of the blocks, still show some signs of dynamical interference, and are thus observed to be somewhat weaker than would follow from the F^2-rule. This effect is called *primary* extinction. Owing to the spread in block orientations the width of a diffraction line of a mosaic crystal is much greater than that of a perfect crystal. The size a crystal must have for the dynamical theory to apply amounts to a few thousands wavelengths. Therefore, the blocks must not be larger than about 1 micron in diameter. Darwin called his hypothetical structure the *mosaic structure.*

Bragg has pointed out that in such a mosaic crystal some additional extinction necessarily occurs, due to the fact that the intensity incident on a particular mosaic block will not be equal to the intensity incident on the crystal surface (in the dynamical theory the intensity is everywhere through the crystal the same (by mutual interference). Those

upper blocks in the path of the X-ray which differ but slightly in orientation from a given block in the interior of a crystal, have already diffracted part of the incident intensity, and, similarly, the outgoing beam diffracted by the block is partly reflected backwards on its way out, when meeting mosaic blocks of suitable setting (Fig. 3). This *secondary* extinction (so named to distinguish it from the primary extinction mentioned above), is most important for strong reflections (like the latter), and it is therefore difficult to separate primary and secondary extinction from each other.

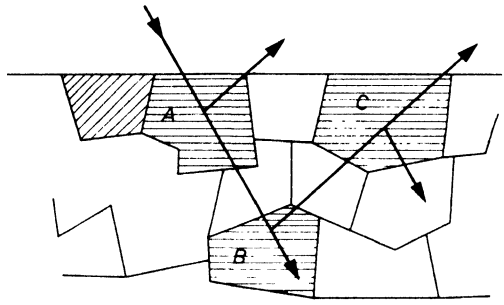


Fig. 3. To illustrate the concept of secondary extinction (after HIRSCH [1957]). The crystal planes in the blocks A, B and C are parallel. A partly shields the incident rays from B, and the rays reflected by B are partly rediffracted by C into the incident direction again. Owing to the random distribution of mosaic blocks, the latter intensity is practically lost.

Several authors have in the past proposed definite models for the mosaic structure (SMEKAL [1933]). The introduction of the concept of dislocations has more or less reconciled the different points of view: the mosaic structure is now regarded as a strongly schematized description of the presence of a spatial distribution of dislocations in the crystal. We shall, nevertheless, occasionally regard the imperfect state as if it were actually composed of slightly misoriented blocks.

The determination of the actual "block size" from intensity measurements rests on the decision whether primary or secondary extinction determines this intensity. In metals it seems that the primary extinction might be negligible compared to the secondary extinction (WILLIAMSON and SMALLMAN [1954]). This would be a happy circumstance, since only the secondary extinction is the result of crystal imperfections. If primary extinction were indeed negligible, an upper limit to the "block size" could be set by observing the *integrated* intensities of

various reflections; this limit is usually found to be about 1–3 microns in metals and somewhat more in exceptionally perfect polar crystals (calcite). If the blocks were larger, the crystals would show much narrower and weaker low-index refractions than are actually observed. The spread in orientation between the "blocks" was, under the same assumption, derived from the observed secondary extinction coefficient; it amounts to a few minutes of arc in aluminium or β-brass. All these deductions should be considered with much reserve, however, as the assumptions are by no means verified.

2.2. STUDIES OF LINE SHAPE AND LINE WIDTH

The most direct X-ray evidence on the mosaic structure has been obtained from the study of so-called *rocking curves*. The crystal to be investigated is rotated through its position of reflection, and the (monochromatic) diffraction peak is investigated with as high a resolving power as possible. In a perfect crystal the reflection is a single narrow peak. When the observed rocking curve has a greater width, or presents more than one maximum, there are certainly misorientations present in the crystal.

This method has been used with some advantage in the study of plastically deformed metal crystals and of crystals annealed after deformation. In the latter the phenomenon of polygonization (Chapter XI) was very clearly observable. This was originally discovered in ordinary Laue diffraction diagrams. The rocking curve of a polygonized crystal consists of discrete, wide peaks, the angles between which can increase to a few minutes of arc.

WARREN and AVERBACH [1952] were the first to subject the line shapes of cold-worked crystals as such to a Fourier analysis. This gives evidence on the distribution of the internal strains. In cold-worked metals the latter are much more important than misorientation effects.

The Laue diffraction technique as used on a single crystal and the phenomenon of *asterism* has already been discussed in Chapter XI. Asterism is essentially a measure for lattice *curvature*, and from the observed peculiar shape of the asterisms in deformed crystals it can be deduced that this curvature arises from dislocations (NYE [1953]). Asterism has also been observed in Debye-Scherrer diagrams of polycrystalline metals; the spots than extend into circular arcs. An important variant of this method has been used recently to determine the

scale of "plastic curvature" in deformed polycrystals more accurately; in this method a very narrow "microbeam" was used. The Debye-Scherrer arcs of deformed metals are then no longer found as continuous rings but they appear broken up into spots ("spotty arcs"). By counting the spots and examining their shape, "particle sizes" and distortions can be deduced. This method has proved powerful in the hands of Hirsch and collaborators (e.g. see GAY *et al.* [1954]).

2.3. INTERPRETATION OF X-RAY DIFRACTION PHENOMENA IN TERMS OF DISLOCATIONS

2.3.1. *Measurable quantities*

The interpretation of the mosaic structure in well-annealed crystals is clearly connected with the concept of crystal boundaries, to be discussed in Chapter XIX. Here already it can be mentioned that the boundary between two crystals with a small orientation difference

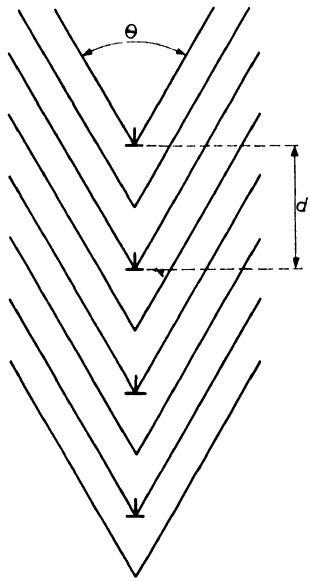


Fig. 4. Tilt boundary considered as a row of edge dislocations.

can be regarded as consisting of a set of dislocations, placed in a regular array. A very simple case of a so-called *tilt* boundary is presented in Fig. 4; it consists of parallel edge dislocations arranged in a plane perpendicular to their glide planes. The two crystals separated by this

boundary are mutually rotated over an angle θ that depends on the spacing d between the dislocations and the Burgers vector b according to

$$\frac{b}{d} = 2 \sin \frac{\theta}{2} \approx \theta . \tag{5}$$

Special types of boundaries, with very small desorientation angles we have already met in Chapter XI.

Other spatial arrangements of dislocations which more or less separate two regions of crystal from another, are the so-called deformation or kinkbands, discussed in Chapters VI and XIII.

With all these arrays a *curvature* of the lattice is usually associated, and misorientations of a few degrees or more are not rare.

Finally, the dislocation network itself, although it does not so much define separate crystal regions, still gives rise to a certain strain distribution. This shares with the other dislocation arrangements the property of being built on a certain well defined *scale*, viz. the grain size or subgrain size, the mean spacing between kinkbands or the maze width of the network.

It is clear that the dimensions of the "mosaic blocks" observed in X-ray diffraction work must reflect, at least in order of magnitude, this scale of the dislocation distribution. It is only when the actual dislocation distribution is known that the mosaic structure can be interpreted in more detail and dislocation densities can be obtained. If, for instance, it is known that a subgrain structure exists, and the mean size of this is t (as can for instance be determined from the observed secondary extinction), an upper limit to the dislocation density is found by assuming that the subgrain orientations are statistically distributed about a mean orientation. If only the standard deviation of the orientation angle is known, β, it is reasonable to suppose that the spread in orientation is also β. In order of magnitude, the dislocation density is then $N_d \approx 3\beta/bt$, where the factor 3 arises from the averaging over all dislocation orientations. If, however, the distribution function of orientation differences between the subgrains is accurately known (e.g. from a study of rocking curves) and the mean angle across the boundaries is α, the dislocation density is exactly $N_d = \alpha/bt$.

In a random network of dislocations, it can be assumed that each dislocation contributes independently of the others to the line broadening, and that the range over which a dislocation exerts its influence is

of the order of t, which now denotes the scale of the network. Then the *mean angle* over which the crystal planes are curved is b/t, and the dislocation density follows as about $1/t^2$. This is no longer true if the dislocations are arranged in such a way that they produce least curvature (rows of positive and negative dislocations alternating). Then the actual density may be an order of magnitude higher than the computed density. The latter is thus a sort of lower limit. Also an upper limit can be estimated. HIRSCH in his review [1957] gave a table of limiting dislocation densities obtained from microbeam X-ray experiments of well-annealed metals, under the assumption that the dislocations are randomly distributed. In Table 1 part of this list is reproduced.

TABLE 1

Limits of dislocation densities in annealed metals (after HIRSCH [1957]).

Metal	*Upper limit of dislocation density* (cm^{-2})	*Lower limit of dislocation density* (cm^{-2})
Al 99.999% recrystallized, grain size 30μ	3×10^8	1.3×10^7
Al 99.99% strain-annealed	10^7	1.5×10^6
Al 99.99% grown from melt	1.1×10^7	2.0×10^6
Al 99.993% bent and annealed	10^8 within subgrains; 5×10^6 in boundaries	5×10^6
Cu 99.999% recrystallized	2×10^7	4×10^6
Ge grown from melt, single cr.	0–10^5	
NaCl grown from melt, single cr.	10^6–10^8	
NaCl polished, single cr.	3×10^8	

From the table it appears that in well-annealed metals and ionic crystals the dislocation density varies between 10^6 and 10^8 cm^{-2}, whereas in the valence crystal germanium this density is much less, between zero and 10^5 cm^{-2}. We come back to this difference in Chapter XXX.

2.3.2. *Deformed crystals*

In deformed crystals there is no such thing as a mosaic structure. The diffraction lines are strongly broadened, which has been interpreted as due to 1) continuous internal strains and 2) to the apparent breaking up of the lattice into "particles": (cells, subgrains or frag-

ments, see Chapter XI). It is difficult to separate the effects of "particle-size" and strain broadening experimentally because both are dislocation effects and therefore intimately related, and conflicting evidence exists therefore on the relative importance of the two. Particle-size broadening is not very important if the particle sizes are larger than a few times 10^{-5} cm. This seems to be the case even in heavily deformed (filed) metals. As to the actual physical interpretation of the particle-size broadening, not much is definitely known. It may in some instances simply reflect the distance between deformation or kinkbands.

Data on the observed broadening can be used to estimate the elastic energy stored in a cold-worked metal lattice. Generally, this is found to be much smaller than the energy measured by calorimetric methods. The discrepancy is probably due to the fact that only *long-range* strain can be detected by the usual X-ray methods; the energy stored in the point defects formed is not observed. As a whole the information yielded by line-broadening studies of deformed metals is rather uncertain. Nevertheless, dislocation densities have been deduced from an investigation of the line shape by Fourier analysis from which the scale of the dislocation distribution can be inferred. Actually, some aid is given here by Wilson's analysis of the scattering of a single dislocation (Chapter V). Densities of the order of 10^{11} cm^{-2} were found in iron, molybdenum and tungsten filings, and a little less in aluminium filings (WILLIAMSON and SMALLMAN [1954], WILLIAMSON and HALL [1953]). These values agree rather well with the particle size determinations of about 10^{-5} cm. Already mentioned in Chapter V was the work of HALL and WILLIAMSON [1951] on the increase of background intensity which led to a much larger density, as other methods also usually do.

The discrepancy is probably due to the non-uniformity of the dislocation distribution present in deformed metals; the occurrence of piled-up groups, for example, may seriously limit the applicability of the simple strain-broadening theory. Point defects, moreover, although they affect the line shape but little, make their influence known especially in the wings of the diffraction lines and in the background scattering, as is the case for dislocation-core scattering.

Some evidence on the existence of piled-up groups can be obtained from the study of Laue or Debye-Scherrer asterism, as this indicates only curvature of the lattice and such curvature is always expected to occur near piled-up groups. Misorientations of several degrees give

rise to diffuse diffraction streaks of well-measurable extension, which break up into discrete small spots after annealing at a high temperature (polygonization). Recent experiments have indicated the presence of such a small-scale substructure directly after deformation but before an annealing treatment had been applied. It does not show up in ordinary Laue diagrams, because of the limited resolving power, but it can be detected for instance by measuring rocking curves. For a more detailed description of peculiar features of the dislocation distribution in deformed metals, the reader is referred to the original papers as cited by Hirsch.

2.3.3. *Microbeam experiments*

Especially suited to detect discrete misorientations on a small scale are X-ray investigations with the aid of a very narrow beam. RIECK [1957] has in this way studied *fragmentation* in bent single-crystal tungsten wires. The asterisms were broken up into small spots, again indicating the existence of a kind of polygonized structure without any annealing having taken place (room-temperature deformation). Another type of fragmentation is found in the already mentioned "spotty arcs". Possibly in both cases the effect is merely due to the presence

TABLE 2

Dislocation densities in cold-worked metals from microbeam experiments

Metal	*Deformation*	*Misorientation within one grain*	*Mean value of excess dislocation density in boundaries* (cm^{-2})	*Max. value of total dislocation density within particles* (cm^{-2})	*Ratio between max. and min. calculated densities*
Aluminium	50% rolled	15°	9×10^9	7×10^8	5
Copper	2% extended	2°	8×10^8	8×10^8	3
Copper	50% rolled	15°	2×10^{10}	4×10^9	10
Gold	50% rolled	20°	2×10^{10}		10
Iron	50% rolled	14°	7×10^9	1.5×10^9	4
Zinc	15% rolled	7°	8×10^8		8

of small regions containing more excess dislocations of one sign than in the bulk of the crystal (HIRSCH [1957]), thus not to polygonization or even fragment formation as this is usually understood. Spotty arcs

occur in many metals deformed at room temperature. The distance between curved regions is generally found of the order of 1μ in heavily worked metals, and is thus appreciably *larger* than the particle sizes as found from line-broadening studies. This is understood by noting that the latter quantity denotes more or less the scale of the individual dislocation distribution whereas the first quantity gives information as to the occurrence of "large scale" inhomogeneities in this distribution.

Microbeam experiments have been also used to obtain dislocation densities in deformed metals, especially were macrobeam methods fail. Table 2 gives some of the results obtained by GAY and coworkers [1953, 1954]. The densities found are remarkably low, and are not considered as very reliable by the present author.

2.3.4. *Cells*

Metals such as aluminium and zinc deformed in *creep* show characteristic X-ray diffraction effects, which can be interpreted as due to a specific subgrain or *cell* structure (WOOD [1948], MCLEAN [1951]). The dimensions of these cells depend on the temperature and rate of strain, they are largest in specimens deformed slowly (0.01%/hr) and at a high temperature. The size of the subgrains varies between 10 and 100μ, and is thus much larger than the scale of fragmentation observed in room temperature work. The angular misorientations vary from a few to ten degrees. Dislocations arranged in walls in densities up to 10^{10} cm^{-2} can explain the observed features. Although it was first thought that here we have to do with a clear case of polygonization, taking place during the creep test itself, it has become apparent that this process cannot be responsible for the whole mechanism of cell formation, but that grain boundary movements also play a part. Cells do not develop during the initial stages but only during steady-state creep (cf. Chapter XII). This stage is associated (at least in aluminium) with a continuous increase of the mutual misorientations of the cells. A full discussion of this phenomenon can be found in MCLEAN's book [1957].

2.4. COMPARISON OF X-RAY DISLOCATION DENSITIES WITH OTHER DETERMINATIONS

Apart from X-ray determinations, a few other, more or less theoretical methods exist by which the density of dislocations in well-annealed metals can be derived. These methods we call theoretical, because the

numerical value of the density is only obtained by applying dislocation theory to interpret the observed effects (the same is of course also true in a sense for the X-ray methods). By comparing the critical stress of dislocation sources with the observed flow stress in metals, MOTT [1952] deducted densities of about 10^6–10^7 cm^{-2} in well-annealed metals. SEEGER [1957], however, obtained values an order of magnitude higher, by considering the flow stress to be determined by the mutual interaction between dislocations rather than by the source lengths (see Chapter VIII). From a consideration of the form of the first part of the stress-strain curve this author found $N_d = 10^8$ cm^{-2} in annealed metals. VAN BUEREN [1957] derived about the same value from electrical resistivity data (Chapter XV), and yield-point studies in carbon-containing iron also lead to this order of magnitude. The dependence of the critical shear stress on the diameter of thin metal crystals (whiskers) again points to 10^8 cm^{-2} as the normal dislocation density. Thus the X-ray determinations seem to result in values of N_d that are smaller by about an order of magnitude than the other determinations.

In cold-worked metals the same apparent discrepancy exists. From the simplest strain-hardening theory imaginable (Taylor theory, see Chapter VIII), a dislocation density of 10^{12} cm^{-2} is estimated in severely cold-worked metals; from the observation of strain-ageing in α-iron (Chapter X) several authors obtained at maximum a few times 10^{11} cm^{-2} (PITSCH [1957]). The spotty-arc method as reported in Table 2 leads, however, to a few times 10^{10} cm^{-2} at the utmost, even in strongly cold-worked metals.

Etch pit counts (see Chapter VI) in metals invariably indicate dislocation densities smaller by another factor of 10 than the results of X-ray determinations. The explanation of *this* discrepancy must probably be sought in the effect of impurities on the "etchability" of dislocations; see also Chapter VI for this topic. By transmission electron-microscopy Bailey (unpubl.) counted 7×10^{10} disl./cm^2 in polycrystalline silver, 31% extended. This confirms the etch pit countings.

3. Special Kinds of X-ray Diffraction

3.1. SMALL-ANGLE SCATTERING

A particular kind of X-ray diffraction that needs mentioning here is *small-angle scattering*. By this is meant the abnormal scattered intensity within the first few degrees of arc from the central beam, which is

often encountered in X-ray diffraction studies. Small-angle scattering can be due to a variety of causes, such as:

1) the presence of small inclusions, of the order of 100 Å or more in diameter, in the crystal;

2) line-broadening effects, as discussed in the previous section, that also act for the central undiffracted beam;

3) small slowly varying strains extending over long distances (a particular case of 2);

4) double diffraction.

GUINIER and FOURNET [1955] have in their book discussed in much detail the small angle scattering due to the first three causes. Particles of (electronic) density different from that of the surrounding medium and homogeneously distributed in it, scatter the incident X-rays within a (small) angle θ according to the exponential intensity distribution:

$$I(\theta) = I(0)\, n^2\, f(R_0, \theta) \exp\{-(16\pi^2/3\lambda^2) R_0^2 \sin^2\theta\}\,. \qquad (6)$$

Here n is the difference between the number of electrons in the particle and in an equal volume of the matrix, f is a form factor and λ is the wavelength of the X-rays. R_0 was called the *radius of gyration;* it is equal to $\frac{1}{2}R\sqrt{2}$ for a spherical particle of radius R, and to $d/2$ for a thin disk of radius R and thickness d. In the former case $f = 1$, in the latter it is $\lambda^2/4\pi^2R^2\sin^2\theta$. Clearly R_0 measures size *and* form of the particle, whereas n only measures its size. To distinguish this kind of small-angle scattering from those mentioned under 2) and 3) is very difficult. We refer to Guinier's book for a discussion.

A very important and remarkable type of small-angle scattering is caused in the following manner (Beeman, see NEYNABER *et al.* [1955]; also HIRSCH [1957] and GEROLD [1958]). Consider the *cone* of diffracted rays belonging to a particular Bragg reflection of one grain of a large-grained polycrystal. Each diffracted ray can in turn be diffracted against a suitably oriented set of lattice planes in another grain. Only strong reflections need be considered in the following. Each ray, once diffracted, thus gives rise to a number of secondary cones, etc. These and analogous cones from other rays cover the whole projection plane more or less homogeneously, and give rise to a (non-observable) background, except in a few particular spots, of which the most obvious one is the *central spot* where this background is exceptionally strong. All cones belonging to two successive reflections through the *same*

Bragg angle must pass through the central spot. It is therefore to be expected that in the immediate neighbourhood of this a certain extra background intensity will occur, due, as is clear, to double refraction (predominantly from low-index planes). In a perfect single crystal this double refraction is not possible except for certain directions of incidence and for the occurrence of a mosaic structure.

Local density variations in the electron distribution on a scale slightly exceeding the X-ray wavelength can be caused by dislocation cores and by voids and inclusions in the crystal, thus in general terms by density variations in the crystal lattice. As these can only have an effect when they extend over more than a few lattice parameters, point defects as such cannot be detected. Around dislocations, the lattice strains contribute appreciably to the small-angle scattering (DEXTER [1953]). Because non-Hookean strains are also present in the dislocation core, which give rise to quite considerable additional density variations, as SEEGER and STEHLE [1956] have shown, the small-angle scattering by dislocations is still enhanced. This second-order core effect is present in edge as well as in screw dislocations (which in the linear approximation give no small-angle scattering at all (SEEGER and HAASEN [1958]). The volume changes due to non-linearity in the dislocation core amount to about 1 atomic volume per lattice plane. LOMER [1957] has derived this order of magnitude in a very simple way by making use of Grüneisen's constant of thermal expansion (which is also a non-linear phenomenon). The dependence of the scattered intensity on the angle is in first instance governed by an exponential expression such as (6) for a perfect dislocation. When the dislocation is extended, as in face centred cubic metals, the actual dependence varies considerably with the angle between the incident X-ray beam and the plane of the stacking fault.

BLIN [1957] has carried out small-angle scattering measurements on plastically deformed Cu, Ni, Al and Zn, and found a very large scattered intensity, many times larger than is theoretically expected from the contribution of dislocation cores and non-Hookean strains. The same discrepancy (of about a factor 100) was observed by ATKINSON [1958] in copper and aluminium. The latter author concluded from this to the preponderance of double Bragg reflection over the scattering by dislocation cores. Double Bragg reflection occurs so strongly in

cold-worked specimens, because of the small desorientations present in the cold-worked lattice which are especially suited to produce double reflection. GEROLD and FRICKE [1958] have observed small-angle scattering in deformed single crystals and found it also there to be caused by double reflection rather than by the presence of voids. All this evidence shows convincingly that density changes and voids only contribute for a small part to the observed small-angle scattering in deformed metals. This is also borne out by the lack of correlation between the amount of small-angle scattering and the mechanical recovery on annealing of the metal (NEYNABER *et al.* [1955]): it is only the orientation differences that count, not the number of defects.

Small-angle X-ray diffraction can also occur when heterogeneities of a *chemical* nature are present in the metal. The statistical composition fluctuations in a disordered alloy produce a kind of weak, non-angular dependent scattering. A small tendency to deviate from the random distribution already produces much larger scattering at very small angles. This type of small-angle scattering becomes especially marked when precipitation takes place, even if only on a submicroscopic scale. Guinier and collaborators have done much work in this respect on age-hardening alloys; more will be said about it in Chapter XXI on precipitation. We merely mention here the detection, by small-angle scattering, of several distinct precipitation stages in aluminium-silver and aluminium-copper alloys (Guinier-Preston zones).

3.2. ANOMALOUS TRANSMISSION

The detection of very small lattice strains or even individual dislocations by the X-ray techniques mentioned in section 2 is very difficult. Line broadening and analogous effects decrease below experimental accuracy if the dislocation density drops below about 10^4 cm^{-2} (average lattice strain between 10^{-3} and 10^{-4}). Therefore, X-ray dislocation density determinations are no longer reliable below this density. Borrmann already in 1925 and recently HUNTER [1958] have drawn attention to a different method, which is able to detect lattice strains of about 10^{-5}. This method is based on the phenomenon of the *anomalous transmission* of X-rays through a (nearly) perfect crystal (Fig. 5). An X-ray wave train arrives upon a slab of perfect crystal at an angle θ_D so that Bragg reflection occurs from a suitable set of lattice planes. Diffraction theory shows that on entering the crystal, this beam can be decomposed into a set of two standing waves, *s* and *c*

in Fig. 5. One of these has nodes in the lattice planes (sine wave), the other has nodes halfway between (cosine wave). If the slab is not too thin, the latter wave, because of its strong interaction with the atoms, becomes completely absorbed. However, if the crystal is perfect and the atoms are very small in size, the interaction of the *sine wave* with the

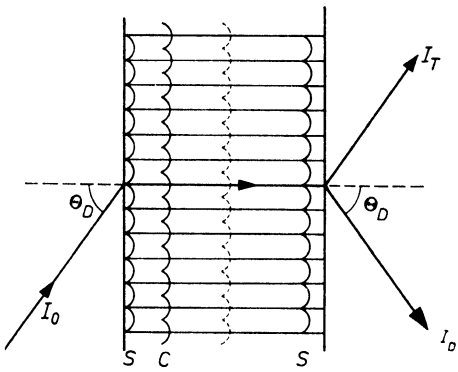


Fig. 5. Anomalous transmission of X-rays. The incident intensity I_0 enters under the Bragg angle θ_D and can be decomposed into a standing "sine wave" s with nodes at the lattice planes and a standing "cosine wave" c, shifted $\pi/4$ in phase. Only the former wave is transmitted with very small absorption in a perfect lattice, and gives rise on the other side to the transmitted wave I_T and the diffracted wave I_D.

lattice is practically zero, and this wave arrives with nearly its full intensity (half the incident intensity I_0) at the other side. There it is again decomposed into the *transmitted* wave I_T and the *diffracted* wave I_D. If there has been no absorption of the sine wave at all, both intensities are equal to $I_0/4$. Hunter showed, by applying the theory of X-ray diffraction to this problem, that for a nearly perfect crystal the intensity of the transmitted beam is given by

$$I_T = \tfrac{1}{4} I_0 \exp\{-\mu_0 t(1 - \varepsilon/\sqrt{(\theta - \theta_D)^2 + 1})\}. \qquad (7)$$

Here μ_0 is the absorption coefficient of the material for X-rays and t the thickness of the sample in the beam direction; the quantity ε in this formula takes account of the finite size of the atoms and the imperfect parallelity. In the ideal case mentioned above $\varepsilon = 1$ and indeed $I_T = I_0/4$ for $\theta = \theta_D$; in imperfect crystals $\varepsilon < 1$. Generally ε depends on the imaginary part of the structure factor only (absorbing index).

Owing to the occurrence of ε in the exponent, the anomalous trans-

mission decreases very rapidly with increasing degree of imperfection, and is only detectable in very perfect crystals. Slight deviations from the perfection, such as the presence of a dislocation here and there, already have a strong effect on the intensity of the anomalously transmitted beam. Hunter has studied the anomalous transmission through slabs of so-called dislocation-free germanium, using the widely spaced {220}-planes as diffracting planes. For these, in a perfect crystal ε was determined as 0.953 (not 1, as the atoms have finite size). The presence of lattice strains of the order of 10^{-5} could easily be detected in these crystals.

The study of the anomalous transmission of microbeams (of about 1 micron in width) should enable the detection of individual dislocations in nearly perfect crystals. In fact, it was reported recently that individual dislocations have been observed by this method in germanium (BORRMANN *et al.* [1958]). The *"Berg-Barrett"* and *Bonse-Kappler methods*, depending on the removal of primary extinction when a dislocation is present in a furthermore perfect crystal, enable dislocations to be viewed only if present in the surface layers (NEWKIRK [1959]). The method of "projection topography", devised by LANG [1959] is more powerful.

4. Neutron Diffraction by Lattice Imperfections

4.1. GENERAL

Neutrons are also diffracted by a crystal lattice since they also represent wave-packets. The wavelength of a neutron of kinetic energy E is given by

$$\lambda = \frac{h}{mv} = \sqrt{\frac{h^2}{2mE}}, \tag{8}$$

where m is the mass and v the velocity of the neutron and h denotes Planck's constant. Neutrons of 1 MeV energy have, according to this formula, a wavelength of only about 0.0003 Å, and are thus not suited for diffraction experiments; but so-called *thermal* neutrons, of about 0.025 eV energy, are represented by waves of slightly less than 2 Å in length. Both kinds of neutrons are readily available from nuclear reactors. The wavelength spectrum of thermal reactor neutrons is depicted in Fig. 6. "Monochromatic" neutrons can for instance be

obtained by reflecting the pile neutron beam against a crystal. Since the reflection angle depends on the wavelength, by using only those neutrons that are reflected at a certain angle a fairly monochromatic beam is then obtained. A certain spread in wavelength is always present, of course, which amounts to several hundredths of an ångstrøm unit in favourable cases. The wavelength is thus much less well-defined than that of X-rays; the wavelength spread of the latter being usually limited to within a few times 10^{-4} Å. Also the relative intensity of a neutron diffraction beam, expressed in terms of a flux of particles or quanta, is much lower, by a factor of about 10^5, than that of a normal X-ray beam.

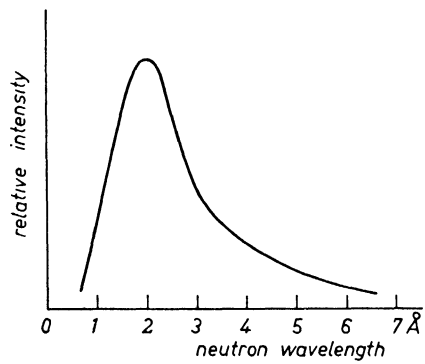


Fig. 6. Spectrum of thermal neutrons emitted by a common type of nuclear reactor (after BACON [1954]).

But more important differences exist. Whereas X-rays are predominantly scattered by the outer-shell electrons of the atoms and the scattering increases with atomic number and depends sensitively on angle, neutrons interact with the atomic *nuclei*. This is partly "*hard-sphere*" scattering (of the type mentioned in Chapter XIV): the nuclei act as impenetrable spheres. The scattering is isotropic and independent of the wavelength for slow neutrons (λ being much larger than the nuclear size). This hard-sphere scattering increases slowly and regularly with the atomic weight. Superposed on it is *resonance* scattering, however. The incident neutron and the scattering nucleus form first a compound nucleus, which may happen to be in one of its excited states. If this is so additional scattering, change of phase angle, etc. occur. *This* scattering is also independent of the wavelength for slow neutrons (the width of the resonance level is always much greater than the

neutron energy), but it varies considerably and *irregularly* with the atomic number. Owing to the phase shift, hard-sphere and resonance scattering amplitudes may even have opposite sign. In Fig. 7 the observed "neutron scattering radius" (which is always of the order of the actual radius of the nucleus) is presented as a function of mass number, and compared to the X-ray-scattering radius. Although neutrons are scattered far less efficiently than X-rays, the irregular dependence on atomic weight sometimes enables the distinction between neighbouring elements in the periodic system or the observation of light atoms that cannot be made by X-ray analysis (see also Table 3).

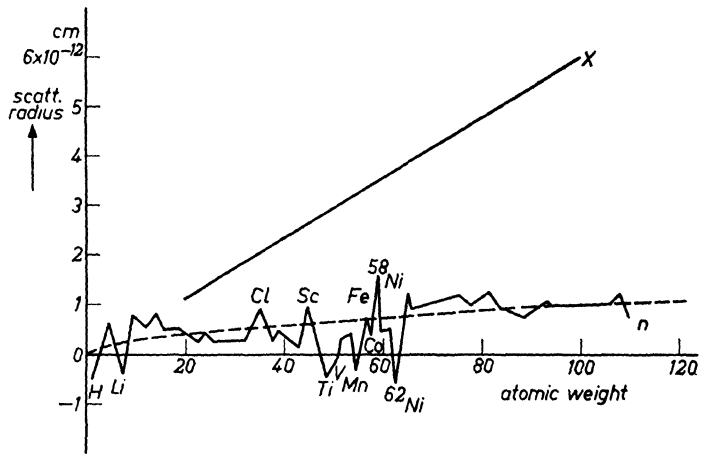


Fig. 7. Neutron scattering radius of atomic nuclei (lower drawn curve) and effective X-ray scattering amplitude (upper curve, sin $\theta/\lambda = \frac{1}{2}$Å^{-1}. After BACON [1954]).

A third factor that contributes to neutron scattering is the interaction between the neutron *spin* and nuclear *magnetic moment*, if present. This is of great importance in ferromagnetic materials. We shall not discuss it here.

4.2. EXPERIMENTAL METHOD

A regular atomic arrangement in a crystal lattice scatters the neutron waves similarly to X-rays. Diffraction is found only at certain discrete scattering angles, governed by Bragg's well-known formula. The technique of observation is different of course, photographic plates being of no avail for example. Because of the very small absorption coefficient of most crystals for neutrons, the neutron beam is weakened

TABLE 3

Neutron scattering amplitude $|b|$ and X-ray form factor $f(\theta = 0)$, and absorption cross-sections σ_n (at 1.08 Å) and σ_x (at 1.54 Å) (after BACON and LONSDALE [1953]).

Element	$\|b\|$ $\times 10^{-12}$ cm	$f(\theta = 0)$ $\times 10^{-12}$ cm	σ_n(1.08 Å) $\times 10^{-24}$ cm^2	σ_x(1.54 Å) $\times 10^{-24}$ cm^2
H	0.38	0.28	0.19	—
Li	0.18	0.84	40	—
Be	0.78	1.13	0.005	20
C	0.66	1.69	0.003	109
N	0.94	1.97	1.1	196
O	0.58	2.25	0.0005	336
F	0.55	2.53	0.01	547
Na	0.35	3.09	0.28	1170
Mg	0.54	3.38	0.04	1630
Al	0.35	3.65	0.13	2170
Si	0.40	3.95	0.06	2810
P	0.50	4.23	0.09	3730
S	0.31	4.5	0.28	4820
Cl	0.99	4.8	19.5	5040
K	0.35	5.3	1.2	9230
Ca	0.49	5.6	0.25	11380
Sc	1.02	5.8	7.1	13800
Ti	0.38	6.2	3.5	16160
V	0.05	6.5	2.8	19100
Cr	0.35	6.8	1.8	22200
Mn	0.37	7.0	7.6	25700
Fe	0.96	7.3	1.4	29800
Co	0.28	7.6	21	34300
Ni	1.03	7.9	2.7	4750
Cu	0.76	8.2	2.2	5620
Zn	0.59	8.5	0.6	6350
Ge	0.84	9.0	1.3	8300
As	0.63	9.3	2.5	9450
Se	0.89	9.6	7.4	10700
Br	0.67	9.8	3.8	12020
Mo	0.66	11.8	1.4	26000
Pd	0.63	12.9	4.0	36400
Ag	0.61	13.3	36	39600
Cd	?	?	2100	43400
Cs	0.49	15.5	17	75900
Ba	0.53	15.8	0.6	81300
W	0.47	20.8	11	51800
Pt	0.95	22.0	5	66000
Au	0.76	22.2	57	69500
Hg	1.31	22.5	210	73600
Pb	0.96	23.1	0.1	82200
Bi	0.89	23.3	< 0.01	86800
Th	1.01	25.2	?	125000
U	0.85	25.8	2.1	138000

almost entirely by reflection only (and large crystals, of several cm^3, are needed). The problem of secondary extinction of the stronger reflections is therefore important in neutron diffraction whereas in X-ray diffraction it is mostly only a correction. By using thin crystals secondary absorption can be avoided, but the reflected intensity is then low (10^{-4} per atomic plane). On the other hand, the large penetration depth of neutrons is, for some investigations, advantageous.

Owing to the small absorption, in a perfect crystal the dynamic diffraction should be small, but the existence of a mosaic structure raises the primary extinction, as in the case of X-rays. Another aspect of the small absorption is the possibility to measure the irregular variation of *transmitted* neutron intensity with wavelength (energy); this variation reflects the distribution of lattice spacings (Fig. 8).

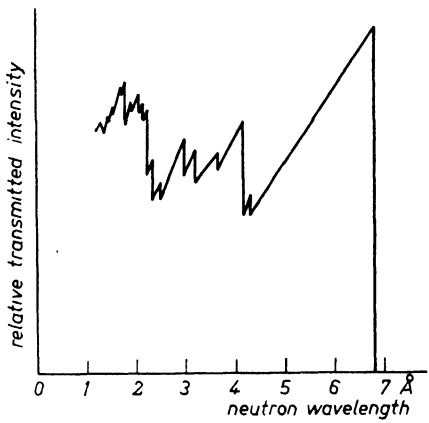


Fig. 8. Irregular variation with wavelength of transmitted neutrons by graphite. The peaks correspond to the onset of coherent diffraction from different sets of crystallographic planes (after BACON [1954]).

Table 3 gives the scattering amplitude of various nuclei for thermal neutrons, together with the corresponding atomic form factors for Cu K_α X-ray scattering, and the respective absorption cross-sections.

The distinction in scattering amplitude between heavy and light elements, which makes the latter unobservable in the presence of the former in X-ray diffraction, is entirely absent in neutron work. Therefore, as has already been said, the neutron diffraction technique is especially suited for the location of light elements in crystals, especially hydrogen (in the form of deuterium) although the limited resolving

power inherent to the method may constitute a serious drawback. In alloys with components differing but slightly in atomic number (e.g. Ni_3Mn or FeCo) neutrons can be used with advantage to distinguish between the two kinds of atoms, and to evaluate, for instance, the state of ordering, which is impossible by X-ray analysis. Also a distinction can often be made between different *isotopes* of the same element, which scatter the X-rays exactly the same way.

Defect crystal structures are being investigated by neutron diffraction techniques but many results have not yet been published. Usually in this field neutrons have no definite advantage over X-rays, except that by suitably choosing the energy and thus the wavelength, point defects and their movements can be detected in principle.

To obtain new insight into the cause of diffraction line broadening, viz. whether it is internal strain broadening or particle-size broadening (see section 2), neutron *transmission* experiments as a function of wavelength were carried out by WEISS [1951]. The discontinuities in transmitted intensity which are connected with the diffraction peaks should be more or less flattened out by continuous lattice distortions, but their form should be *in*dependent of particle size. As the transmission curves of cold-worked and annealed α-brass showed definite distinctions in the sharpness of some (but not of all) diffraction discontinuities, Weiss concluded to strainbroadening as actually present.

Small-angle neutron scattering is not a very suitable research tool for the study of imperfections of larger than atomic size in crystals. Owing to the small reflective power, only reflections by a large number of defects simultaneously can contribute to the small-angle neutron scattering. The phenomenon is of much use for the study of *liquid* structures, however. Recently, small-angle neutron diffraction by cold-worked metals has been measured and compared to the work of Blin, quoted in section 3.1, of small-angle X-ray scattering (note the large difference in wavelength, however). The neutron scattering was found 100–1000 times smaller than the latter and comparable to the theoretical dislocation scattering. It was therefore concluded that the large scattering found by Blin is due to other causes than lattice defects (double reflection, see previous section).

Not only neutrons and X-rays are diffracted by crystalline solids, other fast particles which can be regarded as wave packets also show diffraction. Electrons of 10–50 keV energy have a wave length of

0.1–0.01 Å. Since they are very intensely absorbed, however, only very thin layers (100 Å) can be studied in transmission, or surface layers in reflection. Electron diffraction is therefore especially suited for the study of *surface* effects. Already mentioned in Chapter VI was the spectacular use that has been made by HIRSCH *et al.* [1956] of electron diffraction to visualize individual dislocations and stacking faults in thin metal foils. Since this aspect of diffraction studies falls somewhat outside the scope of this chapter, the reader is referred to the chapter mentioned. It must be emphasized, however, that the underlying principles of the method are the same as those of microbeam scattering, for the dislocation strains are the cause of the decrease in transmissivity. The ability to observe stacking faults is due to an interference phenomenon.

For the sake of completeness we mention here finally the phenomenon of inelastic neutron scattering, caused by the transfer of energy between very low-energy neutrons and lattice vibrations. The study of this can give information as to the normal modes of vibration of a crystal lattice.

5. Dislocations and Magnetic Properties

In ferromagnetic materials the presence of imperfections influences the structure-sensitive magnetic quantities like permeability (or susceptibility), coercive force and hysteresis loss, to a certain extent. Some theoretical thought has been given to this subject by KERSTEN [1956]. Numerous experimental investigations about the influence of plastic deformation on coercive force, and about the effects of precipitation and quenching of iron alloys have been carried out by Köster and his school in Stuttgart. These have been published in a great number of papers in the Zeitschrift für Metallkunde and other German journals; we have already occasionally mentioned parts of this work. As the magnetic properties were predominantly only used to follow the precipitation and deformation processes and not to study the influence of the physical imperfections as such, we shall not discuss this work, however interesting, in detail. BROWN [1940] has carried out a theoretical and experimental investigation on the effect of dislocations on the structure-insensitive magnetic saturation. As could be expected, dislocations affect this quantity but to a small extent. However, the *approach* to saturation is influenced very strongly by lattice imperfections (see below).

To explain the effect of dislocations on the coercive force and permeability, Kersten assumes an interaction to exist between dislocations and *Bloch walls*, the two-dimensional boundaries between Weiss domains of different magnetization (Fig. 9). The dislocations fix the Bloch walls to a certain extent; application of a not too strong magnetic field H which tends to favour the growth of one domain (I) at the cost of the other (II), does not result in the displacement of the Bloch wall as a whole, but only in a *bulging* of the wall between its fixating dislocations. The coercive force is now regarded as the field strength H_c at which this *reversible* Bloch wall motion is replaced by the actual *irreversible* release of the wall from its dislocations (Barkhausen jump).

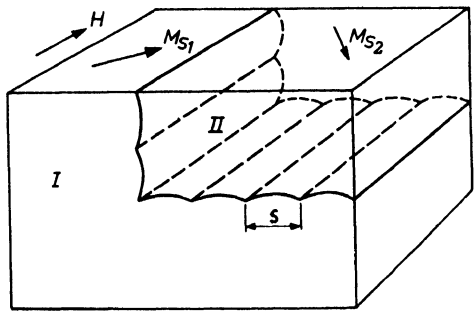


Fig. 9. Bulging of Bloch wall between two magnetic domains I and II with magnetizations M_{s1} and M_{s2} respectively, under the influence of an applied field H. The dislocations, spaced at a mean distance s, restrain the Bloch wall against the force exerted by the field (after KERSTEN [1956]).

Quite similar to the stability criterion for a Frank-Read source (see Chapter III), also here the critical shape of the bulged-out portion of the wall is the semicircular form. This is not surprising, as the two are governed by a line or surface tension respectively that is proportional to the increase in size. The critical field strength was found to depend on the dislocation spacing s as:

$$H_c = \frac{2\sqrt{2}}{\mu_0} \frac{\gamma}{sM_s}, \tag{9}$$

Here γ is the wall surface energy and M_s the magnitude of the saturation magnetization in both domains. μ_0 is the vacuum permeability

(induction constant in the Giorgi-system of units). Agreement with experiment in annealed iron, nickel and their alloys is found by putting $s = 10^{-3}$ cm, leading to the (small) dislocation density of 10^6 cm^{-2}. Not only the coercive force, also the initial permeability μ_i of a ferromagnetic substance is determined by the mobility of Bloch walls, and dislocations should have an influence on this too. Kersten derived the following theoretical relation between H_c and μ_i from his dislocation model:

$$H_c\,(\mu_i - 1) \approx \frac{1}{6} M_s \frac{s}{L}, \tag{10}$$

where L is the average width of the domains, which is usually also of the order of 10^{-3} cm or somewhat greater. Indeed good agreement with experimental values exists in this point and a similar connection with hysteresis losses is present. This need not unambiguously indicate an effect of *dislocations*, however.

The exact interaction mechanism by which dislocations can fix Bloch walls (which extend over several hundreds of Å in width and are thus much wider than the dislocations themselves) is not yet known. Presumably this interaction is of a magnetic rather than of an elastic nature. This, and the low dislocation densities derived, form a critical point in Kersten's theory and have been criticized by DIETZE [1958]. DIETRICH and KNELLER [1957] have investigated the *approach* of the magnetization curve of a ferromagnetic crystal to the saturation values as a function of deformation. They have thus studied the behaviour of the material at a field strength where the magnetization makes but a small angle with the field and Bloch wall motion is *not* important any more. It is rather the *rotation* of the spins of the atoms that now determines the permeability, and the physical quantity which governs the resistance against rotation is the magnetic *crystal anisotropy*, not the Bloch wall energy.

The susceptibility χ of the substance in this region depends on the field strength as

$$\chi = \chi_0 + \frac{C_2}{H^2} + \frac{C_3}{H^3} + \cdots \tag{11}$$

where C_2 and C_3 are constants. The authors mentioned have studied the variations of these constants with dislocation content in single crystals of nickel. They found both constants to increase with the

amount of plastic deformation, and could even detect a close correlation between the amount of change of C_2 and C_3 and the degree and stage of work-hardening (I, II or III), as explained in Chapter VIII. It would carry us too far to go into the explanation of this in terms of the influence of dislocation stresses on the crystal anisotropy. We refer instead to papers by DIETZE [1958] and by SEEGER and KRONMÜLLER [1959].

REFERENCES

ANDREW, E. R. (1953); Phys. Rev. **91,** 425.
ATKINSON, H. H. (1958); Phil. Mag. **3,** 476.
ATKINSON, H. H. and P. B. HIRSCH (1953); Phil. Mag. **3,** 213.
BACON, G. E. (1954); Research **7,** 257 and 312.
BACON, G. E. and K. LONSDALE; Rep. Progr. Phys. **16,** 1.
BLIN (1957); Acta Metall. **5,** 528.
BLOEMBERGEN, N. (1954); Rep. Conf. Defects in Cryst. Solids Bristol, p. 1 (London, The Physical Society).
BLOEMBERGEN, N., E. M. PURCELL and R. V. POUND (1948); Phys. Rev. **73,** 679.
BLOEMBERGEN, N. and T. J. ROWLAND (1953); Acta Metall. **1,** 731.
BORRMANN, G., W. HARTWIG and H. IRMLER (1958); Z. f. Naturforschung **A13,** 423.
BROWN, W. F. (1940); Phys. Rev. **58,** 736.
BUEREN, H. G. VAN (1957); Philips Res. Rep. **12,** 1.
DARWIN, C. G. (1922); Phil. Mag. **43,** 800.
DEXTER, D. L. (1953); Phys. Rev. **90,** 1007.
DIETRICH, H. and E. KNELLER (1957); Metalltagung Düsseldorf.
DIETZE, H. D. (1958); Metalltagung Aachen; (1960) Zs. f. Metallk., to be published.
GAY, P., P. B. HIRSCH and A. KELLY (1953); Acta Metall. **1,** 315.
GAY, P., P. B. HIRSCH and A. KELLY (1954); Acta Crystall. **7,** 41.
GEROLD, V. and H. FRICKE (1958); Metalltagung Aachen; Z. angew. Physik **10.**
GUINIER, A. and G. FOURNET (1955); Small-angle scattering of X-rays (Wiley, New York).
HALL, W. H. and G. K. WILLIAMSON (1951); Proc. phys. Soc. **B64,** 937.
HIRSCH, P. B. (1957); Progress in Metal Physics **6,** 236.
HIRSCH, P. B., WHELAN and R. J. HORN (1957); Phil. Mag. **1,** 670.
HUNTER, L. P. (1958); Proc. Kon. Ned. Acad. Wet. Amsterdam (1958); (1959) J. appl. Physics **30,** 874.
KERSTEN, M. (1956); Z. angew. Physik **8,** 496.
LANG, A. R. (1959); J. appl. Physics **30,** 1748.
LOMER, W. M. (1957); Phil. Mag. **2,** 1053.
MCLEAN, D. (1951); J. Inst. Metals **80,** 507 and many other papers in this journal.
MCLEAN, D. (1957); Grainboundaries in Metals (Oxford, Clarendon Press).
MOTT, N. F. (1952); Phil. Mag. **42,** 713.
NEWKIRK, J. B. (1959); Trans. Metall. Soc. AIME **215,** 483.
NEYNABER, R. H., W. G. BRAMMER and W. W. BEEMAN (1955); Phys. Rev. **99,** 615.
NYE, R. (1953); Acta Metall. **1,** 153.
OTSUKA, E. and H. KAWAMURA (1957); J. phys. Soc. Japan **12,** 1071.
PAKE, G. E. (1956); Solid State Physics **2,** 1.
RIECK, G. D. (1956); Acta Metall **4,** 47.

SEEGER, A. (1957); Acta Metall. **5**, 24.
SEEGER, A. and P. HAASEN (1958); Phil. Mag. **2**, 470.
SEEGER, A. and H. KRONMÜLLER (1959); J. Phys. Chem. Solids **12**, 298.
SHULL, C. G. and E. O. WOLLAN (1956); Solid State Physics **2**, 137.
SLICHTER, C. D., unpublished, reported in PAKE [1956].
SMEKAL, A. (1933); Hdb. der Physik **24**, 795 (Springer, Berlin).
STEHLE, H. and A. SEEGER (1956); Z. Physik **146**, 217.
WARREN, B. E. and B. L. AVERBACH (1952); J. appl. Physics **23**, 497.
WATKINS, G. D. and R. V. POUND (1953); Phys. Rev. **89**, 658. See also G. D. Watkins, Thesis Harvard Univ. 1954.
WEISS, R. J. (1951): Phys. Rev. **83**, 379; (1952): Phys. Rev. **86**, 271.
WHELAN, M. J. and P. B. HIRSCH (1957); Phil. Mag. **2**, 1303.
WILLIAMSON, G. K. and W. H. HALL (1953); Acta Metall. **1**, 22.
WILLIAMSON, G. K. and R. E. SMALLMAN (1954); Acta Crystall. **7**, 574.
WOOD, W. A. and G. R. WILMS (1948); J. Inst. Metals **75**, 693 and numerous other papers in the same Journal.

CHAPTER XVII

INTERNAL FRICTION IN METALS

Internal friction, or mechanical damping, forms a valuable research tool for the study of the properties of lattice imperfections in crystalline solids. The damping properties of a metal are particularly sensitive to the presence of physical imperfections, to the interaction between them and to the interaction with chemical impurities.

In this chapter we shall review the underlying principles of the damping phenomena and discuss the most important damping mechanisms encountered in *metallic* materials. These phenomena will be arranged in the order of the type of lattice imperfection that is held to be responsible.

A discussion of damping in other materials is postponed to the appropriate chapters.

1. Mechanisms of Internal Friction

1.1. RELAXATION AND ANELASTICITY

In perfectly elastic materials, stresses and strains are connected by time-independent relations, which in most cases can be approximated by the *linear* tensor equation (see Chapter IV):

$$\sigma = M \cdot \varepsilon \,, \tag{1}$$

M is the 4th rank tensor of the elastic coefficients. It reduces to Young's modulus of elasticity in simple extension or compression, and to the shear modulus G in torsion of a circular wire. We shall restrict ourselves to these simple cases in the following, that is, we shall regard M, the *modulus of elasticity*, as a *scalar constant* and the stresses and strains as *scalar* functions of the space coordinates (and the time).

A more general kind of *viscous* phenomena is described by an equation which, unlike (1), contains also the time derivatives of stress and strain. We shall limit our attention to the relatively simple case where only the first time derivatives occur and, moreover, the equation is of the first degree. A material to which such a restricted viscous equation applies was called a *standard linear solid* by ZENER [1949] and a *linear visco-elastic solid* by others (NOWICK [1953]).

One can write:

$$\sigma + \tau_1 \dot{\sigma} = M\,(\varepsilon + \tau_2 \dot{\varepsilon}), \tag{2}$$

where the coefficients τ_1 and τ_2 are the so-called *times of relaxation*; dots denote time derivatives.

On the sudden application of a constant load σ_0, the solution of (2) reads:

$$\varepsilon(t) = M^{-1}\sigma_0 + (\varepsilon_0 - M^{-1}\sigma_0)\, \mathrm{e}^{-t/\tau_2}. \tag{3}$$

The strain raises instantaneously to ε_0, and then afterwards gradually to the end value $\varepsilon(\infty) = \sigma_0/M$. This behaviour is depicted in Fig. 1*a*.

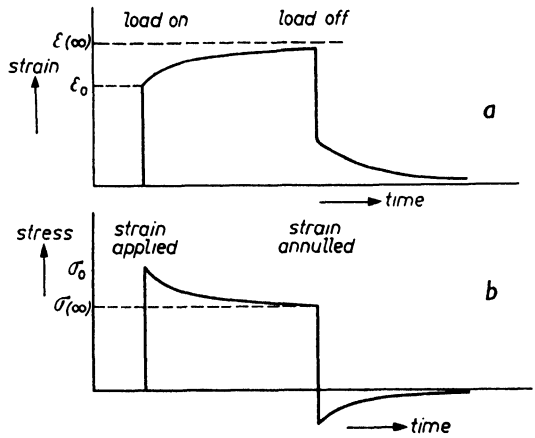


Fig. 1. *a*) Relaxation under constant stress; *b*) relaxation under constant strain.

Reversely, on removing the load σ_0, the strain drops suddenly by the amount ε_0 and then further decreases gradually to zero. The time

dependence of this *after-effect* or *relaxation* is completely governed by the relaxation time τ_2, the "relaxation time under constant stress". Sudden application of a constant *strain* ε_0, results in an instantaneous stress σ_0 followed by a gradual decrease to the end value $\sigma(\infty) = M\varepsilon_0$ (Fig. 1*b*). The time dependence of this decrease is governed by the "relaxation time under constant strain" τ_1.

In both cases the proportionality between stress and strain, *after* complete relaxation has taken place (thus effective for $t \gg \tau_1$ or τ_2), is described by the modulus M, which is therefore called the *relaxed* elastic modulus. On the other hand, a very *rapid* variation of the load results in a rapid variation of the strain, which, as follows from (2), is again proportional to the momentary stress, but now with a coefficient of proportionality

$$M' = \frac{\tau_2}{\tau_1} M . \tag{4}$$

M' is called the *unrelaxed* elastic modulus. It is always greater than or at least equal to the relaxed modulus, since, as there is *dissipation* of energy, $\tau_2 \geqslant \tau_1$ (they usually differ but a few per cents).

Of much interest is the case of a periodically varying load. The relaxation then results in a strain that lags behind the stress. Putting

$$\sigma = \sigma_0 e^{i\omega t} \qquad \text{and} \qquad \varepsilon = \varepsilon_0 e^{i(\omega t - \delta)} ,$$

one finds from (3):

$$(1 + i\omega\tau_1)\sigma_0 = M(1 + i\omega\tau_2)\varepsilon_0 . \tag{5}$$

Thus stress and strain amplitude are now correlated by a complex modulus

$$\mathcal{M} = \frac{1 + i\omega\tau_2}{1 + i\omega\tau_1} M .$$

The tangent of the phaselag angle δ is simply given by the ratio between the imaginary and real parts of the complex modulus. Thus

$$\tan\delta = \frac{\omega\Delta\tau}{1 + \omega^2\tau^2} , \tag{6}$$

where we have put

$$\Delta\tau = \tau_2 - \tau_1 \quad (>0)\,, \qquad \tau = \sqrt{\tau_1\tau_2}\,. \tag{7}$$

The phase lag is maximum for $\omega = 1/\tau$, where it reaches the value

$$\beta = (\tan\delta)_{\max} = \frac{1}{2}\frac{\Delta\tau}{\tau} = \frac{1}{2}\frac{\Delta M}{M}, \tag{8}$$

where ΔM stands for $M' - M$. Fig. 2 depicts $\tan\delta$ as a function of the frequency of excitation, together with the variation of the absolute value of $\mathscr{M}$:

$$|\mathscr{M}| = \frac{1+\omega^2\tau^2}{1+\omega^2\tau_1^2}M\,. \tag{9}$$

It follows directly from (5) that for

$$\omega \ll \frac{1}{\tau}, \qquad |\mathscr{M}| = M\,,$$

the relaxed modulus: there is now sufficient time during each cycle for the relaxation process to take place nearly completely, and no phaselag occurs. On high-frequency loading

$$\omega \gg \frac{1}{\tau}\,;$$

$|\mathscr{M}|$ approaches the limit

$$\frac{\tau_2}{\tau_1}M\,,$$

the unrelaxed modulus, since now practically no time at all is available during one cycle for relaxation to take place.

When δ is small, as is generally the case in practice, $(\tan\delta) \approx \delta$ is often denoted by the symbol Q^{-1}, where Q denotes the *quality factor* of the oscillating system. This terminology is based on the analogy with a damped electromagnetic circuit. The after-effect is accompanied by a *dissipation* of energy,

$$\Delta W = \mathrm{Re}\int_0^{2\pi/\omega} \sigma\,\dot{\varepsilon}\,dt\,,$$

per cycle (Re indicates that the real part of the integral must be taken). Putting as before $\sigma = \sigma_0 e^{i\omega t}$ and $\varepsilon = \varepsilon_0 e^{i(\omega t - \delta)}$, one has

$$\Delta W = \pi \sigma_0 \varepsilon_0 \sin \delta .$$

The *relative* energy dissipation becomes

$$\frac{\Delta W}{W} = 2\pi \sin \delta .$$

By analogy with circuit theory, the quality factor can indeed now be defined as

$$Q^{-1} = \frac{1}{2\pi} \frac{\Delta W}{W} = \sin \delta . \tag{10}$$

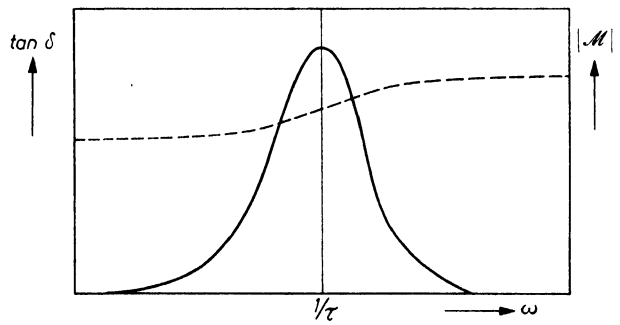


Fig. 2. Schematic variation of tan δ (drawn curve) and $|\mathcal{M}|$ (dotted curve) with frequency for simple relaxation.

When δ is small, sin δ and tan δ are about equal and the above statement becomes self-evident. When δ is *not* small (a theoretical case), from (10) and (6) one derives the general relation

$$Q^{-1}(\omega) = \frac{\omega \Delta \tau}{\sqrt{(1 + \omega^2 \tau^2)^2 + \omega^2 \Delta \tau^2}} . \tag{11}$$

The energy dissipation results in a gradual decrease of the amplitude of vibration of a freely vibrating sample. As a measure of this decrease the *logarithmic decrement* Λ is commonly used. This is directly connected with Q^{-1}, viz. by:

$$\Lambda = \pi Q^{-1} . \tag{12}$$

These considerations can be applied immediately to the damping of

the eigenvibrations of an excited specimen, as commonly used in the internal friction experiments. Sometimes, however, not this damping, but the *attenuation* of acoustic waves, excited in the specimen by external means, is studied. For the case of pure compression or shear waves, damping, logarithmic decrement or Q^{-1}, and attenuation α are closely related.

In this case we apply the simple one-dimensional Newton equation to describe the propagation of a strain wave:

$$\varrho \frac{\partial^2 u}{\partial t^2} = \frac{\partial \sigma}{\partial x},$$

where ϱ is the density of the material, u the local displacement in the x-direction and σ the local stress. One has further

$$\varepsilon = \frac{\partial u}{\partial x}.$$

Finally taking into account equation (3) also, one obtains the following *wave equation* valid in a standard linear solid:

$$\frac{\partial^2 \Psi}{\partial t^2} + \tau_1 \frac{\partial^3 \Psi}{\partial t^3} = \frac{M}{\varrho}\left(\frac{\partial^2 \Psi}{\partial x^2} + \tau_2 \frac{\partial^3 \Psi}{\partial x^2 \partial t}\right). \tag{13}$$

Here Ψ stands for either ε or σ; these obey the same equation (LÜCKE [1956]).

A solution describing an attenuated running wave with frequency ω and wave vector $k = \omega/c$, where c is the velocity of sound in the material, reads

$$\Psi = \mathrm{A} \exp(-\alpha x) \cos(kx - \omega t),$$

where the *attenuation factor* α is given by

$$\alpha = \frac{k}{\omega} \frac{\Delta\tau}{2\tau^2} \cdot \frac{\omega^2\tau^2}{1 + \omega^2\tau^2} = \tfrac{1}{2}k \frac{\omega\Delta\tau}{1 + \omega^2\tau^2}. \tag{14}$$

Apart from the factor $\frac{1}{2}k$, this is the same expression as (6) for the phaselag angle of free vibrations. Thus

$$\alpha = \tfrac{1}{2}kQ^{-1} = \frac{\omega}{2c} Q^{-1}. \tag{15}$$

The assumption that the phaselag angle is small is equivalent to

the previous assumption $\Delta\tau \ll \tau$. The single relaxation time τ then suffices to describe the frequency dependence of the damping. In many instances, the relaxation involves atomic movement by diffusion and this relaxation time is strongly temperature dependent, according to

$$\tau = \tau_0 e^{H/kT},$$

where H denotes an energy of *activation*. Then the frequency of maximum damping also depends on temperature:

$$\omega_{max} = \omega_{max}(\infty) \cdot e^{-H/kT}. \tag{16}$$

Owing to this feature, to observe the whole relaxation curve one need not vary the vibration frequency over many octaves, but only change the temperature over, say, a few tens of degrees, leaving the frequency constant.

The related frequency and temperature dependence of relaxation-type internal friction, together with the absence of amplitude dependence (owing to the linearity of (3)), are quite typical, and serve to distinguish this from other kinds of damping, some of which we shall discuss below.

1.2. RELAXATION SPECTRUM

A relaxation behaviour that conforms exactly to the simple theory discussed above, does seldom occur in practice. An obvious complication results from the occurrence of more than one relaxation time, or even of a continuous distribution of relaxation times, a *relaxation spectrum*.

When the stress-strain relation is linear, the following principle of superposition holds (ZENER [1949]).

If $\varepsilon_1(t)$ is the strain produced by the stress $\sigma_1(t)$ acting alone, and $\varepsilon_2(t)$ the strain produced by the stress $\sigma_2(t)$ acting alone, then $\varepsilon_1 + \varepsilon_2$ is the strain produced by the stresses σ_1 and σ_2 acting simultaneously. Mathematically, we can express the superposition principle for the special case of relaxation as follows. We introduce the *relaxation function* $X(t)$, which gives the strain at time t due to the sudden application, at $t = 0$, of a stress of magnitude unity. Then the strain due to an arbitrarily varying stress is given by

$$\varepsilon(t) = \int_{-\infty}^{t} X(t - t')\dot{\sigma}(t')\mathrm{d}t'. \tag{17}$$

In the interval between t' and $t' + \mathrm{d}t'$, the amount $\dot{\sigma}(t')\mathrm{d}t'$ was added to the stress, and this has given rise to a strain increment at the time t of $\mathrm{d}\varepsilon = X(t - t')\ \dot{\sigma}(t')\mathrm{d}t'$. Integration leads to (17).

The anelastic behaviour of the solid is evidently completely described by the relaxation function X. It is also completely specified by the spectrum of relaxation times, which, in turn, is fully described by the normalized distribution function of relaxation times $F(\tau)\mathrm{d}\tau$. Obviously, $F(\tau)$ and $X(t)$ must be related. It can be shown that the relaxation function is an integral transform of the relaxation spectrum (GROSS [1953]). One has

$$X(t) = X(0) \int_0^\infty F(\tau)\ \mathrm{e}^{-t/\tau}\mathrm{d}\tau\ . \tag{18}$$

When the spectrum consists of a single line, that is, when only one relaxation time τ_r occurs, $F(\tau)$ reduces to the delta function $\delta(\tau - \tau_r)$, and (18) to:

$$X(t) = \mathrm{e}^{-t/\tau_r}\ , \tag{19}$$

an obvious result.

1.3. MECHANICAL HYSTERESIS

There are other mechanisms conceivable that produce energy dissipation and thus internal friction. One of the most important of these is associated with the occurrence of *irreversible changes* of state. This category of damping mechanisms is denoted by the name *hysteresis*. It is exceedingly diverse and cannot be described mathematically as simply as the relaxation phenomena. The irreversibility can in the case we are interested in, namely that of mechanical hysteresis, usually be expressed as a *multivalued* relation between stress and displacement: the displacements in a specimen do not, after removal of the load, return to the values they had before application of that load. Fig. 3 depicts schematically a simple case of hysteresis. Again the amount of damping can be expressed by the quantity

$$Q^{-1} = \frac{\Delta W}{2\pi W}\ ,$$

which also gives the phaselag angle between stress and strain in periodic loading ($\Delta W/W$ is the relative energy dissipation per cycle). As the

stress-strain relation in a solid showing hysteresis is essentially time-independent, the internal friction due to mechanical hysteresis is independent of *frequency*. Since, moreover, hysteresis is a non-linear

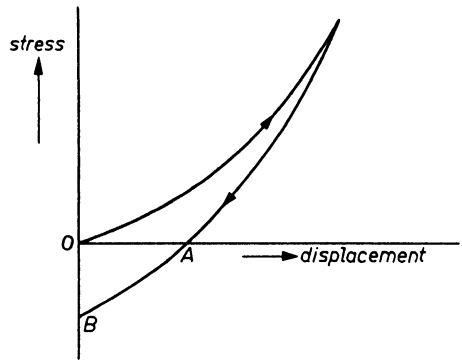


Fig. 3. Schematic illustration of mechanical hysteresis. After unloading from the stressed condition, the displacement does not return to its original zero-value (*O*), but a finite amount of displacement is left (*A*), that can only be removed by applying an opposite stress (*B*).

process, on increasing the strain amplitude, the energy loss per cycle usually increases too. Hysteresis internal friction thus *increases with amplitude*. These two characteristic properties can be directly contrasted with the typical frequency dependence and (in linear solids) amplitude independence of relaxation internal friction.

1.4. DAMPED RESONANCE

A third type of frequency dependence of the internal damping of a vibrating solid can be observed when resonance phenomena take place in that solid. As an example, one can consider the case of dislocations vibrating about their equilibrium position in resonance with the applied stress. Although relaxation or hysteresis damping may very well be negligible in the frequency region concerned (usually high frequency), owing to the strong response of the solid when the frequency of loading approaches the resonance frequency, notable energy dissipation can yet occur, leading to observable damping. At resonance, for instance, the radiation of energy in the form of sound waves by a rapidly oscillating dislocation reaches a measurable level. Damped forced vibrations can usually be mathematically described by the

differential equation:

$$m\ddot{x} + \beta\dot{x} + \varkappa x = Fe^{i\omega t}, \tag{20}$$

when we confine our attention to one-dimensional harmonic systems; m is the mass of the oscillating agent, $\varkappa$ the force coefficient (a constant in harmonic systems) and $Fe^{i\omega t}$ the periodic driving force. $\beta\dot{x}$ is the type of damping term usually assumed to apply (viscous damping). The solutions of (20) have the form:

$$|x| = \frac{|F|}{m\sqrt{(\omega^2 - \omega_0^2)^2 + \frac{\beta^2}{m^2}\omega^2}}, \tag{21}$$

where ω_0 is the resonant frequency: $\omega_0 = \sqrt{\varkappa/m}$.

The internal friction is again measured by the relative power dissipation per cycle. This is found to be

$$Q^{-1} = \frac{1}{2\pi}\frac{\Delta W}{W} = \frac{\omega/\tau_r}{\sqrt{(\omega_0^2 - \omega^2)^2 + \omega^2/\tau_r^2}}. \tag{22}$$

Here we have formally put $\tau_r = m/\beta$, in order to show the close resemblance in frequency dependence between relaxation internal friction as expressed by (11), and damped resonance. Also in the latter case the internal friction goes through a maximum as a function of frequency (namely at $\omega \approx \omega_0$), and it is essentially independent of the amplitude (when (20) applies). There are several differences however. A variation of the relaxation time τ results in a shift of the peak frequency of relaxation internal friction, but it leaves the form of the damping curve essentially unaltered. Only a variation of $\Delta\tau$ influences the latter. In the resonance case, the frequency of maximum damping usually depends little or not at all on the value of the "damping time" τ_r. A variation in the latter, however, results in a narrowing or widening of the resonance peak (the width of which is proportional to $1/\tau_r$). To distinguish in practice between relaxation and resonance damping, the effect of a variation of τ, e.g. by a change of temperature, must therefore be investigated. This has usually a large effect on relaxation damping (rapid increase of ω_{max} with temperature according to (16)) and little on resonance damping (slow decrease of ω_0 with increasing temperature, owing to decrease of $\varkappa$). In other words, the

resonance frequency usually depends much less on the intensive variables of state than the relaxation time.

Resonance and relaxation transform continuously into each other, when the damping coefficient β *together* with the mass coefficient m increases to infinity, but their ratio τ_r remains finite. Equation (20) can, in the absence of an external force, be written as:

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = -\frac{\dot{x}}{\tau_r} - \omega_0^2 x \,.$$

Dividing through

$$\frac{\mathrm{d}x}{\mathrm{d}t} \equiv \dot{x} \,,$$

we have

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}x} = \frac{-\frac{\dot{x}}{\tau_r} - \omega_0^2 x}{\dot{x}} = -\frac{1}{\tau_r} - \omega_0^2 \frac{x}{\dot{x}} .$$

If m and β increase indefinitely, but τ_r remains fixed, ω_0 approaches to zero and we are left with

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}x} = \frac{1}{\tau_r} \,,$$

the solution of which is, of course

$$x = x_0 \, \mathrm{e}^{-t/\tau_r} + x_1 \,,$$

a typical relaxation-type expression. Relaxation is, therefore, to be considered as a special case of damped resonance.

1.5. VISCOUS DAMPING

Damping phenomena that cannot be classified according to one of the foregoing categories will be denoted by the name *viscous damping*. Only very detailed investigations can bring to light what the basic mechanisms are that underly such a process. A typical example of general viscous behaviour is formed by the case of relaxation with a very wide continuous range of relaxation times, or by that of *coupled* relaxations. Neither frequency nor amplitude dependence can be predicted. In general the mechanism is not linear.

Fig. 4 gives, by way of illustration, an example of a substance (glass) in which damping due to a number of causes occurs simultaneously. For a discussion we refer to STEVELS [1950, 1951]. These losses are dielectric rather than mechanical in origin, but the underlying principles are the same.

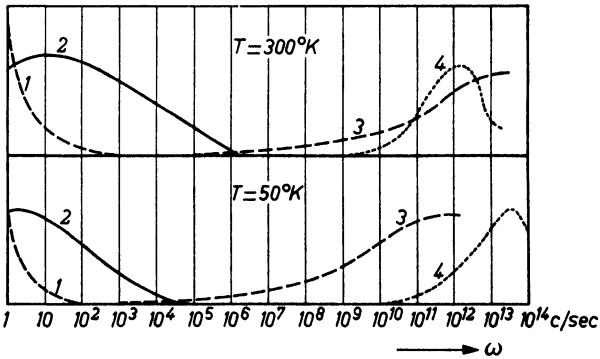


Fig. 4. Schematic diagram of the frequency spectrum of dielectric losses in glass at two temperatures.
1. Socalled *conduction losses*. These correspond to relaxation processes with very long relaxation times. 2. *Relaxation losses*. 3. *Deformation losses*, more or less corresponding to viscous damping and to relaxations with very short relaxation times. 4. *Resonance* (*"vibration"*) losses. On lowering the temperature, the relaxation losses shift to the left, the vibration losses to the right.

2. Observations of Internal Friction

2.1. EXPERIMENTAL TECHNIQUES

The frequency and amplitude dependence of the internal friction observed in metals and other solid materials is exceedingly complex. Moreover, the damping varies greatly with the impurity content, any amount of previous cold work and the crystalline texture of the material, as well as with the temperature. Internal friction has been studied over a frequency range from less than 1 c/sec to 300 Mc/sec. The many experimental methods used can be divided into four categories (NOWICK [1953]).

1) *Torsional pendulum method*

This is suitable for frequencies from 0.1 to 10 c/sec, and produces a maximum strain amplitude of about 10^{-5}. The method is not very

sensitive; very small damping ($Q^{-1} \ll 10^{-4}$) cannot be measured in this way, owing to apparatus losses. The frequency can only be slightly varied. Nevertheless, this method is often used, by reason of its simplicity and the easy interpretation of its results. The material to be examined is brought in the form of a thin wire, which acts as suspension wire for an inertia member. A modification of the method uses the transverse vibrations of a reed.

2) *Damped eigen-vibrations method*

The specimen, (usually in the form of a bar) is excited in one or more of its resonances and the decay of the oscillations is measured. The excitation is achieved by means of a transducer or similar device. Very small damping ($Q^{-1} \geqslant 10^{-7}$) can be measured by this method, as the external disturbing effects are small. The frequency range in which the method can be applied is determined by the practical range in dimensions of the specimen. Measurements have been described between 100 c/s and 300 kc/s; in each experiment the frequency can easily be varied over a few factors of ten, although the excitation, of course, only takes place at discontinuously spaced frequencies. The method is less suited to study the dependence on amplitude since this can only be slightly varied. Another drawback is the intricacy of the stress distribution in the vibrating specimen, which makes interpretation often difficult.

Since the vibration frequency in resonance is determined by the elastic constants (usually Young's modulus) of the specimen (and by its form and density), a change of this modulus makes itself known as a change in resonant frequency. The method discussed here is especially suitable to detect such modulus changes, and thus to give valuable new information on the mechanism of the internal friction, according to the theory explained in section 1.

It has been found that in some specific cases, to be mentioned later (see for instance section 2.2) Young's modulus is an especially sensitive indicator of small changes of the dislocation mobility. It is therefore also that the eigen-vibrations method is much used in practice.

3) *Pulse attenuation method*

The attenuation of ultrasonic wave pulses during propagation through a bar is measured. The frequency range extends above 1 Mc/s, frequency and amplitude can easily be varied, but the calibration of

the amplitude is difficult and the sensitivity of the method is severely limited owing to scattering of the waves ($Q^{-1} \geqslant 10^{-4}$).

4) *After-effect in static loading*

The specimen, usually in the form of a helical spring, is loaded under constant stress, and the gradual increase of strain is observed (Fig. 1*a*). This method obviously applies to damping investigations at frequencies below 1 c/sec; the sensitivity is comparable to that of the torsional pendulum.

A very large number of categories of mechanical internal damping phenomena can be distinguished in metals. The most important of these we shall discuss in the following sections. We shall distinguish between the categories in terms of the particular type of lattice imperfection or elementary process that is presumed to be responsible for it. The danger of this procedure is recognized, since the underlying mechanisms are not known for all damping phenomena.

2.2. DISLOCATION DAMPING

An important source of internal friction in solids are the dislocations. Dislocation damping is usually characterized by a strong dependence on previous plastic deformation. At least six distinct types of internal friction have been observed, caused by six different dislocation mechanisms. Only because they occur in different regions of frequency or temperature, has it been possible to separate them experimentally from each other. The theory of the dislocation damping is still in an unsatisfactory state and more systematic observations are needed to found it on a firmer base. These are difficult to obtain, however, since the experimental results are often disturbed by unverifiable factors. We shall therefore only outline the general theoretical ideas concerning the different types of dislocation damping.

2.2.1. *Dislocation relaxation (low-temperature or Bordoni internal friction)*

BORDONI [1949] discovered the existence of a fairly broad internal friction peak of the relaxation type in single crystals of copper, centred in the temperature region near 90 °K at the measurement frequency of 49 kc/sec. BIRNBAUM and LEVY [1956] and EINSPRUCH and TRUELL [1958] observed a similar maximum, in the same temperature and

frequency region, and in the Mc region respectively, in aluminium that had been slightly cold-worked. The internal friction of lead measured at 95 °K shows a peak at 30 kc/sec. MASON [1955a] found in the same metal a peak near 10 °K at a frequency in the Mc range. NIBLETT and WILKS [1956, 1957] repeated Bordoni's original experiments on Cu single and polycrystals and showed 1) that the damping occurred preferentially after slight cold-work, 2) that after heavy working the

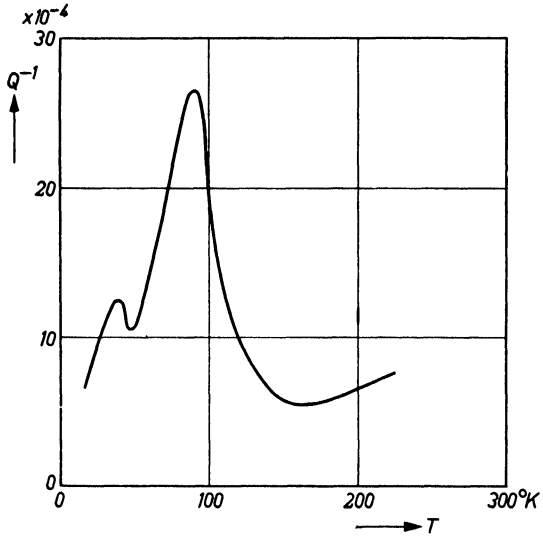


Fig. 5. Bordoni internal friction in copper, strained 2.2% (after NIBLETT and WILKS [1957]). Frequency 1100 c/s.

damping decreased again and 3) that annealing below the recrystallization temperature had only little effect; annealing above it resulted in a disappearance of the damping peak (together with a large decrease of mechanical hardness). These authors were also able to dissolve *two* relaxation peaks (after more than 2% prestrain), centred around 35° and 70 °K respectively at 1100 c/sec (Fig. 5). See also BORDONI [1960].

This damping is approximately independent of the strain amplitude, whereas the frequency of maximum damping depends exponentially on the temperature. MASON [1955a] derived an activation energy governing the latter variation of 0.05 eV in lead; in copper this energy is 0.122 eV. EINSPRUCH and TRUELL [1958] derived 0.11 eV for the activation energy in aluminium. All this evidence points rather unambiguously to the existence of a relaxation type of internal friction

in face-centred cubic metals, in which dislocations play an essential role.

To explain the low-temperature, medium- and high-frequency damping by some kind of relaxed motion of dislocations, Mason has envisaged the following process. A dislocation segment, pinned at its ends by impurity atoms or other dislocations, oscillates as a whole to and fro between two neighbouring equilibrium positions (adjacent valleys between atom rows). The activation energy for this process is proportional to the height of the barrier between the two equilibrium positions, that is, the Peierls barrier (see section IV–7.2) and to the length of the dislocation segment. In view of the latter fact it is surprising that the *observed* activation energy has such a well-defined value; further that it does not depend on the impurity content of the material, nor on the temperature. NIBLETT and WILKS's [1956] experiments showed that addition of a few hundredths per cent of phosphorus or bismuth to the copper does not affect the shape and the location of the internal friction peaks at all. SEEGER [1956] pointed out that the length of an oscillating segment could be determined by the dislocation properties itself. A dislocation line in thermal equilibrium in a crystal and lying parallel to a close-packed lattice direction is not straight, but contains a number of kinks (Fig. 6) in which its position changes from one

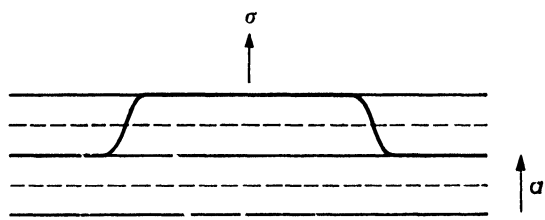


Fig. 6. Kinks in dislocation line lying along a lattice direction. Dotted lines denote potential (Peierls) barriers between fulldrawn positions of minimum energy.

atomic valley to another. (A dislocation along an *arbitrary* direction contains very many kinks already by geometrical reasons, as it is built up out of pieces along close-packed directions.) Each kink contributes to the free energy of the crystal an amount U_k, thus the distribution of lengths of segments between kinks will undoubtedly contain a factor exp $(-U_k/kT)$. A shear stress σ on the dislocation of Fig. 6 tends to move it over into the next valley, that is, it tends to move

the kinks outwards laterally. The distribution function is thus altered, or, in other words, U_k depends on the stress. An alternating stress produces rapid fluctuations in the distribution of kinks, which are accompanied by energy dissipation. As kink formation in dislocations along close-packed directions is temperature controlled, a thermally activated relaxation damping is the result.

Seeger argues that, at a given stress, a *critical separation* d_{cr} exists between kinks: pairs with a larger separation will expand outwards under the action of the stress, those with a smaller separation tend to annihilate owing to their mutual attraction. The activation energy observed in the damping is the energy needed to move *this* critical length of dislocation over the Peierls barrier.

The observation of *two* damping peaks in face-centred cubic crystals was explained by Seeger as caused by the existence of two types of dislocations along the close-packed ⟨110⟩ directions: pure screw dislocations, and dislocations with Burgers vectors inserting an angle of 60° with the dislocation axis. Only the small fraction of the dislocations that lie along ⟨110⟩ form kinks by a thermally activated process. The observed internal friction is due to just this small fraction of the dislocations. When the dislocation density increases by cold work, this *fraction* presumably decreases, since the mutual interaction of the dislocations tends to arrange them in other than close-packed directions. The internal friction as a function of prestrain will then go through a maximum.

Seeger's calculation runs along the following lines. The differential equation for the "transverse" deviation y under the action of a stress of a dislocation of Burgers vector b lying along the x-axis is written as:

$$m_d \frac{\partial^2 y}{\partial t^2} + \frac{\partial W}{\partial y} - b\sigma y - W \frac{\partial^2 y}{\partial x^2} = 0 . \qquad (23)$$

Here m_d represents the "dislocation mass", $W(y)$ the line energy per unit length, and σ the shear stress component along b. W is approximated by the simple expression

$$W = W_0 + W_1 \cos \frac{2\pi y}{a} ,$$

where a is the lattice spacing in the y-direction. W_0 is the line tension of a straight dislocation in equilibrium (see section III–2) and W_1 is the Peierls energy per unit length (section IV–8.2). Time independent

solutions of (23) are in first approximation ($W_1 \ll W_0$) solutions of the ordinary differential equation

$$W_0 \frac{d^2y}{dx^2} - \frac{2\pi W_1}{a} \sin \frac{2\pi y}{a} + b\tau y = 0 \tag{24}$$

By a procedure analogous to that followed in determining the energy of a crystalline dislocation in the Peierls approximation (section IV–8.2), from the solution $y(x)$ of equation (24) the energy of a single kink is found to be

$$U_k = \frac{4a}{\pi} \sqrt{W_0 W_1} . \tag{25}$$

For the critical separation distance between kinks of opposite sign Seeger equated the work done by the applied stress in bowing out a dislocation segment, to the interaction energy between the kinks, and found

$$d_{cr} = \frac{a}{2} \sqrt{\frac{W_0}{W_1}} \log \left(\frac{32 W_1}{\tau a b} \right) . \tag{26}$$

The total energy needed to displace a segment of length d_{cr} into the nearest equilibrium position is equal to the difference in energy of the segment somewhere midway between the two positions and its energy in the equilibrium position. In the former situation the dislocation contains two "half kinks", the combined energy of which is about equal to the energy of one full kink, U_k. The total energy difference is then

$$2W_1 d_{cr} + U_k . \tag{27}$$

Inserting (25) and (26) into (27), one finds, at least in this approximation, for the activation energy of the oscillatory motion:

$$H = a \sqrt{W_0 W_1} \left[\frac{4}{\pi} + \log \left(\frac{32 W_1}{\tau a b} \right) \right] . \tag{28}$$

DONTH [1957] has extended this calculation and put it on a firmer mathematical basis. From a comparison of the observed values of H with the theoretical expression, BORDONI *et al.* [1959] derived a value of the Peierls energy W_1 (or the equivalent quantity: Peierls stress $\tau_p = 2\pi W_1/ab$). They found $\tau_p = 5 \times 10^{-4} G$ in copper and $5 \times 10^{-5} G$ in lead. In copper especially, this seems a high value compared to

semi-experimental estimates, even at the low temperature concerned. The influence of dislocations not lying along ⟨110⟩ directions must always be taken into account, however, which have a much smaller effective Peierls stress.

Dislocation relaxation experiments provide information on the magnitude of the Peierls force, that is, indirectly, on the interatomic forces in solid materials. Its detailed study is therefore of great theoretical and practical importance.

2.2.2. *High-temperature dislocation damping*

At temperatures above 400 °C a marked increase of the medium and low-frequency damping in copper is noted, especially when the internal friction is observed at very small strain amplitude, where it increases exponentially with the temperature (the energy of activation varies considerably from specimen to specimen, and lies between 0.8 and 2.1 eV in copper (SACK [1950]); in aluminium the value 1.5 eV has been quoted (KÊ [1950a]). These values are of the order of the activation energy of self-diffusion. The magnitude of this internal friction decreases inversely proportional to the frequency. Also a certain but not very well-defined amplitude dependence exists: The damping decreases slightly with increasing amplitude. Cold work enhances the damping considerably, especially in single crystals, where the energy of activation itself increases too with the amount of prestrain. No distinct intercorrelation between high-temperature damping and impurity content has yet been recognized in literature.

The effect of cold-work again indicates dislocations as the cause. WEERTMAN [1955], WEERTMAN and SALKOVITZ [1956] and Sack have tried to explain this internal friction in terms of the damped oscillatory motion of the dislocations in a field of internal stress, made possible by the action of a thermally activated process. WEERTMAN [1957] takes impurity atoms or precipitates as the source of internal stress, and dislocation *climb* as one of the activated processes. Sack on the contrary regards the interaction between dislocations as the cause of the stress and sees their mutual intersection as the activated process. NOWICK [1953], following ZENER [1949] regards the damping as a kind of general relaxation of the internal stress across slip bands. A superficial resemblance to grain-boundary internal friction in polycrystals, of which stress relaxation across boundaries is the cause (section 2.3.3), is indeed present.

The inverse-frequency variation and the exponential temperature dependence can, in either model, be understood as the natural result of the action of one or two thermally activated relaxation processes with *very long* relaxation times, so that at the measuring frequency $\omega\tau \gg 1$. Then formula (11) goes over into

$$Q^{-1} \to \frac{\omega \Delta\tau}{\omega^2\tau^2} \to \left(\frac{\Delta\tau}{\tau}\right) \frac{1}{\omega\tau} . \tag{29}$$

If we put, as usual, $\tau = \tau_0 \exp(H/kT)$, we obtain:

$$Q^{-1} \approx \left(\frac{\Delta\tau}{\tau}\right) \frac{1}{\omega\tau_0} e^{-H/kT} , \tag{30}$$

a formula that describes the features of high temperature damping very well indeed ($\Delta\tau/\tau$, being the relaxation strength, is T-independent). A possible explanation for the observed amplitude dependence was indicated by NOWICK [1953] in terms of the existence of a mutual coupling between the relaxation processes, which in general gives rise to non-linear behaviour. It is also possible, however, that the observed amplitude dependence results from the occurrence of other, non-linear, damping phenomena in the same region of frequency and temperature (e.g. dislocation hysteresis, section 2.2.4), or from a very large value of $(\Delta\tau/\tau)$. In Weertman's first model, the dislocations were considered to move through a lattice containing impurity atoms or rather precipitate particles. As in the theory of MOTT and NABARRO [1948] of precipitation hardening (section X–2), the dislocation lines try to "wriggle through" the precipitates. They will only be able to do this when the stress σ acting on them is large enough to bend them into loops of radius R smaller than the mean distance Λ between the particles. As R and σ are related according to the well-known formula

$$\sigma = \frac{Gb}{R} ,$$

there exists a critical stress $\sigma_c = Gb/\Lambda$, below which thermal fluctuations must aid the dislocations to overcome the obstacles. A similar activation process governs the creep of precipitation-hardened alloys. Mott and Nabarro deduced for the latter a relaxation time of the order of

$$\tau_c \approx 10^{-2} e^{H/kT} \text{ seconds.}$$

Taking $H = 1$ eV (a small value), at 400 °C the relaxation time already amounts to 10^5 sec, thus formula (29) indeed applies. It is still an open question where the small frequency factor (100) comes from. Another difficulty of this theory is that a strong dependence on impurity content and structure would be expected. This has not been studied accurately.

In Sack's model, the activation energy is that for jog (constriction) formation, which depends on the stress. Hence, a certain variation of the damping with the amplitude is present. If dislocation climb or stress relaxation over slip bands were responsible for the high-temperature internal friction, the activation energy would coincide with that for self-diffusion. Although in some experiments such a correspondence has been observed, in others it was entirely absent.

A fully satisfactory explanation of this type of internal friction has not yet been given.

BOULANGER [1949] observed a quick rise of the low frequency internal friction of cold-worked and *polygonized* aluminium at temperatures near the melting point. FRIEDEL [1957] explains this in terms of creep (notably vacancy creep, see section XII–4) of the subgrains. There is again an inverse frequency proportionality.

2.2.3. *Dislocation resonance*

Theoretically, it must be expected that near the high frequency end of the range investigated in internal friction experiments, that is, in the (ultrasonic) range from 3–300 Mc/sec, a third kind of dislocation damping shall occur, associated with the resonance motion of dislocations within their equilibrium potential valleys between atom rows. The resonance frequency of a dislocation segment of length l is in first approximation determined by the line energy; it is given by the simple expression (MASON [1955]):

$$\omega_0 = \frac{1}{l}\sqrt{\frac{2G}{\varrho(1-\nu)}}, \tag{31}$$

where ϱ denotes the density of the material. It amounts to about 1000 Mc/sec for $l = 10^{-4}$ cm in the noble metals. This frequency is barely reached experimentally, but the resonance peak will presumably be very broad (by reason of the large radiation damping), and some

absorption will already be notable at frequencies in the megacycle range. In this case, where $\omega \ll \omega_0$, expression (22) reduces to

$$Q^{-1} \to \frac{\omega\tau}{\omega_0^2} \tag{32}$$

and the damping becomes proportional to the frequency (thus the attenuation is proportional to ω^2, according to (15)). In the ultrasonic range and in the Mc range at very low strain-amplitudes, an amplitude-dependent, frequency-proportional type of internal friction has indeed been observed in slightly cold-worked metals (see the review of data by GRANATO and LÜCKE [1956]). It has a frequency-independent tail in the kc-range, which, however, may not be connected with it: it is difficult to separate there *resonance* damping from the hysteresis losses to be described in the next section. A detailed study of this effect concerns the attenuation of ultrasonic waves in deformed germanium and silicon; less detailed investigations have been carried out on the metals Pb, Cu, Al and Zn.

The resonance damping is usually found to decrease with increasing impurity content, although exceptions to this rule are known. Annealing of a cold-worked specimen results in a lowering of the damping, although it can never be fully depressed. Even light inadvertent handling already produces a considerable increase of the resonance damping.

The dependence of this type of internal friction on the temperature has not yet been well studied. It is certain that at high temperatures the damping is much stronger than at low temperatures. From the few data known (Granato and Lücke), a preliminary estimate of the energy of activation follows that lies between 0.15 and 0.4 eV in lead, between 0.6 and 0.8 eV in aluminium and is approximately 1 eV in copper. The temperature dependence seems to vary sensitively with the amount of previous cold-work.

Granato and Lücke proposed a detailed theory, in which the damped motion of dislocation segments pinned at their ends by impurity atoms is considered as the basic process. An important role in this theory is played by the distribution function of the lengths of the segments. The authors mentioned took for this a function originally proposed by KOEHLER [1952]:

$$\mathcal{N}(l)\mathrm{d}l = \frac{n^2}{N} \mathrm{e}^{-\frac{n}{N} l} \mathrm{d}l \,, \tag{33}$$

where $\mathcal{N}(l)\mathrm{d}l$ is the number of segments of length between l and $l + \mathrm{d}l$ per cm³ of the crystal, and n is the number of impurity atoms *bound* to the length N of dislocation in that volume. Formula (33) was obtained by considering the impurities to be randomly distributed along the dislocation lines (Poisson distribution).

Starting from a differential equation of the form (20), the authors derived for the damping due to one loop of length l:

$$\Delta(l) = \frac{4(1-\nu)}{\pi^3} l^2 N \frac{\frac{\beta}{m} \omega\omega_0^2}{(\omega_0^2 - \omega^2)^2 + \frac{\beta^2}{m^2}\omega^2} \tag{34}$$

(in the terminology used in section 1.4).

The decrement resulting from the complete distribution is then

$$Q^{-1} = \int_0^\infty l\Delta(l)\mathcal{N}(l)\mathrm{d}l \,.$$

In the special case that we are interested in, viz. that the resonance frequency ω_0 of all actual looplengths is much higher than the measuring frequency ω, the integral could be approximated by

$$Q^{-1} = c_1 N\lambda^4 \frac{\beta\omega}{Gb^3}\left[1 + c_2 \lambda^2 \frac{m\omega^2}{Gb^3} - c_3 \lambda^4 \frac{\beta^2\omega^2}{G^2b^6}\right]. \tag{35}$$

where c_1, c_2 and c_3 are numerical constants between one and ten, and $\lambda = N/n$ is the average loop length in cm. The variation of Q^{-1} with frequency is shown in Fig. 7. Equation (35) resembles a similar equation already derived by KOEHLER [1952].

Since with increasing cold-work N increases but λ presumably goes through a maximum, as also the number of pinning points increases, the damping is expected to go through a maximum too. Indeed, a decrease of the resonance damping after very heavy deformation has been observed. Annealing results in diminished damping, because it produces precipitation of impurities on the dislocations, therewith decreasing λ; secondly, dislocations are enabled to find more stable

positions, with higher resonant frequencies, and thirdly the dislocation density may be lowered. The effect of temperature on the damping is twofold. An increase of the temperature results in a decrease of the number of locked impurity atoms. If W is the mean binding energy of such an atom to a dislocation, the distance λ between bound atoms

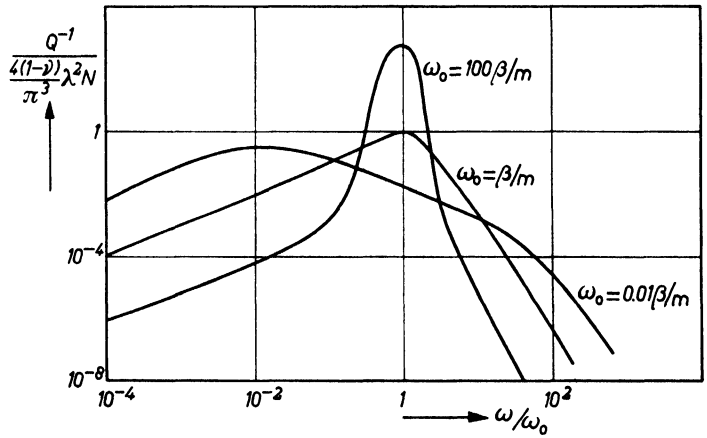


Fig. 7. Normalized frequency curves of damped dislocation resonance according to GRANATO and LÜCKE [1956]. Normalization factor is the coefficiënt of (34) for a loop of average length $\lambda^2 N$; parameter is $\frac{\omega_0}{\beta/m}$.

will, in equilibrium, be proportional to exp $(-W/kT)$. The internal friction itself should then depend on the temperature through an energy of activation of about $4W$, according to the use of formula (35). This is to be compared to the observed energies of activation, of about 0.4 to 1 eV in face-centred cubic metals, where theoretically the binding energies of substitutional impurity atoms to dislocations amount to a few tenths of an electron volt (Chapter IX).

On the other hand, a variation of λ also results in a variation of the resonant frequency. At higher temperatures the resonant frequencies will thus be shifted to lower values, which also leads to an increase of the observed damping. This effect will be especially notable in the very high frequency range. A shift of the attenuation *vs* frequency curve with the temperature in agreement with this has indeed been observed in germanium.

Although there thus seems to exist rough qualitative agreement between the experiments in the high-frequency range and the theory

of Lücke and Granato, many details remain as yet unexplained. The low-frequency tail, for instance, is undoubtedly due to a different cause. It is not as yet possible to test this theory quantitatively, as the number of parameters that can be varied arbitrarily in it is rather large. Many observations are forthcoming at the time of publication of this book, and from these the parameters may be determined.

2.2.4. *Dislocation hysteresis*

By far the most important kind of internal friction originating from the presence of dislocations in metals is the *frequency-independent, strain-amplitude dependent* type of loss (NOWICK [1954]). This occurs most conspicuously in the medium frequency range at a not too small strain amplitude. The internal friction depends again sensitively on previous cold-work. It is so strong that in many damping investigations of the past it was regarded as the *only* dislocation internal friction. It is accompanied by a strong ΔE (that is, modulus) effect. The strain-amplitude dependence points to a non-linear process underlying this type of damping. It is therefore not surprising that its explanation has been sought in terms of irreversible dislocation motion. We shall call this internal friction *dislocation hysteresis.*

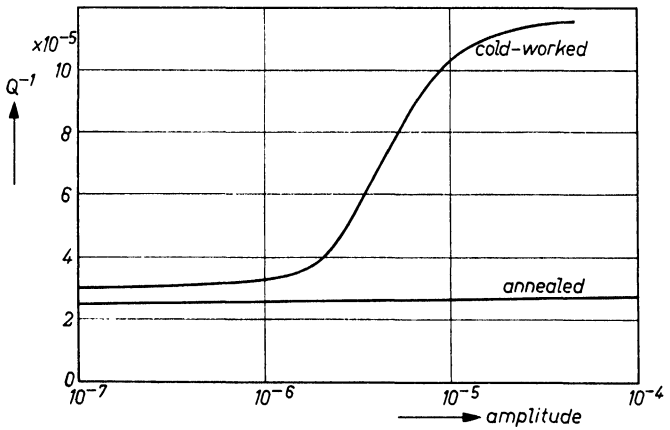


Fig. 8. Strain-amplitude dependent internal friction (dislocation hysteresis) in a slightly cold-worked aluminium crystal, compared to the damping after annealing (schematic).

A quantitative description of the amplitude dependence is difficult. Fig. 8 shows an example. Below 10^{-6} strain amplitude the damping

becomes constant, presumably because other damping causes overshadow it. At high strain amplitudes ($> 10^{-4}$) the amplitude variation again levels off. Granato and Lücke tried to fit the observed amplitude dependence below 10^{-5} by the empirical formula, confirmed later by BESHERS [1959]:

$$Q^{-1} = \frac{C_1}{a} e^{-C_2/a}, \tag{36}$$

where a denotes the amplitude and C_1 and C_2 are parameters. Stressing considerably below the yield stress already produces a large increase of the damping, such as by a factor 40 in pure copper single crystals compressed to about $^3/_4$ of the yield stress. Even ordinary handling already causes a significant raising of the damping level. In this respect the damping much resembles the resonance internal friction. Impurities have a very pronounced effect. Lead containing 0.65 at.% bismuth shows only some strain-amplitude dependent damping above amplitudes of 10^{-5}, whereas the amplitude dependence can already be detected at $a = 10^{-6}$ when only 0.05% Bi is present (see e.g. WEERTMAN and SALKOVITZ [1955]). At a given strain amplitude the damping depends exponentially on the concentration of impurities, which means that in (36) C_2 is to be taken proportional to this concentration.

Irradiation with neutrons or fast electrons, or quenching from a high temperature, reduces and sometimes even completely suppresses dislocation hysteresis (THOMPSON and HOLMES [1956]). This is probably due to the formation of vacancies and interstitials near dislocations which immobilize these dislocations. The damping returns only after annealing at a fairly elevated temperature, between 200 °C and 400 °C in copper (stages IV and V of Chapter XV). Annealing at a lower temperature only accelerates the rate of suppression of the damping. Neutron irradiation can already suppress the damping at helium temperatures however (BLEWITT [1957]); only small irradiation doses (10^{12}–10^{15} neutrons/cm^2) are needed for it. All these experiments concern the change of modulus rather than the actual damping.

The dependence of the damping on temperature is again through an energy of activation, which is rather small, however, being of the order of or less than 0.1 eV in copper and aluminium. No definite dependence of this energy on the previous amount of cold-work has been observed.

According to Lücke and Granato and others, dislocation hysteresis

is due to the stress-induced break-away of dislocation segments from impurity atoms to which they were locked, followed by the irreversible collapse of the long loops thus formed (Fig. 9). Once a few neighbouring loops have been liberated, longer pieces, of dislocation

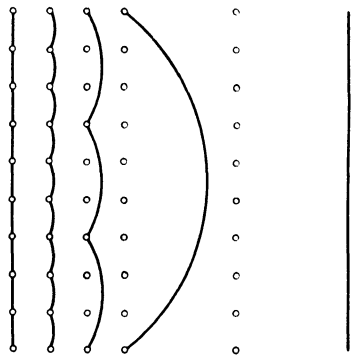


Fig. 9. Successive stages in the irreversible break-away of dislocation segments from impurity atoms.

become unstable in the applied stress field. An impurity atom is bound to a (long straight) dislocation by an attractive force which, in the first order elastic approximation of Cottrell (see Chapter IX) and neglecting orientation dependence, is:

$$f_b \approx \alpha G \eta b^2 ,$$

where η is the relative dilatation around the foreign atom and α is a numerical constant of order 10. The force exerted by a shear stress τ along the Burgers vector on a piece of dislocation of length l is in the same approximation (section IV–7)

$$f_\tau \approx \tau b l .$$

Consider now a dislocation consisting of segments of length l_1, l_2, l_3,... (Fig. 10). The dislocation will break loose from the impurity atom at A between the segments l_1 and l_2 when there $f_\tau > f_b$, thus when:

$$\tau > \frac{\alpha G \eta b}{l_1 + l_2}.$$

Conversely, break-away in a given field of stress will occur from those atoms only that connect dislocation segments of lenghts l_1 and l_2 such that

$$l_1 + l_2 > \frac{\alpha G \eta b}{\tau} = l_m \,. \qquad (37)$$

The break-away gives rise to an irreversible strain, and thus to hysteresis in the sense of section 1.3.

In practice, the process is more complicated. When (37) is fulfilled, the loop will not only break loose at A, but afterwards successively at all other locking points, as certainly $l_1 + l_2 + l_3 > l_m$, etc. The process will thus presumably be a *catastrophic* one.

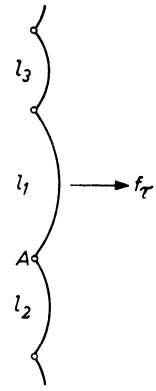


Fig. 10. Break-away of dislocations from locking points.

A dependence of the damping on the strain (thus on the stress) amplitude follows immediately. The larger the shear stress τ, the smaller the limiting length l_m, and progressively shorter loop lengths can contribute to the irreversible strain. For a mathematical description of the process the reader is referred to the paper by Granato and Lücke. The result for the *irreversible strain* can be written as follows

$$\varepsilon_{irr} = g(\tau) e^{-\frac{\alpha G \eta b}{\lambda \tau}} \,. \qquad (38)$$

Here λ is as before the mean length of dislocation between impurities in the non-stressed situation. The rate of damping, Q^{-1}, can be found from this formula by evaluating the integral

$$\int_{-\tau}^{+\tau} \varepsilon_{irr} d\tau \,.$$

An exponential expression like (38) of course results. As the stress τ occurs in the exponent, the exponential amplitude dependence of (36) indeed finds an explanation. Also the proportionality of the constant C_2 in (36) with the impurity content emerges, namely through the factor λ in the denominator of the exponent.

The temperature dependence first enters in the same constant C_2, namely again through λ, and to a lesser extent also in C_1, since this, like $g(\tau)$ of (39), also varies with λ and with l_m (37). The exponential dependence dominates, of course. The binding energy of an impurity atom to a dislocation is found from these damping observations to be about 0.1 eV in face-centred cubic metals. This is of the same order as followed from the resonance damping.

Excess impurities and physical defects (vacancies, interstitials) produced by irradiation or quenching can block the dislocation movement very effectively. From the observed decrease of damping with impurity content it follows that to suppress it completely, 10^{15} to 10^{16} impurities per cm^3 are needed. This corresponds to a full occupation of all available dislocation sites. The observations of THOMPSON and HOLMES [1956] on the necessary neutron dose to suppress the damping confirm this. Seeger (private communication) has objected to the basic principles of the hysteresis model used here, on the ground that a unidirectional elastic prestress would be expected to affect the break-away process in an unsymmetrical way, whereas the contrary has been observed. However, it is not quite certain that such an effect would indeed show up in such simple measurements of the energy dissipation as have been made so far. Although the motion of dislocation segments indeed becomes asymmetrical in time, the energy dissipation per cycle remains in first instance unaltered. The break-away stress is asymmetrically effected, though.

2.2.5. *Transient cold-work internal friction*

By the name *transient (cold-work) internal friction* we shall denote the increase of the low- and medium-frequency internal friction that occurs immediately after moderate deformation of a metal at not too high a temperature. This damping shows already appreciable recovery when the metal is kept at a temperature far below that of mechanical recovery or recrystallization. The best studied examples are α-iron and β-brass (FUSSFELD [1951]) and also rocksalt (GORDON and NOWICK

[1956]). Pure cold-drawn α-iron wires show an internal friction increase of about a factor 2 at 1 c/sec. After storing the metal for a few hours at 20 °C, this increase has practically disappeared (KÖSTER [1940b]). Similar observations have been made on wires of copper and silver. Aluminium formed an exception in so far that the recovery took place only at a relatively elevated temperature. In NaCl this damping was studied as a function of irradiation with X-rays (GORDON and NOWICK [1956]), see also Chapter XXVII). An exponential recovery rate was observed.

The transient damping in metals has been observed in the 1 c/sec range (KÊ [1950b], KÖSTER *et al.* [1940]) and in the 10 kc/sec range (KÖSTER [1940a]); it seems to be fairly independent of frequency. A definite amplitude dependence is absent, but accurate measurements on this point have yet to be carried out. The damping increases with the amount of previous cold-work at a less than linear rate. The temperature dependence of the damping is difficult to evaluate, as the recovery becomes very rapid at elevated temperatures, and measurements below room temperature have not yet been performed. The estimated energy of activation for the annealing process is about 0.9 eV in copper, and about 1 eV in aluminium, but only 0.5 eV in β-brass and even less in α-iron ($\leqslant$ 0.4 eV). In the f.c.c. metals these energies resemble in magnitude activation energies for the diffusion of point defects (vacancies) in the metals quoted. The transient damping is presumably caused by the motion of the dislocations, present immediately after the cold-work in irregular arrays. The recovery would then be due to the gradual transfer of these dislocations to more stable equilibrium positions, as a result of the diffusion of point-defects to them. The very rapid annealing in the b.c.c. metals like β-brass and α-iron as compared to the same process in f.c.c. metals could then be understood, because dislocation climb is presumably much more easy in these metals than in close-packed materials.

GRANATO, HIKATA and LÜCKE [1958] have recently attempted to explain quantitatively the transient internal friction along the lines here mentioned, thus as a kind of *strain ageing*. They used the model already discussed in section 2.2.4 to explain dislocation hysteresis, and studied the change with time in the distribution function of loop lengths, and thus of the damping, owing to the migration of point defects to dislocations. Only elastic interaction is considered and formula X–(4) is applied to describe the accumulation of point defects in

the dislocation lines. The function $g(\tau)$ of (38) is shown to change according to a $t^{2/3}$-law which is in fair agreement with the observations on rocksalt (GORDON and NOWICK [1956]). Irradiation of the latter material produces an increased recovery rate, which could be well described by Granato *et al.*:

$$Q^{-1} \sim a \exp\left[-b(1 + t^{2/3})\right].$$

For a further discussion the reader is referred to the original papers.

NOWICK [1955] has proposed a different less quantitative model, in which the diffusion of point defects does not lead to pinning, but to spatial rearrangement of the dislocations; a typical *recovery* phenomenon more in keeping with the suggestion that *climb* is responsible for the decay of the damping.

2.2.6. *Deformation hysteresis*

A type of internal friction that bears resemblance to transient cold-work internal friction and to dislocation hysteresis (possibly the effects are but different forms of the same phenomenon, occurring under different circumstances) occurs at low frequencies and large strain amplitudes (1 c/sec at $a \gg 10^{-5}$) in slightly cold-worked metals (Fe: FAST and VERRIJP [1961], Cu: DARLING [1957]). Fast and Verrijp observed the strongly amplitude dependent damping of 0.1% deformed pure iron wires. Storing the metal at room temperature resulted in a rapid lowering of the internal friction. This behaviour is quite comparable to that of the transient internal friction discussed in the previous section. However, the two types differ in their amplitude dependence as well as in the fact that the latter type of damping has only been observed after moderate or heavy deformation. In view of the strong amplitude dependence, we propose the name *deformation hysteresis* for this type of internal friction.

The amplitude dependence can be described as quadratic; the damping tends to saturate at an amplitude of about 10^{-4} (Fig. 11). The dependence on the amount of previous cold-work has not yet been determined. There exists some variation with the grain size. Most interesting is the variation of the deformation hysteresis with the temperature of deformation and of measurement respectively. Wires deformed and measured at −50 °C showed but half as much damping as the same wires deformed and measured at 20 °C (Fig. 11). On storing

the latter wires at 20 °C, the damping decreased in about 100 hours to half its value (Fig. 12). On storing the former wires at −50 °C, the damping (at −50 °C) actually *increased* with time (10% after 7 hours). Wires deformed, measured and stored at −30 °C showed neither increase nor decrease of the internal friction. Deformation and measurement at −100 °C resulted in practically no damping at all. Wires

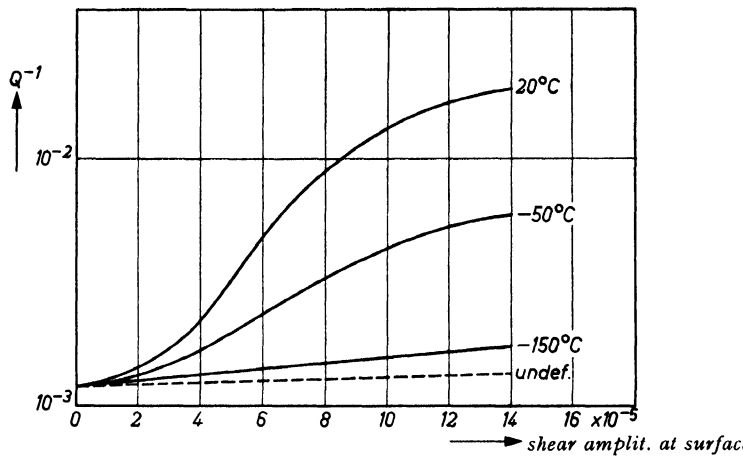


Fig. 11. Deformation hysteresis at 1 c/sec of pure polycrystalline α-iron, 0.1% plastically strained, as a function of amplitude and temperature of measurement.

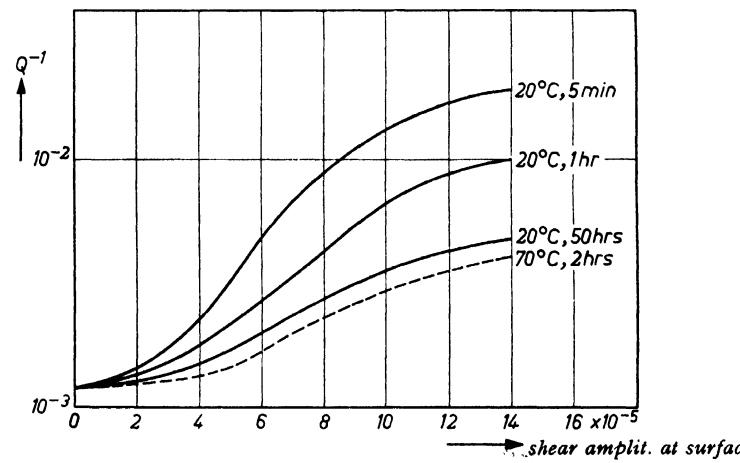


Fig. 12. Influence of annealing, at 20 °C and 70 °C respectively, on the deformation hysteresis of iron measured at room temperature and 1 c/sec.

deformed at 20 °C and measured within half an hour at −50 °C again showed a lowering of the damping level by a factor of 2 compared to the 20 °C damping. Now, however, on storage at −50 °C, the damping *de*creased a little more (Fig. 13). Evidently, the increase observed in the experiment: "deformed and measured at −50 °C", had now already taken place very quickly during the room temperature period. It was observed quite generally that deformation and measurement at a low

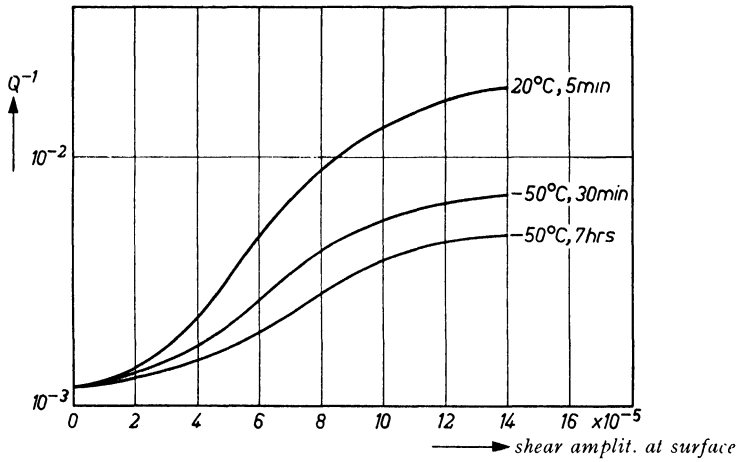


Fig. 13. Deformation hysteresis at 1 c/sec of iron wire deformed 0.1% at 20 °C, after storing at −50 °C for various times.

temperature, followed by a second measurement at a higher temperature, resulted in an instantaneous increase of the damping in the latter case, followed first by a slight gradual further increase and afterwards by a gradual decrease. All processes occurred more rapidly the higher the temperature. Deformation at a low temperature followed by a short higher temperature "anneal" thus resulted in an *in*crease of the low-temperature damping.

When quenched from 800 °C in water and deformed and measured afterwards, the wires show a much faster decrease of the damping on storage than unquenched wires. This shows that excess vacancies promote blocking of the damping agent. Solute carbon atoms have a similar effect.

It will be clear that we must distinguish between three processes.

1) The damping process itself, whatever it may be, evidently takes place at a stronger scale at higher temperatures. Whether here an

energy of activation occurs, or whether the temperature dependence is only associated with the change of the elastic constants, thermal expansion, etc., is not known. If an activation energy is present, it must be rather small, of the order of 0.01 or 0.02 eV.

2) A release process which results in a gradual increase of the internal friction. The effectiveness of this process may be a function of the temperature, its rate certainly is. The time necessary to achieve maximum damping depends on the temperature presumably through an activation energy of about a few tenths of an eV. We evidently have to do here with a *gradual release* of the damping agent.

3) A recovery process, which is equivalent to a *gradual blocking* of the damping agent. It is characterized by an energy of activation of about 0.5 to 0.6 eV, as follows from the curve of Fig. 12.

The interpretation of the observed effect as due to irreversible motion of dislocations, formed in the material during cold-work, is rather obvious. To explain the gradual loosening, it might be assumed that at low temperatures the dislocations are piled up against some barriers and cannot contribute to the damping before they are again dispersed—a process which is more easy at higher temperatures. The decrease can then be ascribed to pinning of the dislocations by point-defects or, alternatively, to climb motion to more stable equilibrium positions, made possible by the diffusion of point defects to the dislocations. A $t^{\frac{1}{3}}$-law should then describe the time dependence of the damping recovery, and this has indeed been observed.

2.3. INTERNAL FRICTION NOT EXCLUSIVELY ASSOCIATED WITH DISLOCATION MOTION

A number of internal friction processes exist that are probably connected with dislocation movement, but the mechanisms of which are not yet clear. They occur mostly in the extreme frequency- or temperature-ranges, have been little investigated and can not yet be classified as pure relaxation, hysteresis or resonance phenomena. It is not even possible to distinguish rigorously between the various categories, since observational data on temperature, amplitude and frequency dependence are scarce.

2.3.1. *Visco-elastic after-effect and microcreep*

A kind of elastic after-effect that bears a superficial resemblance to the high-temperature dislocation relaxation (section 2.2.2), owes its occurrence to slow, thermally activated motion of groups of dislocations under stress in slightly contaminated metals. The phenomenon occurs at very low frequencies or static loading and elevated temperatures, in contrast to high-temperature relaxation that is found at low and medium frequencies. It is but *little* sensitive to cold-work. The character of this *visco-elastic after-effect* resembles much that of creep; it has been termed *microcreep* in the case of static loading (COTTRELL and JASWON [1949]). Dislocations having an atmosphere of impurity atoms around them (see section IX–3) are retarded in their motion; and a phase lag occurs between stress and dislocation strain. The relaxation time governing this retardation is the jumping time of the impurity atoms (in the stress field of the dislocation); the activation energy is that for their diffusion ($\geqslant 1$ eV usually).

PEARSON and ROTHERHAM [1956], in the course of their measurements, at 1 c/sec and elevated temperatures, of the grain boundary internal friction of copper and silver (section 2.3.3), observed a background effect that can be interpreted as visco-elastic damping. It was independent of previous cold-work, but decreased with increasing grain size of the specimens. At the frequency 0.4 c/sec the damping was somewhat higher than at 1.5 sec. The authors themselves interpreted the effect as a kind of grain boundary viscosity, in view of the approximate equality of the activation energies governing the temperature dependence of the grain boundary damping and of this background damping (section 2.3.3). No systematic measurements of the visco-elastic damping at very low frequencies exist as yet.

2.3.2. *Temperature-activated hysteresis*

At elevated temperatures the low frequency damping of copper, aluminium (and many other metals) was observed to increase rapidly (KÊ [1947]), the increase being governed by an activation energy of 0.25 eV and 0.35 eV respectively (Fig. 14). BORDONI and NUOVO [1954] observed a similar phenomenon in lead stressed in the Mc range. One has, obviously

$$Q^{-1} = c\mathrm{e}^{-U/kT} \; ; \tag{39}$$

U is 0.15 eV in lead. The process occurs also at ultrasonic frequencies and is clearly independent of the frequency. No definite influence of previous plastic deformation has been detected. MASON [1955a]

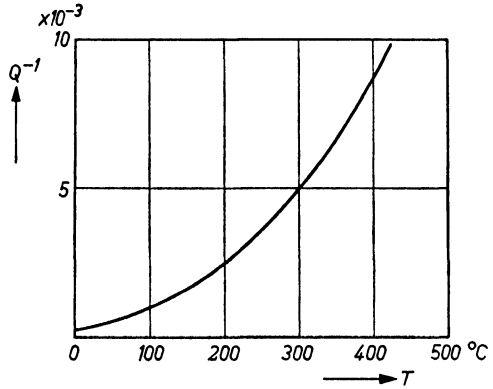


Fig. 14. Temperature-activated hysteresis in an aluminium single crystal stressed at 0.8 c/sec (after KÊ [1947]).

named the effect *temperature-activated hysteresis*, although we may have to do with a more general viscous phenomenon in view of the absence of a pronounced amplitude dependence. Its cause is not known. Mason proposed a mechanism rather resembling that of dislocation hysteresis. However, as the strong amplitude dependence expected for this is absent, this explanation appears improbable. The common occurrence of the effect seems to point to a general property of the metal instead of to dislocations as the cause of the effect. This internal friction should be clearly distinguished from visco-elastic after-effect, as the temperature dependence is much more rapid in the latter case.

2.3.3. *Grainboundary relaxation*

Comparison of the 1 c/sec damping of single and polycrystalline aluminium (KÊ [1947]), revealed the presence of a relaxation process in the latter material near 300 °C, that was absent in the former (Fig. 15). In many other metals the same difference has been noted. Simultaneous temperature and frequency variation together with the absence of amplitude dependence point unambiguously to a relaxation mechanism that is obviously related to grainboundaries. It has been

called *grainboundary relaxation.* In Al, the activation energy governing the peak frequency variation with temperature amounts to 1.4 eV, which is close to the energy of self-diffusion in this metal. However,

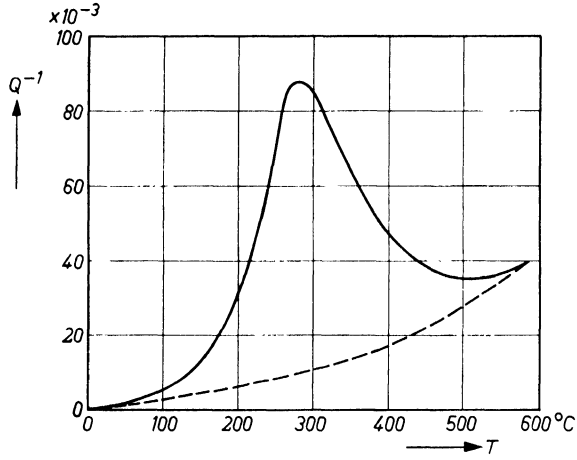


Fig. 15. Grain boundary relaxation in polycrystalline aluminium at 1 c/sec, compared to single crystal damping (dotted curve, see fig. 14). (After Kê [1950]).

PEARSON and ROTHERHAM [1956] found the grainboundary relaxation in *pure* copper and silver to be characterized by an energy of activation much *smaller* than that of self-diffusion (by about a factor 2), so that it becomes comparable to that for the visco-elastic after-effect (section 2.3.1). In *impure* copper and silver, however, the activation energies for damping and self-diffusion were again about equal, *independent* of the nature of the alloying element.

The correlation with the phenomenon of grainboundary self-diffusion (section XVIII–5.1) lies at hand: this is also characterized by an energy of about half the volume self-diffusion. It therefore seems warranted to explain the internal friction as a sort of stress-induced grainboundary self-diffusion.

Various authors have tried to explain this relaxation in terms of grainboundary *slip.* In Kê's original paper it was assumed that the matter in a narrow zone near the grainboundary is in a disordered state, and that under influence of the stress the atoms slide past each other. The activation energy for such a process would evidently be the same as that for grainboundary self-diffusion. MOTT [1948] suggested

that a grainboundary surface should contain "islands" of good fit between the adjacent crystals, separated by lineshaped regions of bad fit. This model is more in agreement with the modern dislocation model of a grainboundary. The elementary slip process in the boundary is a shear motion of the two crystal surfaces constituting one of the islands of good fit, accompanied by a disordering of the atoms around the island. The activation energy for this process should be proportional to the number n of atoms contained in the island, and it should depend on the applied stress σ and on the reduced temperature T/T_m. The energy of activation must namely be zero at the melting point T_m and equal to nL_F at the absolute zero, where L_F is the latent heat of fusion. Mott concluded to the following approximate expression for the activation energy U:

$$U = nL_F\left(1 - \frac{T}{T_m}\right)\log\left(\frac{\sigma V}{kT}\right), \tag{40}$$

where V is the molar volume.

A direct correspondence between U and the diffusion energy does not exist in this theory.

The natural variation in the nature and dimensions of the disordered regions in Kê's model as well as that in Mott's, leads to a fairly wide distribution of relaxation times, and indeed, the grainboundary peak is rather broad. However, the viscosity of the grainboundary matter in Kê's model is much too small to be accounted for by the presence of molten zones.

The effect of impurities on the grainboundary internal friction are not yet understood either. Their effect on boundary *diffusion* might, from a theoretical point of view, be very well much greater than on *slip*. ZENER [1941, 1958] has developed a theory of stress relaxation across viscous matter in grainboundaries (viscous grainboundary slip), and derived quite generally for the ratio between relaxed and unrelaxed moduli the value $3(1 + \nu)/2(3 + \nu)$, which is about 0.60 for $\nu = \frac{1}{3}$. The ratio's observed in practice vary between 0.6 and 0.8.

2.3.4. *Point defect internal friction*

HASIGUTI and coworkers [1953] have observed the occurrence of a number of relaxation type internal friction peaks in a great variety of low-temperature cold-worked metals. These peaks disappeared on

room temperature annealing, and their occurrence was found to be fairly universal, in as far as crystal structure, polycrystallinity etc. did not affect them. Only care had to be taken to deform the metals in question (iron, nickel, copper, silver, aluminium, magnesium, titanium) at sufficiently low temperatures, where no diffusion of point defects can take place (e.g. the temperature of liquid air). On annealing, the two most important damping peaks gradually disappear in stages, comparable (also numerically) to the recovery stages of the electrical resistivity discussed in Chapter XV. This led the discoverers to the conclusion that the damping should be due to the stress-induced motion of the same point defects which are responsible for the resistivity increase. As the symmetry of the strain field of a single defect in a cubic metal is to high to allow stress-induced migration, an explanation in terms of divacancies and other combined defects must be invoked. Indeed, the activation energy describing the shift of frequency with temperature of one of the peaks in copper was found as 0.5 eV, which may well be the diffusion energy of a divacancy (see section XV–5). Further, the activation energy for relaxation of any peak was always equal to that for recovery.

A remarkable feature of this complex of damping phenomena is that the values of the activation energies of the two most important peaks were found to be approximately independent of the metal studied.

For a relaxation effect due to *interstitials* we refer to section 2.4.2.

2.4. INTERNAL FRICTION RELATED TO THE PRESENCE OF FOREIGN SOLUTE ATOMS IN THE METAL

A particularly important group of internal friction phenomena is associated with the stress-induced diffusive motion and ordening of solute impurity atoms in crystals. A large literature already exists on this topic, and as we are primarily interested, not in chemical, but in physical lattice imperfections, we shall severely restrict the discussion here.

2.4.1. *Snoek-damping in body-centred cubic metals*

In α-iron containing carbon or nitrogen in *solid solution* (as obtained after quenching), a typical relaxation peak occurs near room temperature at frequencies around 1 c/sec (Fig. 16). The explanation was first given by SNOEK [1941]. The interstitial carbon or nitrogen atoms can

occupy three distinct types of octahedral positions, each characterized by a local tetragonal deformation of the lattice. The directions of tetragonality of the three positions are mutually orthogonal and coincide with the cube directions. Application of a unidirectional stress favours the occupation of those sites whose tetragonality matches best

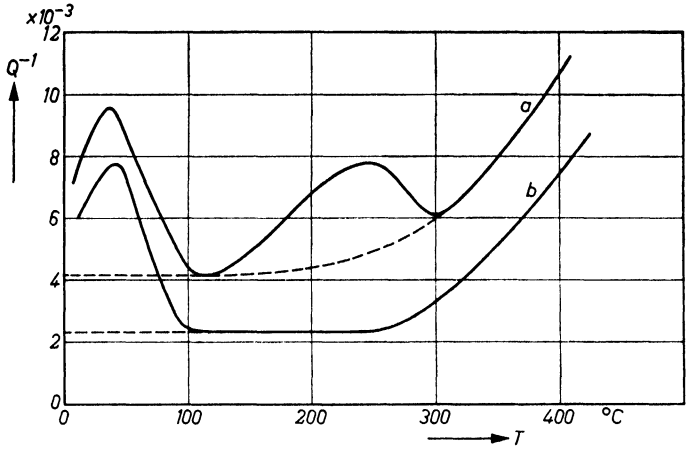


Fig. 16. Snoek damping (near 40 °C) and Köster damping (near 230 °C) in ferrite containing carbon. The Köster damping occurs only in the deformed specimen *a*, the Snoek damping is also present in non-strained metal (*b*). The dotted parts refer to *pure* iron. The increase of the Snoek damping after deformation is only apparent and is actually due to a rise in background; the peak height remains approximately the same. The increase of damping at high temperatures is due to the effects discussed in sections 2.2.2, 2.3.1, 2.3.2 and 2.3.3. (After KÖSTER *et al.* [1954].)

with the overall strain in the lattice. Preferential diffusion of the solute atoms will thus take place to these sites, resulting in an elastic after-effect. As only one atomic jump is needed to change from one type of site to another, the relaxation time for this process will be of the order of the inverse jumping frequency of the interstitial atoms. This is indeed about 1 per second at room temperature in the case considered. No dependence at all exists on the strain amplitude, and the activation energy determining the frequency variation of the relaxation peak with temperature is that for interstitial diffusion of the solute atoms. It is about 1 eV for C and N in α-Fe.

It should be mentioned here that the study of this kind of internal friction yields one of the most accurate methods of determining the diffusivity of interstitial atoms in b.c.c. metals in the room tempera-

ture region, where ordinary diffusion techniques would require millions of years.

The height of the damping peak, $Q^{-1}{}_{max}$, was observed to be strictly proportional to the solute content and thus serves as a measure for the solubility at room temperature. The solubility at other temperatures can also be determined in this way, namely by measuring the Snoek damping of *quenched* materials.

The Snoek damping is independent of the state of cold-work. It occurs only in b.c.c. and less symmetric materials, for in f.c.c. crystals the interstitial atoms produce no tetragonal distortions. Much experimental work has been done on iron and iron alloys by WERT [1950], FAST *et al.* [1950, 1953, 1954], DIJKSTRA [1949] and many others, and by KÊ [1948] on tantalum and its alloys. POWERS and DOYLE [1956] observed very interesting effects when oxygen and nitrogen were present together in tantalum. A number of damping peaks corresponding to *pairs* of interstitial atoms is then observed, a discussion of which has been given by MEIJERING [1957].

A very important application of this damping method is the measurement of the effect of cold-work in the precipitation of carbon and nitrogen in α-iron (strain-ageing, section X–1). HARPER [1951] was able to show, by measuring the decrease in height of the damping peak as a function of time and temperature, that the solute atoms diffuse to dislocations and form "preprecipitates" there. This diffusion process should be governed by a time dependence of the form $(Dt)^{2/3}$, which is typical for dislocation-impurity interaction (see, however, Chapter X). To name but one other example, WERT [1954] studied by internal friction methods the mutual influencing of the rates of precipitation of carbon and nitrogen in α-iron.

2.4.2. *Snoek damping in face-centred cubic metals and alloys*

SEEGER *et al.* [1960] pointed out that an interstitial in the *interstitialcy* configuration (see section 2.2), owing to its tetrahedral symmetry can cause a Snoek type relaxation even in f.c.c. metal lattices, whereas a normal, cubic, interstitial can not. They could indeed observe the relevant damping peak in cold-worked or n-irradiated nickel: $Q = 0.8$ eV.

ROZIN and FINKELSTEIN [1953] observed at 1 c/sec a damping peak in an austenitic 25–20% chrome-nickel steel containing some carbon, at about 300 °C. They explained this as due to the jumping of carbon

atoms between octahedral interstitial positions surrounded by different arrangements of the various substitutional metal atoms. KÊ and TSIEN [1956] studied the same phenomenon in γ-iron stabilized by the addition of 10% or more manganese. Pairs of C-Mn atoms will occur, and these rotate under the influence of a periodically varying stress.

A very interesting sequel to these experiments form the observations of KÊ, TSIEN and MISEK [1955] on *pure* nickel with some carbon, of KÊ and YANG [1957] on γ-Fe with only 1.7% manganese, and by WU and WANG [1958] in f.c.c. Fe-Ni alloys containing some C. In all cases a strong damping due to the carbon was found at about 240 °C, and this can no longer be explained at all by the theory of "alloy pairs". The peak in the latter case shifted with frequency according to an activation energy of 1.5 eV, its height was approximately proportional to the square of the carbon content. It decreased in height after annealing at 300 °C. The authors interpreted this in terms of rotating *pairs of carbon atoms*, that had associated because they were both attracted to a lattice vacancy. The activation energy would then be of the order of the diffusion energy of carbon in austenite, whereas the binding energy between a C-atom and a vacancy is of the order of 0.1 eV only. An extrapolation from high temperature diffusion data leads to reasonable agreement.

2.4.3. *Stress-induced ordering in substitutional solid solutions*

Substitutional solid solutions not only can show internal friction of the kind discussed in the previous section, thus when they also contain foreign atoms in interstitial solution; the presence of two (or more) different types of atoms occupying lattice positions only can already give rise to stress-induced ordering, by reason of the asymmetrical distortions around, for instance, a pair of neighbouring-like atoms. When the stress is periodic relaxation may again occur and an internal friction peak may be observed. The damping peak, in the typical case of Ag-Zn, lies near 200 °C at 1 c/sec. The relaxation time and its temperature variation depend intimately on the diffusivities of the two components (NOWICK [1953]).

In general, diffusivities computed from this kind of internal friction in substitutional alloys differ from those determined by other means, in that the energies of activation come out significantly too *low*. A very detailed discussion of the experimental and theoretical aspects of

stress-induced ordering and relaxation has been given by CHILDS and LE CLAIRE [1954] and LE CLAIRE and LOMER [1954] respectively.

Stress-induced ordering internal friction in substitutional alloys can bring to light a difference in diffusivity of the two components of an alloy, when this exists. Thus it gives additional information on the *Kirkendall effect* (Chapter XVIII).

Further, in cases where self-diffusion data are known, internal friction results can aid to separate the contributions from the formation and the migration respectively of vacancies to the self-diffusion energy, as the damping depends only on the migration energy. The system silver-zinc has been investigated in this respect (see NOWICK [1953]).

The damping will, of course, differ according to the state of order that is already present in the alloy. GOERING and NOWICK [1958] showed the relevant damping in Cu_3Au to be more than 5 times smaller in the ordered state than in the disordered state.

2.4.4. *Köster damping in interstitial solid solutions*

Already in Snoek's original paper (SNOEK [1941]) on the internal friction of ferritic iron, it was mentioned that, apart from the room-

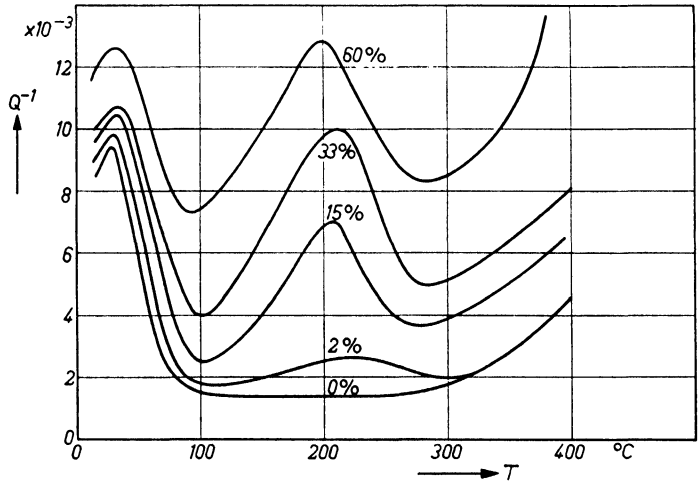


Fig. 17. Influence of pre-strain (expressed in % elongation) of the low-frequency damping of deformed (ferritic) steel wires. (After KÖSTER *et al.* [1954]).

temperature damping, another maximum occurred near 200 °C (Fig. 16). This kind of damping has been extensively studied by Köster

and his associates (see e.g. KÖSTER, BANGERT and HAHN [1954]), and we denote it by the name *Köster damping*. It is again a relaxation effect, and a close correlation exists between the Snoek- and the Köster damping. However, whereas the former is practically insensitive to plastic deformation (apart from trivial effects such as strain ageing), the latter *only* occurs in plastically deformed material, and increases with the amount of cold-work (Fig. 17). This increase can, when due account is taken of the simultaneous rise of the background damping level, be very well represented by a square root relation:

$$\tan \delta \sim \varepsilon^{\frac{1}{2}}, \tag{41}$$

where ε stands for the plastic strain.

The saturation value of the Köster damping as a function of the concentration of C and N also increases with the amount of cold-work. Formula (41) has been derived while correcting for this secondary effect.

Annealing of a previously cold-worked metal results in a decrease of the Köster damping. This is partially due to a decrease in the concentration of solute atoms, and partially to a decrease of dislocation density (the latter effect is predominant at higher annealing temperatures, the former in the lower temperature region).

The Köster damping is, like the Snoek damping, characterized by an energy of activation. This was found to be 1.6 eV in iron containing only nitrogen, and slightly lower in iron containing both carbon and nitrogen. (The activation energy for Snoek damping is only about 1 eV.)

The elementary process underlying the Köster damping is not known at present. Undoubtedly it is due to interaction between the interstitial solute atoms and the physical imperfections generated by the plastic deformation. It seems probable at first sight that this interaction is between dislocations and solute atoms. There are some objections to this, however. Inspection of the effects of annealing on Snoek (S) and Köster (K) damping simultaneously, shows that (Fig. 18):

After annealing between 100–200 °C for 1 hour: S disappears,
K not affected.
,, ,, ,, 200–300 °C ,, 1 ,, : S remains absent,
K decreases to half.
,, ,, ,, 300–400 °C ,, 2 ,, : S returns to $\frac{1}{3}$,
K disappears.

After annealing between 400–500 °C for 1 hour: S returns to $\frac{1}{2}$,
K remains absent.
,, ,, ,, 500–600 °C ,, 1 ,, : S returns in full,
K remains absent.

Evidently, the quenched-in solute atoms responsible for the Snoek damping precipitate between 100 and 200 °C, where dislocations undoubtedly play a role. However, no effect is observed on the Köster damping. The recovery of the Snoek damping on the higher anneal is retarded by a few hundreds of degrees, compared to the decrease of the Köster damping.

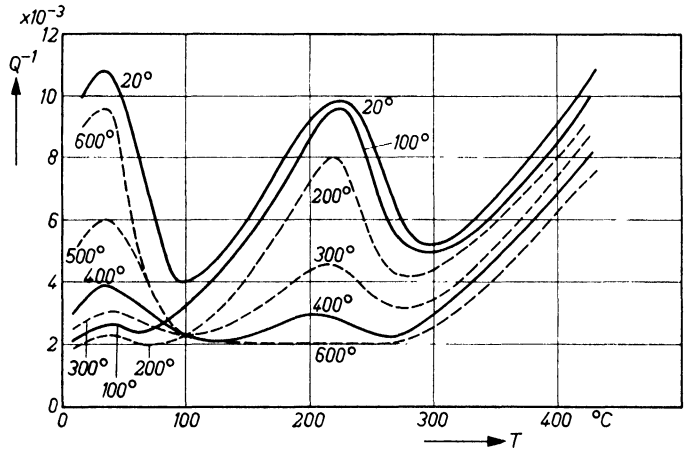


Fig. 18. Effect of annealing (for 1 hour) at various temperatures on the Snoek and Köster damping of deformed ferritic steel wires. (After KÖSTER *et al.* [1954].)

After the 500 °C anneal the Snoek damping has only returned to half of its original value, whereas the Köster damping has already disappeared to within 10%. Evidently the blocked atoms can only for about one fifth of their number contribute to the latter. This may be due to a decrease of the dislocation density, although a large decrease of the Köster damping already occurs at a temperature at which the dislocation density is not expected to change much (300 °C).

Recrystallization removes the Köster damping and causes the Snoek damping to return in full.

Köster originally proposed the stress-induced diffusion of the solute atoms captured in the stress field of the dislocation, as the cause of the phenomenon. One would then expect (cf. BERGHOUT [1958]) the acti-

vation energy (about 1.5 eV) to be less than that for diffusion in the perfect lattice (1 eV), instead of greater, as observed.

The Köster damping might also be due to the interaction between vacancies and solute atoms, more or less of the kind as suggested by Kê as taking place in face-centred cubic γ-iron also (section 2.4.2). The close agreement between the numerical values of the energies of activation may not be accidental. Vacancies and solute C or N-atoms undoubtedly attract each other, and composite lattice defects such as C-C-vacancy agglomerates may very well be stable. These produce non-cubic lattice strains and thus can cause damping. The observed activation energy may then be associated with diffusion of vacancies instead of with carbon or nitrogen diffusion. Annealing at a fairly low temperature suffices to dissolve the agglomerates, owing to the small binding energy (0.1 eV).

2.4.5. *Internal friction due to phase transformations*

A large number of damping phenomena find their cause directly or indirectly in internal transformations. Thus, for example, in the neighbourhood of precipitation zones or pre-precipitation zones (dislocations), stress-induced ordening may take place, giving rise to a relaxation type of internal friction. The amount of damping depends usually on the temperature; it is largest in that range of temperature where the precipitation process occurs. Plastic deformation, by its ability to provide new precipitation nuclei, greatly enhances this type of internal friction, which is in most metals observed in the 1 c/s range at temperatures between 200 °C and 400 °C (see NOWICK [1953]).

Another type of internal friction belonging to this category is the damping associated with the order-disorder transitions in ordering alloys (Chapter XXI). Near to the transition temperature for ordering, the damping due to stress-induced ordering may reach extremely high values (in the alloy CuAu, for instance). A third cause of internal friction is connected with diffusive or martensitic phase transformations proper. This damping is generally characterized by a typical temperature variation, that follows closely the variation of the thermodynamic quantities governing the transformation. Examples have been given by SCHEIL and others [1944] for the $\alpha - \gamma$-transition in iron alloys and by KÊ and MA [1956] for the martensitic transformation. The influence of a previous plastic deformation is usually large.

2.5. THERMAL, MAGNETIC AND ELECTRONIC DAMPING IN METALS

For the sake of completeness, we mention here in addition a number of internal friction phenomena that do *not* depend on physical lattice defects, but that are caused by the specific thermal, magnetic or electric properties of metals.

2.5.1. *Thermoelastic damping*

One of the troublesome background effects that limit the accuracy of low-frequency damping experiments is the thermoelastic damping. It is caused by temperature gradients induced in the metal when it is inhomogeneously deformed, owing to the thermoelastic effect (the "reverse" of ordinary thermal expansion). The temperature gradients can only be equilibrated by a flow of heat, which results in a further deformation. We have here, evidently, a typical after-effect, the relaxation time of which is governed by the heat conductivity $\varkappa$ of the material and by its relevant dimension d. One has:

$$\tau = \frac{d^2 c^2}{\pi^2 \varkappa}, \tag{42}$$

where c is the heat capacity per unit volume.

In polycrystalline metals the fluctuations between the stress levels in the various grains already cause thermoelastic damping if the applied stress itself is homogeneous; in single crystals the effect is absent in this case. It exists in full force, however, in bending experiments. The typical frequency range where it occurs lies between 1 and 1000 c/s; it only varies slowly with the temperature (through $\varkappa$).

2.5.2. *Magnetoelastic relaxation and magnetic aftereffect (in ferromagnetic materials)*

Eddy currents set up in a periodically stressed material through the intermediate of magnetostriction also cause energy dissipation and therefore damping. The relaxation time of this after-effect is the time necessary for the eddy currents to "diffuse" gradually inwards, thereby allowing the local magnetization to relax and the elastic deformation to adjust itself. A distribution of relaxation times exists, all of which are, inter alia, determined by the quantity $d^2\chi/\varrho$ where d is again the important dimension, ϱ is the electrical resistivity, and χ is the magnetic susceptibility.

The effect is largest in a partially magnetized material, since then all stress variations can result fully in changes of the magnetization. There is no damping when the material is magnetically saturated. In field zero, a domain structure still exists in ferromagnetic materials, the variations of which under the influence of the stress can produce damping. This *micro*-effect is found to be proportional to the frequency over a very wide range, covering the low and medium frequencies. It occurs most conspicuously in the kc/sec range. The earlier mentioned *macro*-effect is zero without applied magnetic field, and otherwise shows a wide damping maximum in the kc range. The temperature has an influence only through the variation of the electrical conductivity. The micro-effect is very sensitive to the presence of internal stresses.

2.5.3. *Ferromagnetic (magneto-mechanical) hysteresis*

In the low-frequency region ferromagnetic materials, still another type of damping exists, which has hysteresis character. It is due to stress-induced, sudden, irreversible motions of the Bloch walls separating magnetic domains, again through the intermediate of magnetostriction. As a hysteresis phenomenon, this damping is essentially independent of frequency (but can only be observed at low frequencies owing to the disturbing influence of magnetic relaxation at higher frequencies). In a magnetic field the Bloch walls are fixed and the damping disappears.

The magnetic hysteresis depends on the strain amplitude in a complicated manner, being proportional to the amplitude at low amplitudes, reaching a maximum afterwards and finally decreasing inversely proportionally to the square of the amplitude, at large amplitudes. For further particulars we refer to the papers by BOULANGER [1949]. In a very extensive study of a great number of alloys, COCHARDT [1956] has discovered some specific *high-damping* alloys (occurring in certain regions of the Fe-Ni-Co ternary diagram), the properties of which are largely determined by ferromagnetic hysteresis. These alloys are of particular value for some technical applications in which suppression of vibrations is of advantage (turbine blades).

2.5.4. *Electron attenuation of ultrasonic waves and magneto-acoustic resonance*

The statistical distribution of electrons in a metal (or semiconductor) over the Fermi surface can follow relatively slow deformation im-

mediately, since the electrons can move quickly enough in momentum space to adjust themselves at any moment to the strain. The relaxation time τ is of the same order as that occurring in the theory of electrical conductivity. However, at very high (ultrasonic) frequencies the deformations of the Fermi surface lag behind those of the lattice. A strain ε in a given direction, suddenly applied, produces a momentary deformation of the Fermi surface in that direction of the same order of magnitude. The number of electrons, n, per unit volume moving in this direction then increases by the factor $(1 + \varepsilon)$, and the "kinetic pressure" of the electron-gas changes by ${}^{8}/_{15} n E_F \varepsilon$ (MORSE [1955]), where E_F is the Fermi energy. These electrons are subsequently

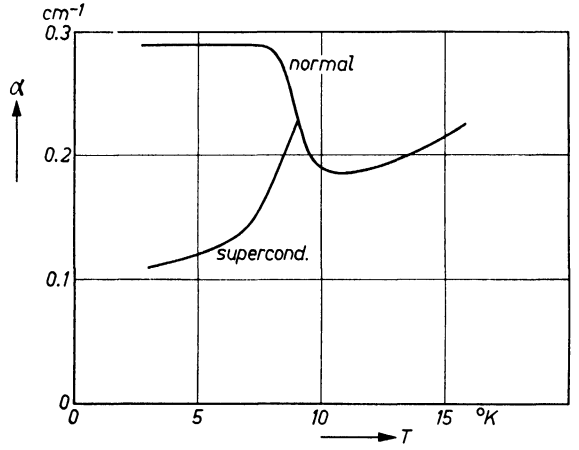


Fig. 19. Electronic attenuation of 26 Mc/s ultrasonic waves in lead as a function of temperature (BÖMMEL [1954]). Above 10 °K dislocation (Bordoni) damping becomes notable.

redistributed over the Fermi surface, which results in a stress relaxation by this amount. Ultrasonic waves are therefore attenuated. The attenuation coefficient is (for relatively low frequencies compared to $1/\tau_s$):

$$\alpha = \frac{4}{15} \frac{\omega^2 m E_F \sigma}{\varrho e^2 c^3}, \tag{43}$$

where the relaxation time has already been expressed in terms of the conductivity σ, and ϱ and c are density and velocity of sound respectively. In the 10–100 Mc/s range this becomes of the order of a few tenths per cm at temperatures below 10 °K in pure metals with low

residual resistance. A similar expression has been derived by KITTEL [1955].

The damping has indeed been observed by BÖMMEL [1954], who showed that at very low temperatures the ultrasonic attenuation in pure lead at 26–28 Mc/sec increases, and decreases again suddenly below the superconductive transition (Fig. 19). The same effect occurs in tin (MASON [1955b]). In both cases the attenuation indeed varies proportionately to the square of the frequency.

The effect of superconductivity points clearly to the conduction electrons as being the cause. Mason discussed the damping in terms of general viscoelastic electron motion rather than of linear after-effect.

PIPPARD [1956] pointed out that if the mean free path of the electrons becomes comparable to the wavelength of the ultrasonic waves, peculiar phenomena would be expected to occur. This would happen at very high frequencies (10^{11} c/sec) and for longitudinal waves only (because of the density variations associated with them). In the presence of a magnetic field the electron paths are "curled up" and interference of this kind occurs at much lower frequencies. BÖMMEL [1955] observed a damping of 20 Mc ultrasonic waves in lead that varied periodically with the magnetic field strength (of a few hundred gauss). PIPPARD [1957] termed this *magnetoacoustic resonance*. When the diameter of the electron orbit becomes approximately an odd number of half wavelengths of the ultrasonic waves, at low enough temperatures the electron is stabilized for several hundreds revolutions within the same wave valley. It moves then in a helicoidal path together with the wave and takes up much of the ultrasonic energy (a kind of cyclotron resonance). The study of this effect gives insight into the shape of the Fermi surface of a metal and is a subject of many recent researches.

3. Survey

In Table 1 the most important characteristics of the various types of internal friction discussed in this chapter have been brought together. For those interested in the specific properties of lattice imperfections, the categories a) *dislocation relaxation*; b) *dislocation resonance* and *dislocation hysteresis*; c) *transient cold-work internal friction* and *deformation hysteresis*; and d) *Köster damping* are the most promising internal friction phenomena. They inform us, more or less indirectly, about respectively: the Peierls force and thus the inter-

atomic forces; the binding energy of foreign atoms to dislocations; the diffusivity of point defects and the general aspects of the interaction between impurity atoms and imperfections. However, in the present state of the theory, only few of the conclusions drawn from the multitude of internal friction data can be regarded as completely conclusive, which is very unsatisfactory indeed.

TABLE I

Internal friction in metals

	Section	*Region of temp. and frequency*	*Type*	*Frequency dependence*	*Amplit. dependence*	*Temperature dependence; energy of activation U*	*Cold-work dependence*	*Impurity dependence*	*Mechanism*	*Inference*
Dislocation damping	**2.2**									
Dislocation relaxation	2.2.1	$T < 100$ °K; kc–mc	relaxation	peak(s)	none	U (of τ) = 0.05–0.1 eV (Cu)	through maximum after few % strain	small	relaxation of disloc. over Peierls barrier	Peierls barrier height
High-temperature dislocation damping	2.2.2	$T > 300$ °C; c–kc	relaxation	$\sim 1/\omega$	undefined, small	$U = 0.8$–2 eV (Cu)	large, also act. energy varies	?	relaxation of disloc. in internal stress field	
Dislocation resonance	2.2.3	$T \geqslant 0$ °C; 1–1000 mc	damped resonance	$\sim \omega$	none	$U = 0.2$–1 eV (Ge)	very large	impurities suppress	resonance of disloc. in equilibrium positions	binding energy imp. to disloc.
Dislocation hysteresis	2.2.4	$T \geqslant 0$ °C; kc–mc	hysteresis	none	parabolic	$U = 0.1$–0.4 eV (Cu)	very large	impurities suppress	break-away of disloc. from impur.	binding energy imp. to disloc.
Transient cold-work internal friction	2.2.5	$T \leqslant 200$ °C; c–mc	viscous	none	none?	$U = 0.5$–1 eV (annealing) Fe	large	?	rearrangement of disloc.	diffusion of point defect
Deformation hysteresis	2.2.6	$T \leqslant 100$ °C; c	hysteresis or viscous	?	parabolic	$U \approx 0.1$ eV; ann. ≈ 0.5 eV; third process present (Fe)	large	impurities suppress	hysteretic dislocation motion?	interaction point defects and dislocations?

Table 1 (*continued*)

	Section	*Region of temp. and frequency*	*Type*	*Frequency dependence*	*Amplit. dependence*	*Temperature dependence; energy of activation U*	*Cold-work dependence*	*Impurity dependence*	*Mechanism*	*Inference*
Non-exclusive dislocation damping	**2.3**									
Visco-elastic after-effect and microcreep	2.3.1	$T > 250$ °C; c or static	viscous	decreases with ω	?	$U \approx 1.5$–2 eV (Sn)	very little	impurities affect U	dislocations drag solute atoms or vacancies	interaction imp.-disl. vacancy diffusion
Temperature activated hysteresis	2.3.2	$T > 250$ °C; all frequ.	hysteresis or viscous	none	none?	$U = 0.15$–0.35 eV (Cu)	none	?	?	?
Grainboundary relaxation	2.3.3	$T > 250$ °C; c	relaxation	peak (broad)	none	U (of τ) 1–2 eV (Cu)	none	impurities affect U	grainbound. slip, diffusion	grainb. diffusion
Point defect internal friction	2.3.4	$T <$ room T c–kc	relaxation	several peaks	none	U (of τ) 0.3–0.5 eV (Cu)	only after cold-work	?	stress-induced migration of composite point defects	diffusion of point defects
Damping due to solute atoms and interstitials	**2.4**									
Snoek damping in int. b.c.c. alloys	2.4.1	room T; c	relaxation	peak	none	U (of τ) 1 eV (Fe-C)	none	proportional to conc.	stress-induced diffusion	diffusivity and solubility; precipitation
Snoek damping in int. f.c.c. alloys	2.4.2	< 200 °C; c	relaxation	peak	none	U 1.5 eV (NiC); 0.8 eV (Ni)	none; enhanced after c.w. and irr.	parabolic; small	stress-induced diffusion	ass. of def. and imp.; rel. of interstitialcies
Stress-induced ordering in subst. solid solutions	2.4.3	≈ 200 °C; c	relaxation	peak (broad)	none	U (of τ) 1 eV (Ag-Zn)	slight	none	stress-induced diffusion	diffusion, Kirkendall effect
Köster damping	2.4.4	≈ 200 °C; c	relaxation	peak	none	U (of τ) 1.5 eV (Fe)	large, $\sim \varepsilon^{1/2}$	saturates	stress-induced diffusion	interact. imp.-disloc.-point defects?

Table 1 (*continued*)

	Section	*Region of temp. and frequency*	*Type*	*Frequency dependence*	*Amplit. dependence*	*Temperature dependence; energy of activation U*	*Cold-work dependence*	*Impurity dependence*	*Mechanism*	*Inference*
Transformation damping	2.4.5	T_{trans} c-kc	mostly relaxation	peak (broad)	slight	narrow T region	large	variable	stress-induced ordering, or stress relaxation	kinetics of transformation
Damping due to thermal, magnetic and electronic causes	**2.5**									
Thermoelastic damping	2.5.1	all T; c–kc	relaxation	peak (broad)	none	slow	slight	slight	thermo-elastic heat flow	
Magnetoelastic relaxation	2.5.2	below Curie temp; kc	relaxation	$\sim \omega$ (micro); peak (macro)	none	slow	large (micro); slight (macro)	medium	eddy currents, magnetostriction	domain structure; magnetization processes
Ferromagnetic hysteresis	2.5.3	below Curie temp; c	hysteresis	none	maxim.	little	slight	medium	irrev. displ. of Bloch walls	domain structure; magnetization process
Electron attenuation	2.5.4	< 10 °K; > 10 mc	relaxation	$\sim \omega^2$	none	increases with decr. T	?	?	electron redistribution over Fermi surface	shape of the Fermi surface
Magnetoacoustic resonance	2.5.4	< 10 °K; > 10 mc; magn. field	resonance	periodic	none?	increases with decr. T	?	?	resonance between electron orbits and ultrasonic waves	shape of the Fermi surface

REFERENCES

BERGHOUT, C. W. (1958); Acta Metall. **6,** 1.
BRESHERS, D. N. (1959); J. appl. Phys. **30,** 252.
BIRNBAUM, H. and M. LEVY (1956); Acta Metall. **4,** 84.
BLEWITT, T. (1957); J. appl. Physics **28.**
BÖMMEL, H. E. (1954); Phys. Rev. **96,** 220.
BÖMMEL, H. E. (1955); Phys. Rev. **100,** 758.
BORDONI, P. G. (1949); Ric. Sci. **19,** 851.
BORDONI, P. G. (1954); J. acoust. Soc. America **26,** 495.
BORDONI, P. G. and M. Nuovo [1954]; Nuovo Cim. **11,** 127.
BORDONI, P. G. (1960); Suppl. al Nuovo Cim. **17,** nr. 1, p. 43.
BOULANGER, CH. (1949); Rev. de Métall. **46,** 321.
CHILDS, B. G. AND A. D. LE CLAIRE (1954); Acta Metall. **2,** 718.
COCHARDT, A. W. (1956); Trans. AIME **206,** 1295.
COTTRELL, A. H. and M. A. JASWON (1949); Proc. Roy. Soc. **A199,** 104.
DARLING, H. S. (1957); J. Inst. Metals **85,** 1779.
DONTH, H. (1957); Z. f. Physik **149,** 111.
DIJKSTRA, L. J. (1949); J. Metals **1,** 252.
EINSPRUCH, N. G. and R. TRUELL (1958); Phys. Rev. **109,** 653.
FAST, J. D. and L. J. DIJKSTRA (1950); Philips techn. Rev. **13,** 180.
FAST, J. D. and J. L. MEIJERING (1953); Philips Res. Rep. **8,** 1.
FAST, J. D. and M. B. VERRIJP (1954); J. Iron Steel Inst. **163,** 24.
FAST, J. D. and M. B. VERRIJP (1961); Philips Res. Rep. **16,** 100.
FRIEDEL, J. (1957); Les Dislocations (Gauthier-Villars, Paris).
FUSSFELD, H. I. (1951); Phys. Rev. **82,** 769.
GOERING, W. A. and A. S. NOWICK (1958); Trans. AIME **212,** 105.
GORDON, R. B. and A. S. NOWICK (1956); Acta Metall. **4,** 514.
GRANATO, A., A. HIKATA and K. LÜCKE (1958); Acta Metall. **6,** 470.
GRANATO, A. and K. LÜCKE (1956); J. appl. Phys. **27,** 583 and 789.
GROSS, W. (1953); Mathematical structures of the theories of viscoelasticity (Herrman, Paris).
HARPER, S. (1951); Phys. Rev. **83,** 709.
HASIGUTI, R. R. (1953); Proc. Int. Conf. Theor. Physics, Kyoto and Tokyo, p. 577.
HASIGUTI, R. R. and T. HIRAI (1951); J. appl. Phys. **22,** 1084.
KÊ, T. S. (1947); Phys. Rev. **71,** 533; **72,** 41.
KÊ, T. S. (1948); Phys. Rev. **74,** 9 and 914.
KÊ, T. S. (1950a); J. appl. Phys. **21,** 414.
KÊ, T. S. (1950b); J. Metals **2,** 575.
KÊ, T. S. and Y. L. MA (1956); Scientia Sinica **5,** 19.
KÊ, T. S. and C. T. Tsien (1956); Scientia Sinica **5,** 625.
KÊ, T. S., C. T. TSIEN and K. MISEK (1955); Scientia Sinica **4,** 501.
KÊ, T. S. and P. W. YANG (1957); Scientia Sinica **6,** 623.
KITTEL, C. (1955); Acta Metall. **3,** 295.
KOEHLER, J. S. (1952); Imperfections in nearly perfect crystals, p. 197 (Wiley, New York).
KÖSTER, W. (1940a); Z. Metallk. **32,** 282.
KÖSTER, W. (1940b); Arch. Eisenh. wesen. **14,** 271.
KÖSTER, W., L. BANGERT and R. HAHN (1954); Arch. Eisenh. wesen. **25,** 569.
KÖSTER, W. and K. ROSENTHAL (1938); Z. Metalk. **30,** 345.
LE CLAIRE, A. D. and W. M. LOMER (1954); Acta Metall. **2,** 731.
LÜCKE, K. (1956); J. appl. Phys. **27,** 1433.

MASON, W. P. (1955a); Bell. Syst. techn. J. **34**, 903.
MASON, W. P. (1955b); Phys. Rev. **97**, 557.
MEIJERING, J. L. (1957); La Diffusion dans les Métaux, p. 97. (Philips techn. Library, Eindhoven).
MORSE, R. W. (1955); Phys. Rev. **97**, 1716.
MOTT, N. F. (1948); Proc. phys. Soc. **60**, 391.
MOTT, N. F. and F. R. N. NABARRO (1948); Conf. on Strength of Solids at Bristol (London, The Physical Society).
NIBLETT, D. H. and J. WILKS (1956); Phil. Mag. **1**, 415.
NIBLETT, D. H. and J. WILKS (1957); Phil. Mag. **2**, 1427.
NOWICK, A. S. (1951); J. appl. Phys. **22**, 925.
NOWICK, A. S. (1953); Progress in Metal Physics **4**, 1.
NOWICK, A. S. (1954); J. appl. Phys. **25**, 1129.
PEARSON, W. and R. T. ROTHERHAM (1956); J. Metals **8**, 894.
PIPPARD, A. B. (1956); Phil. Mag. **46**, 1104.
PIPPARD, A. B. (1957); Phil. Mag. **2**, 1147.
POWERS, R. W. and M. V. DOYLE (1956); Acta Metall. **4**, 233.
ROZIN, K. M. and B. H. FINKELSTEIN (1953); C. R. Acad. Sci. U.S.S.R., **91**, 811.
SACK, H. S. (1950); Symp. Plastic Deform., Pittsburgh.
SCHEIL, E. and G. REINACKER (1944); Z. Metallk. **36**, 63.
SEEGER, A. (1956); Phil. Mag. **1**, 651.
SEEGER, A., P. SCHILLER and H. KRONMÜLLER (1960); Phil. Mag. **5**, 853 and Z. Naturf. **15a**, 740.
SNOEK, J. (1941); Physica **8**, 711.
STEVELS, J. M. (1950); Philips Res. Rep. **5**, 23.
STEVELS, J. M. (1951); Philips Res. Rep. **6**, 53.
THOMPSON, D. O. and D. K. HOLMES (1956); J. appl. Physics **27**, 713.
THOMPSON, D. O. and D. K. HOLMES (1959); J. appl. Physics **30**, 525.
WEERTMAN, J. (1955); J. appl. Phys. **26**, 202.
WEERTMAN, J. (1957); J. appl. Phys. **28**, 193.
WEERTMAN, J. and E. I. SALKOVITZ (1955); Acta Metall. **3**, 1.
WEERTMAN, J. and E. I. SALKOVITZ (1956); J. appl. Phys. **27**, 1251.
WERT, C. A. (1950); Phys. Rev. **79**, 601.
WERT, C. A. (1954); Acta Metall. **2**, 361.
WU, T. and C. WANG (1958); Scientia Sinica **7**, 1029.
ZENER, C. (1941); Phys. Rev. **60**, 906.
ZENER, C. (1949); Elasticity and anelasticity of metals (Univ. of Chicago Press).
ZENER, C. (1958); Suppl. al Nuovo Cimento No. **2**, page 560.

CHAPTER XVIII

DIFFUSION IN CRYSTALS

1. General Theory of Diffusion

Diffusion is an extraordinarily diverse subject, and of great importance for many branches of solid-state physics. In this chapter we shall direct attention only to those aspects of the diffusion phenomenon which are connected with physical crystalline imperfections. We shall confine our discussion to metals; many conclusions remain valid without alteration for other crystalline solids, and the typical differences will be discussed in the appropriate chapters of Part III.

Diffusion is essentially an irreversible flow of matter. Therefore, the thermodynamic treatment of diffusion brings with it some very tricky and as yet not wholly resolved difficulties. The irreversibility is shown clearly in the general form of the one-dimensional diffusion equation (Fick's law):

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left(D \frac{\partial c}{\partial x}\right). \tag{1}$$

which is linear in the time derivative. This equation expresses the phenomenological statement that everywhere and always the rate of accumulation of the diffusing atoms is equal to the gradient of the flux of matter, S. The latter quantity is assumed to be proportional to the local concentration gradient, the proportionality constant being the *diffusion coefficient* (diffusivity) D, of dimensions cm^2/sec. This coefficient measures the tendency of the diffusing atoms to *yield* to a difference in concentration; D is thus directly related to the atomic mobility. Normally, D is regarded as independent of concentration, and then the diffusion equation simply reads

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}. \tag{2}$$

The thermodynamical treatment of diffusion (cf. LE CLAIRE [1949]) starts not by considering the concentrations, but rather the Gibbs free energy G of a phase containing a number of components, the molar concentrations of which are c_i. Denoting the partial derivatives of G to these concentrations by μ_i, the *chemical potentials*, one has when temperature, pressure etc. remain constant, for the change in G owing to a change in composition

$$dG = \sum_i \mu_i dc_i . \tag{3}$$

Thermodynamic equilibrium is attained when each μ_i is constant throughout the system, irregardless of phase. When this is not the case, diffusion occurs to equalize the potentials. Now, a difference in μ_i at two sites not necessarily signifies a similar difference in c_i; it can be associated with no concentration difference at all or even a difference in the opposite sense. To study the tendency for diffusion, it is therefore more fundamental to regard the potentials than the concentrations. "Uphill" diffusion and related phenomena then appear as natural consequences of special configurations. The decrease of the free energy (per mole) of the system due to the displacement of 1 mole of the component i over the distance dx is simply

$$dG = d\mu_i = \frac{\partial \mu_i}{\partial x} dx + \text{second order terms.} \tag{4}$$

The driving force *per atom* on the substance i is thus:

$$F_i = \frac{1}{N}\left(-\frac{dG}{dx}\right) = -\frac{1}{N}\frac{\partial \mu_i}{\partial x}, \tag{5}$$

where N is Avogadro's number. If the average velocity of the atom under unit force is M_i (the *mobility* of the species i), then the average diffusion drift velocity is

$$v_i = -\frac{M_i}{N}\frac{\partial \mu_i}{\partial x} . \tag{6}$$

The mass flow in the direction of x, that is, the number of atoms crossing unit area perpendicular to x in unit time, is then

$$j_i = n_i v_i = -\frac{n_i M_i}{N}\frac{\partial \mu_i}{\partial x}, \tag{7}$$

where n_i is the number of atoms i per unit volume $= c_i A$, A being the total number of atoms per unit volume.

Further, in non-ideal solutions, the chemical potential μ_i can be expressed in the *activity coefficient* γ_i by using the relation:

$$\mu_i = g_{0i} + RT \log (c_i \gamma_i) . \tag{8}$$

where R is the gasconstant $= Nk$, k being Boltzmann's constant. In ideal dilute solutions $\gamma_i = 1$. Hence, as g_{0i}, the molar free energy of the pure substance i, is of course independent of x:

$$j_i = - n_i M_i \frac{RT}{N} \left\{ \frac{\partial \log (c_i \gamma_i)}{\partial x} \right\} = - A M_i kT \left\{ 1 + \frac{\partial \log \gamma_i}{\partial \log c_i} \right\} \frac{\partial c_i}{\partial x} . \tag{9}$$

The rate of change of concentration in a given element of volume at the site x is of course again proportional to the negative gradient of the mass flux, thus:

$$\frac{\partial n_i}{\partial t} = A \frac{\partial c_i}{\partial t} = - \frac{\partial j_i}{\partial x} = A M_i kT \left\{ 1 + \frac{\partial \log \gamma_i}{\partial \log c_i} \right\} \frac{\partial^2 c_i}{\partial x^2} . \tag{10}$$

One obtains again the thermodynamic diffusion equation:

$$\frac{\partial c_i}{\partial t} = \frac{\partial}{\partial x} \left(D_i \frac{\partial c_i}{\partial x} \right) . \tag{11}$$

where now

$$D_i = M_i kT \left\{ 1 + \frac{\partial \log \gamma_i}{\partial \log c_i} \right\} . \tag{12}$$

In the case of *self-diffusion* (e.g. transport of radioactive tracer atoms of the same kind as the atoms of the parent substance in a concentration gradient) γ is independent of c. In this case, as in the case of an *ideal solution* ($\gamma = 1$), the diffusion constant is simply given by $D = MkT$ or $D_i = M_i kT$ respectively. Then the well-known Einstein equation relating diffusivity and mobility is obtained.

Equations (11) and (12) can be directly compared with the phenomenological equation (1).

It must be noted here that the derivation of (12), in which we have followed LeClaire, is only allowed when it is assumed that the elementary atomic process that constitutes the diffusion mechanism involves the net motion of only *one* atom. Such is the case in all diffusion mechanisms we shall discuss in the following.

It is seen from equation (12) that each atomic species in the system has its own diffusion coefficient. Thus in principle different components diffuse at unequal rates. Of course, this possibility is realized only when indeed the migration of each atom takes place independently of the other atoms (no mutual exchange mechanism). It has actually been observed in several cases: the phenomena related to unequal diffusivities are denoted by the name *Kirkendall effect.*

Differences in the two diffusion coefficients in binary alloys only arise from differences in the atomic mobilities M_i. The term in (12) between brackets is necessarily the same for the two components in the alloy, in accordance with the Gibbs-Duhem equation.

2. Atomic Theory of Diffusion

Diffusion in metals (and in most other types of crystalline solids) can take place by several distinct atomic processes. We are here interested only in mechanisms where lattice defects play a role, thus

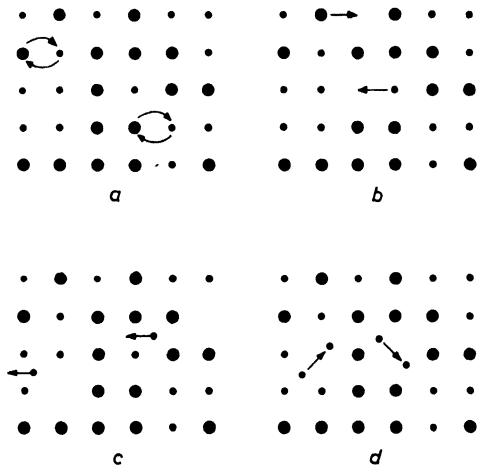


Fig. 1. Four mechanisms of diffusion in a binary compound:
a) Direct exchange; *b*) vacancy diffusion;
c) interstitial diffusion; *d*) interstitialcy diffusion.

not for instance in direct atomic interchange (Fig. 1*a*). The most important defect mechanism of diffusion in crystals is the vacancy mechanism.

2.1. VACANCY DIFFUSION

An atom can jump from its lattice position into a neighbouring vacant site thereby leaving a vacant site behind. Another atom (or the same) can then jump into this vacant site in its turn, whereby it shifts the vacancy to a third (or back to the first) position. If we consider the simplest case imaginable, that of *self-diffusion*, when only one atomic species is present, what actually happens is thus that the *vacancy* migrates through the lattice (Fig. 1*b*). In the presence of a concentration or rather a potential gradient, the average *drift* of many vacancies will be so directed that after a sufficiently long time the required net transport of mass has taken place.

The jumping probability of an atom, and in particular the difference in jumping probabilities of different atoms (because it is this that results in a net mass flow), depends on the vacancy concentration, on the diffusive force (gradient of the chemical potential) and on the height of the free-energy barrier between two neighbouring atom positions, A and B, respectively before and after the jump (Fig. 2*a*). The height of the barrier will depend on the type of crystal, the nature of the migrating atom, the diffusion mechanism, the interatomic forces, etc. The energy to overcome the barrier must be provided by thermal vibrations. This is one of the reasons why atomic diffusion requires an energy of activation, the other is that the equilibrium concentration of defects also depends on the temperature.

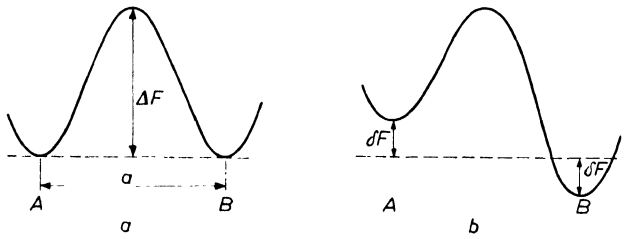


Fig. 2. To illustrate the derivation of the equation for the atomic diffusion coefficient.

Let ΔF be the height of the barrier expressed as a free energy. As is well-known the application of the simple reaction rate theory (cf. the book of GLASSTONE, LAIDLER and EYRING [1941]) then yields in reasonable approximation the probability for migration from A to B. The average number of transitions per second (considering)

an ensemble of processes A $\rightleftarrows$ B) is

$$n = \alpha \nu_0 \exp\left(-\Delta F/kT\right), \tag{13}$$

where α is a numerical geometrical factor depending on the precise diffusion geometry, and ν_0 is the average atomic vibration frequency (which is about 10^{13} sec^{-1} in most crystals). To derive formula (13) in an exact way is a very difficult procedure, which requires the formalism of the theory of stochastic processes, and consequently the validity of (13) is still a matter of much controversy. The interested reader is referred *inter alia* to a classical paper by KRAMERS for a thorough discussion of the difficulties [1940]. We shall use (13) in the following derivation.

The transition probability is altered when a potential gradient $\partial\mu/\partial x$ of the diffusing particles is present (for simplicity we suppose the diffusion to be one-dimensional, which makes no essential difference to the problem). The gradient of the free energy, that is, the driving force per atom is then

$$-\frac{1}{N}\frac{\partial \mu}{\partial x}$$

and Fig. 2*a* is changed into Fig. 2*b*. The potential energy of the positions A and B are no longer the same, they differ by the amount $2\delta F$, where

$$\delta F = \frac{1}{N}\frac{\partial \mu}{\partial x}\frac{a}{2}, \tag{14}$$

a being the distance between A and B, the jumping distance, and the same difference applies to the effective barrier height in the two directions. The average number of transitions per second from A to B will be larger than the number of those in the opposite direction, by the amount

$$\Delta n = \alpha \nu_0 \left(e^{-\frac{\Delta F - \delta F}{kT}} - e^{-\frac{\Delta F + \delta F}{kT}} \right).$$

When $\delta F \ll kT$ as is usually the case, one has

$$\Delta n = \frac{\alpha \nu_0}{NkT} a \frac{\partial \mu}{\partial x} e^{-\Delta F/kT}. \tag{15}$$

The drift velocity of the atoms considered is simply $v = a\Delta n$, and

comparison with formula (6) gives for the atomic mobility

$$M = \alpha \nu_0 \frac{a^2}{kT} e^{-\Delta F/kT} . \tag{16}$$

Since the factor

$$\left(1 + \frac{\partial \log \gamma}{\partial \log c}\right)$$

is unity for self-diffusion we then find from (12) for the self-diffusion coefficient the general formula

$$D = \alpha \nu_0 a^2 e^{-\Delta F/kT} . \tag{17}$$

We must still investigate the coefficient α more closely. First, it contains a geometrical factor which takes account of the fact that, if an atom can jump equally probably in m directions (for example $m = 6$ in the face-centred cubic lattice), then, on the average, of every jump of the atom in A the distance moved in the direction of B is different. Hence α contains a factor of the order of $1/m$ (vacancy diffusion in cubic lattices).

Further, in the vacancy mechanism of diffusion considered here, the atom at A can only jump to B when the atomic site B is *empty*, thus when a vacancy is at B. The probability for this is c_v, the vacancy concentration in the material. Now, in thermal equilibrium, c_v is given by (see Chapter II)

$$c_v = e^{-E_v/kT} , \tag{18}$$

where E_v is the free energy of formation of a vacancy.

Therefore, the factor α also contains the term $e^{-E_v/kT}$. For the diffusion coefficient (in the case of a vacancy mechanism) we thus find finally the expression:

$$D = \frac{1}{m} \nu_0 a^2 e^{-\frac{E_v + \Delta F}{kT}} . \tag{19}$$

The two free energies E_v and ΔF often contain quite appreciable entropy terms $-T\Delta S_v$ and $-T\Delta S$. These give rise to the occurrence of a temperature independent factor in (19). Usually, (19) is written as

$$D = D_0 e^{-Q/kT} , \tag{20}$$

where Q, the energy of activation for self-diffusion, stands for the sum of the *energies* of formation and migration of vacancies, and where

$$D_0 = \frac{1}{m} \nu_0 a^2 e^{(S_v + \Delta S)/k} \tag{21}$$

is denoted as the pre-exponential factor.

2.2. INTERSTITIAL DIFFUSION

Another mechanism of diffusion is *interstitial diffusion*. An atom occupying an interstitial position in the lattice, either because it belongs there (interstitial solid solution, for instance C in α-Fe), or after it has left a vacancy on the substitutional site it came from, migrates as such from one interstitial position to another through the lattice, until it may or may not end by occupying again a normal lattice site (Fig. 1*c*).

The jump from an interstitial site to a neighbouring one of course also requires an activation free energy, say $\Delta F'$, which in some circumstances may differ considerably from the substitutional migration energy ΔF, i.e. in the noble metals, where it is much smaller than the latter. However, the formation energy of an interstitial atom in a pure metal or substitutional solution is always quite large compared to that of a vacancy and the fraction of atoms that is in interstitial positions is therefore quite small. The resultant diffusitivity will then also become small, in spite of the large mobility. A calculation resembling that above can be set up, which leads to $(\frac{1}{2})E_i + \Delta F'$ as the activation energy for interstitial selfdiffusion in substitutional alloys. The factor $\frac{1}{2}$ enters when E_i stands for the total formation energy of a Frenkel defect. In interstitial compounds, E_i is per definition zero. Because of this, examples of systems showing interstitial diffusion are indeed provided by some (interstitial) alloys containing a component the atomic radius of which is very much (20% or more) smaller than of the other components (Pb-Au, Pb-Ag, Fe-C, and alloys containing hydrogen as a component, to mention but a few). Also in the substitutional alloys Ge-Cu and Ge-Ni interstitial diffusion is found, presumably because in the germanium lattice interstitial migration is exceptionally favoured (see Chapter XXXI).

Self-diffusion according to an interstitial mechanism has never been detected with certainty.

2.3. INTERSTITIALCY DIFFUSION

A variant of interstitial diffusion is diffusion by the *interstitialcy* mechanism, already mentioned in Chapter XIV. An atom on an inter-

stitial site pushes a neighbouring atom from its normal lattice position into the next interstitial site and itself occupies the substitutional site (Fig. 1*d*). The activation energy for this simultaneous motion of two atoms, in which the total number of interstitials is *not* changed, can, in principle, be small and this mechanism is much more probable than the former.

2.4. CHEMICAL DIFFUSION, SPECIAL MECHANISMS

The previous discussion was mostly concerned with self-diffusion, the migration of atoms through a crystal in the absence of a chemical concentration gradient. Without appreciable changes the treatment can be applied to chemical diffusion as well, only then care must be taken in the correct evaluation of the activation free energy for migration, and in the judicious use of factors like

$$\left(1 + \frac{\partial \log \gamma}{\partial \log c}\right).$$

It will be clear from formula (12), and has been pointed out by several authors (see e.g. SEITZ [1948, 1951]; LE CLAIRE [1953]), that a diffusion coefficient measured in a *chemical potential gradient*, such as in ordinary chemical diffusion experiments, need not be equal to the coefficient pertaining to the same composition, but measured for instance by a radioactive tracer method, in the absence of any concentration gradient whatsoever. It is only in the latter situation that no net vacancy current appears and the theory holds as described. If a net vacancy flow exists (e.g. in the Johnson mechanism mentioned below, and the diffusion processes discussed in sections 3 and 4), complications occur which lead to still other deviating values of D. FARA and BALLUFFI [1959] have made a start with the theoretical investigation of the effect of non-equilibrium vacancies on diffusion.

There exists an interesting method of determining diffusion coefficients by constructing a concentration gradient which is *in*dependent of the time. Such a *steady state* can for instance be obtained in gas permeation experiments, where a gas diffuses through a thin wall, the pressure on both sides of which is kept constant. The method has been applied to solid state diffusion too, however, it does not seem to lead to new results there that cannot be obtained by the conventional techniques.

BARDEEN and HERRING [1951] and COMPAAN and HAVEN [1956]

have drawn attention to a complication in the geometry of crystallographic diffusion. This arises because the statistical probability of an atom to jump in a particular direction depends on the place it came from in the previous jump. The atom has a larger probability to jump "back", which leads to the introduction of a *correlation factor* in the description of the macroscopic drift of atoms in a concentration gradient, at least if this takes place according to a *defect* mechanism. A determination of this factor yields information about the atomic mechanism of diffusion. This is practically impossible in metals, but very feasible in polar crystals. We refer therefore to Chapter XXV for a further discussion.

The activation energy for diffusion, which can usually be determined with an accuracy of about 0.1 eV, is commonly of the order of a few eV. The experimental accuracy of the determination of D_0 is much smaller, and the uncertainty easily amounts to a factor 10. The values computed by theory for Q in the various defect mechanisms are not very reliable either. It is therefore difficult to decide on the basis of a determination of the diffusion coefficient alone, which atomic mechanism plays a part. Nevertheless, by comparing subsidiary factors, it has become pretty certain that in nearly all metals substitutional diffusion takes place by a vacancy mechanism; direct atomic exchange does not occur and interstitial(cy) migration only takes place under special circumstances, which have already been mentioned earlier. It is usually characterized by a relatively small value of Q and a large value of D_0.

If a certain association probability exists between the diffusing atoms and vacancies, such as for instance may occur in an alloy with a component that is present in low concentration, the diffusivity by a vacancy mechanism may be appreciably enhanced (*Johnson effect*). The association may be there because the lattice distortions around the impurity atom are partially released by the presence of the vacancy. Not only can the equilibrium concentration of the vacancies in the material now be appreciably enhanced, also the migration energy of the complex may as well be lower than of the individual atom. The mechanism has been invoked with succes to explain the observed exceptionally large diffusivities in many substitutional alloys, in which it could be expected on good grounds that no interstitial diffusion can take place. (JOHNSON [1939], see also WAGNER [1938].)

The most direct evidence on the atomic mechanism of diffusion in metals comes from the Kirkendall effect, to be discussed in section 4.

Excellent and brief treatments of the mathematical and thermodynamical aspects of the diffusion problem in crystals have been given by LE CLAIRE [1949, 1953] and by SEITZ [1951]; more extended discussions can be found in monographs such as those of BARRER [1941], JOST [1952] and SEITH [1955]; the latter book primarily contains experimental data. A review of modern developments is that of BIRCHENALL [1958].

3. Influence of Pressure and Temperature

Application of an external hydrostatic pressure to a diffusion couple influences the rate of diffusion. One can formally define an *activation volume* of the elementary diffusion process according to:

$$v_a = -NkT\left(\frac{\partial F}{\partial p}\right)_T . \tag{22}$$

where F is the free energy of activation for diffusion, and thus equal to $E_v + F$ (according to (19) when a vacancy mechanism applies); p is the pressure and N is Avogadro's constant. The physical meaning of this activation volume is that it can be regarded as the sum of the (molar) volume of the N lattice defects taking part in the diffusion, and the further dilatation caused by their motion (NACHTRIEB [1958], see also BIRCHENALL [1958]). It should thus be positive when a vacancy mechanism, negative when an interstitial mechanism applies. If, in the first case, there were no elastic relaxation around a vacancy and no additional squeezing through the saddle point, v_a would be equal to the molar volume. From the pressure dependence of the diffusion coefficient it is found, almost universally in a number of solid materials (lead, sodium, α-white phosphorus), that the activation volume amounts to about half of the molar volume. TICHELAAR [1959] was able to verify this by elastic relaxation measurements. In liquid materials the ratio is only a few percents. From this it is concluded that a considerable amount of relaxation across a vacant lattice site takes place in most metals. A theoretical discussion of this interesting topic has been given by KEYES [1958]. BOSMAN *et al.* [1960] tried to measure the activation volume for the diffusion of *interstitial* C or N in α-Fe but found only very small values, in contrast to Keyes' theory.

NACHTRIEB [1958] was able to correlate the change of diffusion coefficient under pressure with that of the melting point; these changes were opposite, so that along the solidus line in the p-T-diagram the

diffusion coefficient would be a constant, as it should be according to the classical Eyring rate theory. In a solid whose melting point decreases with increasing pressure, the diffusivity should therefore increase (Meyering, unpublished).

An interesting possibility for deciding in principle experimentally whether interstitial or vacancy diffusion takes place has been pointed out by SHOCKLEY [1953], and was later discussed more extensively by LE CLAIRE [1954] and BRINKMAN [1954]. It has not led to success, however. The *simultaneous* presence of a temperature and a concentration gradient results in a flow of matter that is different for different kinds of defects. Consider a bar of material in which both a temperature gradient $\mathrm{d}T/\mathrm{d}x$ and a concentration gradient of *vacancies* $\mathrm{d}c_v/\mathrm{d}v$ are present (for simplicity we again confine ourselves to the one-dimensional case). Two lattice planes 1 and 2, distant by a lattice spacing a, have

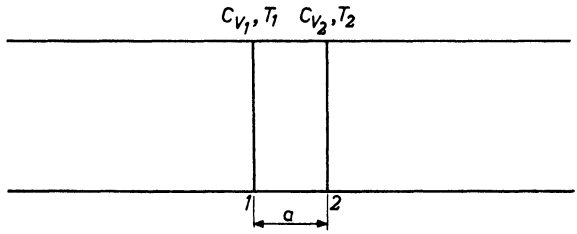


Fig. 3. On thermal diffusion.

the temperatures T_1 and T_2 respectively, whereas the vacancy concentrations are c_{v_1} and c_{v_2} (Fig. 3). The jumping frequency of a vacancy from 1 to 2 is $\nu_0 \exp(-\Delta F/kT_2)$, where ΔF is again the free energy for migration of a vacancy: T_2 is used because it is the atoms at plane 2 that must jump. Conversely, the jumping frequency of the vacancies from 2 to 1 is $\nu_0 \exp(-\Delta F/kT_1)$. We thus have for the flow of vacancies from 1 to 2 per unit area:

$$j_{12} = c_{v_1} \nu_0 \, \mathrm{e}^{-\Delta F/kT_2}, \tag{23a}$$

and for the flow from 2 to 1:

$$j_{21} = c_{v_2} \nu_0 \, \mathrm{e}^{-\Delta F/kT_1}. \tag{23b}$$

The *net* flow from plane 1 to plane 2 is thus:

$$j = j_{12} - j_{21} = c_v \nu_0 a \mathrm{e}^{-\Delta F/kT} \left(\frac{\Delta F}{kT^2} \frac{\mathrm{d}T}{\mathrm{d}x} - \frac{\mathrm{d}c_v}{\mathrm{d}x} \right),$$

where c_v and T denote the mean values of these quantities between the planes. The presence of a minus sign between the brackets implies that a temperature gradient in the *same* direction as the concentration gradient *counteracts* the flow of vacancies to some extent. In the presence of a temperature gradient a stationary state builds up, where the net flow of vacancies is zero and a vacancy gradient in the same direction is established, of magnitude

$$\left(\frac{\mathrm{d}c_v}{\mathrm{d}x}\right)_{st} = \frac{\Delta F}{kT^2}\frac{\mathrm{d}T}{\mathrm{d}x}. \tag{24}$$

If, however, enough sources and sinks (dislocations) are present in the material to ensure that everywhere and always the *equilibrium* concentration (18) of vacancies is maintained, a steady flow of vacancies takes place:

$$j_0 = \frac{\nu_0 a}{kT^2}(\Delta F - E_v)\, \mathrm{e}^{-(\Delta F + E_v)/kT}\frac{\mathrm{d}T}{\mathrm{d}x}. \tag{25}$$

It depends on the relative magnitude of the energies of formation and migration whether this flow is to higher or to lower temperature.

If, instead of vacancies, interstitial atoms are considered, *both* terms between brackets in (23) are preceded by a minus sign. Now the temperature gradient *promotes* downhill flow of interstitials, and in the steady state an interstitial gradient develops *against* the temperature gradient. If equilibrium conditions are maintained everywhere, there must exist a steady flow of interstitials always in the direction of *decreasing* temperature.

The effects here predicted have never been verified experimentally (SHEWMON [1958]). The reason for this is that, besides the kinetic considerations just given, another factor, arising from the irreversible nature of the phenomenon, should be taken into account. This concerns the amount of energy *transported* in each jump by the migrating atoms. Instead of (25), proper irreversible thermodynamics leads to a more general formula in which the free energy of migration ΔF is replaced by a free energy ΔQ, which is the transported energy mentioned above. At first sight both energies might indeed appear to be the same, but on closer consideration there are two contributions to ΔQ. First, not all of the energy ΔF taken up by the migrating atom, before it makes the jump, is transported to the next position; only a fraction

β is, and the rest is given up to the lattice atoms surrounding the saddle point. Secondly, when the migrating atom jumps into a vacancy, a certain amount of energy is released into the lattice. This is of the order of E_v, and hence:

$$\Delta Q = \beta \Delta F + E_v .$$

Inserting into (25) shows j_0 to be proportional to $\beta \Delta F$. If now β is very small, the observed invariance of j_0 can be understood.

4. Kirkendall Effect

In alloys it is not a priori certain that the various components diffuse at equal rates.

Only when the direct atomic exchange mechanism of Fig. 1*a* occurs, are the rates equal per se, but if a defect mechanism prevails, this need not be so. Consider a two-component alloy AB in which vacancy diffusion occurs. It may well be that one species of atoms, say A, changes place with a vacancy more easily than the other species, because of a difference in the respective barrier heights. In this case, the average drift velocity v_A of A in a potential gradient is larger than that of B, v_B, and a *net transport* of mass occurs during the diffusion: one side of the diffusion couple grows at the cost of the other. In a one-dimensional diffusion experiment, the velocity with which the lattice planes in the diffusion zone pass a rigid reference plane is then:

$$v = -(v_A - v_B), \tag{26}$$

where v_A and v_B (absolute values) are measured *with respect to the moving lattice planes*. The transport velocity of A *with respect to the fixed reference plane* is now

$$v'_A = v_A + N_A v = N_B v_A + N_A v_B \tag{27}$$

and, of course (absolute values)

$$v'_B = v'_A . \tag{28}$$

Here N_A and N_B stand for the fractional concentrations of A and B. Inserting the simple definitions

$$v_A = -D_A \frac{\partial N_A}{\partial x} \quad \text{and} \quad v_B = -D_B \frac{\partial N_B}{\partial x} \tag{29}$$

(concentration gradients are taken as the cause of the diffusion), one finds, since

$$\frac{\partial N_A}{\partial x} = -\frac{\partial N_B}{\partial x}:$$

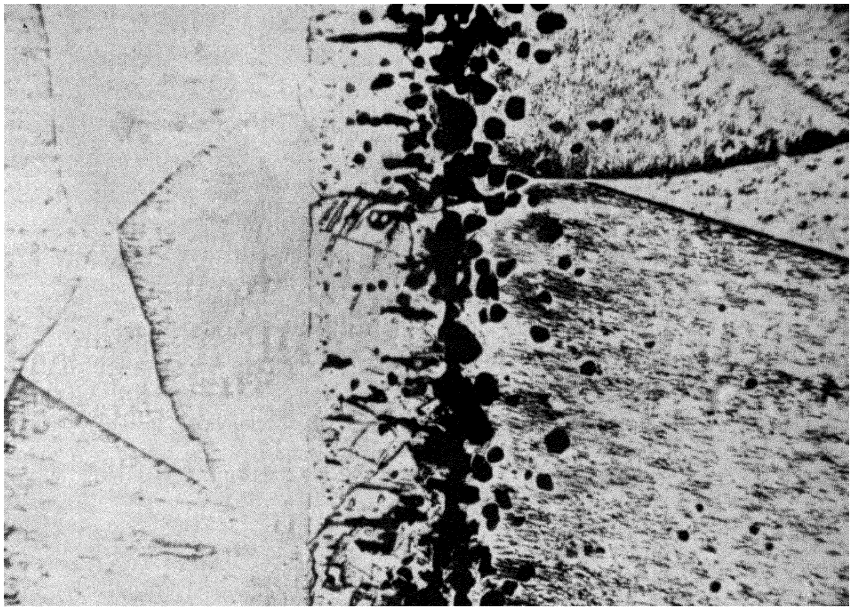
$$v'_A = -(N_A D_B + N_B D_A)\frac{\partial N_A}{\partial x}. \qquad (30)$$

$$v'_B = +(N_A D_B + N_B D_A)\frac{\partial N_B}{\partial x}.$$

The quantity $D = N_A D_B + N_B D_A$ is called the *interdiffusion* coefficient. The phenomenon of a net mass flow, or the motion of the lattice planes as a whole, has been called *Kirkendall* effect. It was discovered by Smigelskas and Kirkendall in the α-brass-copper system in 1947. The detection of a Kirkendall effect, say by the observation of the displacement of *markers* (e.g. thin wires of an inactive material or hardness indentations), points unambiguously to a defect mechanism for diffusion. It has been observed in nearly all accurately studied single-phase face-centred cubic and many body-centred cubic binary alloys, and in multiphase alloys as well: no direct exchange of atoms appears to take place to any considerable extent in these systems (see BOLK [1959]).

Generally associated with the Kirkendall-effect is the formation of *pores* and a lateral contraction of the sample on the side of the more rapidly diffusing component, a short distance from the original interface (Fig. 4). On the other side of that boundary, a slight bulging of the material is often observed. In some systems the porosity is exceptionally large, e.g. in 50/50 Au-Ag samples. The porosity zone moves as the diffusion proceeds, so as always to occur at a *particular composition*. It consists of larger or smaller holes, which are presumably formed by the condensation of excess vacancies. They are certainly not caused by the collection of dissolved gaseous impurities, as has been proved by taking materials of different degrees of purity and heating them in different atmospheres. A comparison of the observed net mass flow (shift of markers) with the computed flow on the basis of the measured difference between the partial diffusion coefficients has shown that the formation of pores might indeed be one of the

Pores



Welding interface

Fig. 4. Pores aside from the original interface of a Au-Ag diffusion couple, that have originated as a result of the Kirkendall effect, on the side of the more rapidly moving (Ag) component. (After W. SEITH [1955], p. 140.)

methods in which the necessary volume change during the diffusion is taken up.

To explain the formation of pores, BRINKMAN [1955] has proposed a quantitative theory which has as an essential element the establishment of a system of elastic stresses in the diffusion zone, due to the occurrence of volume changes. Although the degree of superconcentration of vacancies is but low and the probability that a few of them will meet to form a void is small, they will nevertheless tend to collect into macroscopic pores because the total elastic energy of the stress system is decreased by doing this. There exists a critical size above which voids will grow, which in the absence of a stress is rather large, but which, when the stresses are considered, becomes sufficiently small for nucleation to take place. The nucleation is further enhanced by the presence of impurities and dislocation arrays.

Other structural changes consequent to the occurrence of a Kirkendall effect have been investigated in detail by DOO and BALLUFFI [1958]. In the diffusion zone of copper-α-brass couples they found new dislocations to be generated, which were subsequently arranged into subboundaries, whereas simultaneously the phenomena of recrystallization and twinning took place locally, the latter two invariably together. All processes occurred most conspicuously at low diffusion temperatures and were explained in terms of the non-equilibrium concentration of vacancies that must be maintained to enable the diffusion rate of the two components to build up (see below). The excess vacancies are assumed to be generated at dislocations present in the diffusion zone, and cause the latter to climb. It is *this* process which is held responsible for the observed dislocation rearrangement, which bears thus some resemblance to subboundary formation by annealing at high temperatures (Chapter XI).

The mathematical description of the Kirkendall effect is difficult; we shall only indicate the outlines, basing ourselves on a vacancy mechanism. If vacancies exchange more easily with one component, say A, of a one-dimensional diffusion couple AB than with the other, B, then the A atoms will diffuse more rapidly than the B atoms. A Kirkendall effect will be observed, and inert markers will shift towards A. However, the flow of A atoms being greater than the opposite flow of B atoms, there will necessarily occur a counter flow of vacancies towards the component A. Consider a band of width Δx perpendicular to the diffusion direction (Fig. 5). The number of atoms

A leaving this band per unit time is proportional to

$$\left(\frac{\partial j_A}{\partial x}\right) \Delta x ,$$

the number of atoms B entering it is proportional to

$$\left(\frac{\partial j_B}{\partial x}\right) \Delta x ,$$

where j_A and j_B again stand for the partial fluxes of A and B. The number of vacancies left over in the band per unit time is thus:

$$n_v = \frac{\partial}{\partial x}(j_A - j_B)\,\Delta x = \frac{\partial}{\partial x}\left\{(D_B - D_A)\,\frac{\partial c}{\partial x}\right\}\Delta x . \qquad (31)$$

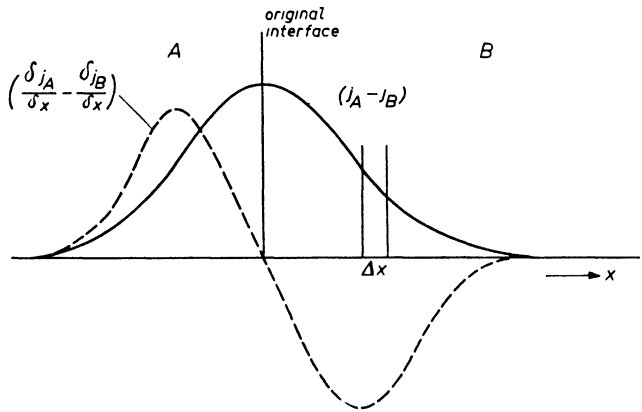


Fig. 5. Net mass flow j_A–j_B and its derivative (dotted curve) in a diffusion experiment. The generation rate of excess vacancies is proportional to $\frac{\partial}{\partial x}(j_A - j_B)$, it has a maximum where the dotted curve reaches its maximum (porosity zone).

(When this difference is negative, it means that vacancies are absorbed instead of formed.) Now, the porosity as has been observed is presumably greatest in that region where n_v has it maximum, thus where $\partial^2 j_A/\partial x^2 = \partial^2 j_B/\partial x^2$.

In a plot representing the net mass flow $j_A - j_B$ as a function of the coordinate (Fig. 5), the slope of this curve (drawn line) is thus a measure for the local production rate of vacancies, and this rate will be maximum where the dotted curve has a maximum on the side of A,

and minimum at the minimum of the curve on the side of B. There thus arises a *gradient* in the vacancy concentration, and this in its turn causes diffusion of vacancies in the sense *opposite* from that originating from the proper Kirkendall effect. With this diffusion of course an atomic flow is again connected and it is *this* flow that causes the movement of the inert markers, if present.

The concentration gradient of vacancies is maximum near the original plane of separation of A and B, and is zero at the inflexion points mentioned. Markers placed between these points will thus move against the direction of the fastest moving component, which is the commonly observed result. *Outside* the inflexion points, however, the vacancy concentration gradient is oppositely directed, and markers placed there will move (much more slowly, to be sure) *in* the direction of the fastest moving component. That is, the movement of the lattice planes changes its sense at the inflexion points, as was first mentioned by HEUMANN [1957]. BOLK [1959] has pointed out that this conclusion depends on the assumptions that $D_A > D_B$ over the complete range of concentration and that vacancies are in local equilibrium everywhere through the diffusion zone. In the course of the diffusion process, the net mass-flow curve gradually becomes flatter, the inflexion points move outwards, and the region of "normal marker movement" extends. The "retrograde mass flow" has been observed by SEITH and coworkers [1955] in the system Au-Ag. Gold diffuses more rapidly than silver and markers will thus move to the silver side in the inside diffusion zone and to the gold side in the outside regions. The phenomenon is of course somewhat complicated by the simultaneous formation of pores which brings about another change of dimensions of the diffusion couple.

The question of the origin of the vacancies is of great interest. The thermal equilibrium concentration of these defects is normally much too small to account for the observed mass flow. There must be sources of vacancies in the slowest moving component and sinks in the fastest moving component (the vacancy flow is opposite to the mass flow). Possible sources and sinks are the surface, grain boundaries if present and especially jogs in dislocations of not pure screw character. In view of the absence of any dependence of the Kirkendall effect on the size and grainsize of the specimen, dislocation sources seem the most probable, as we have already mentioned previously. The density of dislocations in the material must be such that the rate of production

of vacancies is everywhere sufficiently high to account for the observed net mass flow. The latter is in usual cases of the order of 10^{17} atoms/cm^2/sec. The width of the diffusion zone being of the order of 1 mm, this entails a vacancy production of 10^{18} vacancies per cm^3 per second. If the dislocation density is 10^8 cm^{-2}, there are 10^{15} atomic sites per cm^3 on dislocation lines. When 1% of these are jogs, then each jog must emit 10^5 vacancies per second. Putting the energy of formation of a vacancy equal to 1 eV, the thermal equilibrium rate at which jogs emit and absorb vacancies at a temperature of 500 °C is about 10^7 sec^{-1}. Thus indeed the number and productivity of jogs are sufficient to account for the required supply of vacancies during the diffusion process.

Of course the occurrence of a Kirkendall effect, porosity, eventually also elastic and plastic strains, etc. invalidates the applicability of the simple kinetic theory of chemical diffusion presented in the section 2. Also the thermodynamic treatment of section 1 breaks down, since no equilibrium exists with respect to the vacancy concentration. The rate of flow of each component no longer depends on its own chemical potential only, but also on the vacancy concentration, etc. A solution of this problem may be obtained by the methods of irreversible thermodynamics (LE CLAIRE [1953]).

5. Diffusion along Grain Boundaries

5.1. GRAIN BOUNDARY DIFFUSIVITY

Surface imperfections such as grain boundaries and line imperfections such as dislocations can be regarded as extended regions of enhanced lattice disorder. It is therefore probable that they will provide especially suitable diffusion paths. Indeed, many observations exist regarding the enhancement of diffusion rates when grain boundaries are present. Table 1 (after LACOMBE [1957]) gives the results of several measurements of self-diffusion along grain boundaries, expressed in terms of effective diffusivities, and characterized again by an activation energy and a pre-exponential factor (for definitions see the mathematical treatment below).

The activation energies for grain boundary diffusion are found to be roughly half of those for volume diffusion. Therefore, in experiments conducted at *low* temperatures grain boundary diffusion becomes

especially notable, and can be separated observationally from volume diffusion.

TABLE 1

Coefficients of self-diffusion in the volume and along grain boundaries

Metal	*Volume self-diffusion*		*Grain boundary self-diffusion*	
	D_0 ($cm^2 \cdot sec^{-1}$)	Q (eV)	D_0 ($cm^2 \cdot sec^{-1}$)	Q (eV)
Silver	0.895	2.0	0.03	0.88
Nickel		2.6–2.8		0.4–0.7
Lead	6.56	1.21	0.81	0.68
Zinc	$\parallel c$ 0.13; $\perp c$ 0.48	$\parallel c$ 0.95; $\perp c$ 1.06	0.22	0.62

A review of experimental and theoretical work on grain boundary diffusion has been given by LACOMBE [1957]; in the following we shall mention only some of the most important aspects.

An approximate phenomenological theory of diffusion through a grain boundary has been given by FISHER [1951]; mathematically more satisfactory treatments have been developed by Whipple and others, without physically very different results, however. Fisher considered a grain boundary in the form of a plane slab of constant

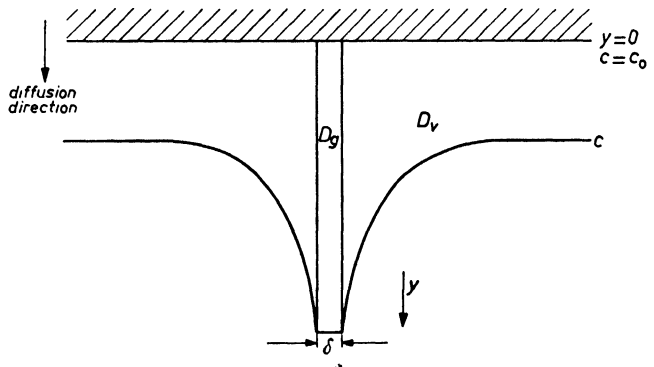


Fig. 6. Mathematical model of diffusion through a grain boundary (after FISHER [1951]); c equiconcentration curve, δ width of boundary, D_g grain boundary diffusivity, D_v volume diffusity. The x-direction is horizontal, the z-direction perpendicular to the plane of drawing.

width δ (Fig. 6), infinitely extended in the (yz)-plane (of the boundary) and filled with a material of diffusivity D_g, very much greater than the diffusivity D_v of the bulk material in which it is embedded. The

material fills the half plane $y > 0$. Both diffusivities are considered as independent of the concentration, which moreover is assumed to be constant laterally across the narrow boundary. Fick's law (1) again governs the diffusion of atoms from the plane $y = 0$ into the material; the boundary conditions are rather formidable. However, under the initial condition that the concentration at $y = 0$ remains constant, c_0, in time, and that *no* lateral diffusion from the grain boundary into the bulk occurs, volume diffusion and boundary diffusion can be dealt with separately. In the bulk one has then the well-known one-dimensional solution

$$c_v(y) = c_0 \operatorname{erfc}\left(\frac{y}{2\sqrt{D_v t}}\right), \tag{32}$$

and in the boundary the similar solution

$$c_g(y) = c_0 \operatorname{erfc}\left(\frac{y}{2\sqrt{D_g t}}\right), \tag{33}$$

where

$$\operatorname{erfc} x = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt\,. \tag{34}$$

The assumption that no lateral diffusion occurs from the boundary into the bulk cannot be maintained, however. Fisher considered the contribution to the concentration in the bulk due to this lateral diffusion as a separate two-dimensional diffusion problem; the boundary conditions are determined by (33). This is allowed if indeed grain boundary diffusion is much more rapid than volume diffusion. He obtained for the additional volume concentration arising from lateral diffusion in this approximation:

$$c'_v = c_0 \exp\left\{-\left(\frac{4}{\pi}\frac{D_v}{t_e}\right)^{1/4} \frac{y}{\sqrt{D_g \delta}}\right\} \cdot \operatorname{erfc}\left(\frac{x}{2\sqrt{D_v t}}\right). \tag{35}$$

This is usually much greater than (33) at distances x not very far from the boundary. Most of the material is thus first transported along the grain boundary and from there it is distributed laterally into the bulk. In a narrow section of width Δy parallel to the surface and at a distance y from it, the total quantity of diffusing material that has arrived after a time t from the surface, is now given by:

$$A_{v,g} = nR^2 \Delta y c_v + 2nR\Delta y \int_{-\infty}^{+\infty} c'_v dx \,, \qquad (36)$$

when each cm² of the surface is divided into n regions of average area R^2 by the grain boundaries nearly perpendicular to it. (It is assumed that the material contained in the grain boundaries itself does not count.) Carrying out the integration and neglecting the first term (direct volume diffusion) altogether, one finds:

$$A_{v,g} = 4.5nR\Delta y c_0 \sqrt{D_v t} \exp\left\{-\left(\frac{4D_v}{\pi t}\right)^{1/4} \frac{y}{(D_g\delta)^{1/2}}\right\} . \qquad (37)$$

If only volume transport had occurred, the quantity of matter that has arrived in the layer Δy were simply:

$$A_{v,v} = nR^2 \Delta y c_v = nR^2 \Delta y c_0 \operatorname{erfc}\left(\frac{y}{2\sqrt{D_v t}}\right) . \qquad (38)$$

Formulae (37) and (38) can now be directly compared to interpret the experimental findings. By this comparison it is then found that grain boundary diffusion will only be experimentally detectable when $D_g/D_v \geqslant 10^5$ and the grain size is relatively small ($R \leqslant 10^{-3}$ cm when $D_g/D_v = 10^5$; $R \leqslant 0.2$ cm when $D_g/D_v = 10^6$). In less favourable cases grain boundary diffusion does not add significantly to volume diffusion within measurable ranges of time and temperature. It is therefore not surprising, that only in carefully performed experiments can preferential grain boundary diffusion be demonstrated, and then only at relatively low temperatures where the diffusion is slow.

How the experimental results in Table 1 have been obtained the following example shows. HOFFMAN and TURNBULL [1951] have found the self-diffusivity of silver in single crystals to be representable by

$$D = 0.895 \exp\{-2.0\,(\text{eV})/kT\}\ \text{cm}^2/\text{sec} \,,$$

whereas at low temperatures the diffusivity in polycrystals could be represented by the apparent diffusion coefficient

$$\bar{D}_v = 2.3 \times 10^{-5} \exp\{-1.15\,(\text{eV})/kT\}\ \text{cm}^2/\text{sec} \,.$$

By applying Fisher's analysis, the investigators were able to interpret

the difference by assuming the presence of a separate grain boundary diffusivity of

$$D_g = 0.03 \exp \{ - 0.88 \text{ (eV)}/kT \} \text{ cm}^2/\text{sec} .$$

The other data in Table 1 were obtained by a similar analysis.

5.2. EFFECT OF ORIENTATION

Many authors, among whom ACHTER and SMOLUCHOWSKI [1951], have studied the effect of grain boundary *orientation* and of the orientation *difference* between the neighbouring grains on the grain boundary diffusivity. They have investigated also the possibility of *anisotropic* grain boundary diffusion. It was found that the grain boundary penetration of silver in "columnar" polycrystalline copper, of which all grains had a cube direction in common parallel to the boundaries, is maximum when the orientation difference between the grains amounts to 45°, and presents a shallow minimum at 53°. Below 20° or above 70°, preferred grain boundary penetration was no longer detectable. Quite similar behaviour (both maximum and minimum present) was found for the penetration of radioactive iron into 4% silicon-iron (HAYNES and SMOLUCHOWSKI [1955]) (Fig. 7).

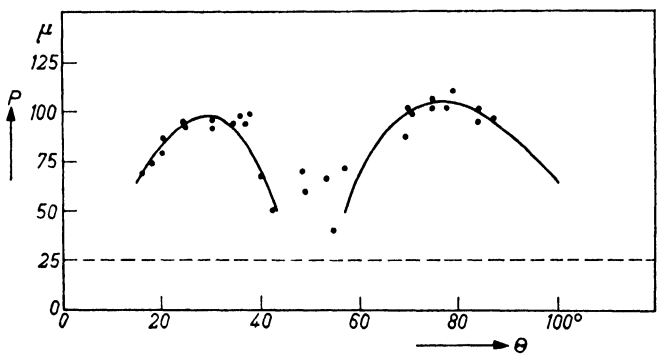


Fig. 7. Penetration depth P of radioactive iron into a grainboundary of Si–Fe as a function of disorientation angle θ (between two cube directions) of a bicrystal with a common (011)-plane (after HAYNES and SMOLUCHOWSKI [1955]).

The self-diffusion rate of silver in silver through tilt boundaries (see below) having a [100]-direction in common was found to be appreciably larger *along* that direction than *perpendicular* thereto (difference in penetration about 2–5 times, COULING and SMOLU-

CHOWSKI [1954]). The anisotropy depends on the orientation difference between the grains, however; it was most outspoken when this difference was smaller than 20° (or larger than 70°), and disappeared at an orientation difference of 45°. The anisotropy of the grain boundary self-diffusion of lead in various types of boundaries has been studied in much detail by Burgers, Tiedema and coworkers (see e.g. OKKERSE *et al.* [1955]). Their results, as well as those mentioned above, can all be interpreted in terms of the *dislocation model* of a grain boundary. This dislocation model will be fully treated in Chapter XIX. Here it may suffice to mention the simplest type of grain boundary, the tilt boundary, in which two grains are rotated with respect to each other about a common crystallographic direction (see Fig. 2 of next chapter). This boundary can be regarded as a wall of equidistant edge dislocations parallel to the rotation axis, of the type already encountered in the discussion of polygonization. Between the orientation difference of the tilted grains, θ, and the spacing d of the dislocations along the boundary the relation exists

$$2 \sin \theta/2 = b/d \,, \tag{39}$$

where b is the Burgersvector of the dislocations. This dislocation model loses its significance when d becomes smaller than b, as then a clear distinction can no longer be made between the individual dislocations. The boundary must from then on rather be regarded as a zone of alternating regions of good and bad fit (see Chapter XIX).

When the boundary plane has a symmetric crystallographic position between the grains, it can be built as a particularly simple dislocation array with only slight lattice strains in between the imperfections. It is to be expected that the grain boundary penetration will be small at these particular disorientation angles, as the local disordering is but small. Such a case is indeed present for the 53° boundary in silver, quoted above.

The anisotropy of grain boundary diffusion as found by Okkerse c.s. in lead and by Couling and Smoluchowski in silver can also easily be explained in terms of the dislocation model, as it springs to the eye that diffusion along the dislocation lines might well be more easy than the migration of atoms perpendicular to the dislocations. All observations so far are in excellent agreement with this supposition: no preferred grain boundary diffusion was detected at all in a direction perpendicular to the dislocations in a simple tilt boundary, but the

penetration was maximum along the direction of the dislocations.

An obvious test of the influence of grain boundary dislocations on diffusivity is the study of the preferred penetration into *small-angle* grain boundaries, where the dislocation model is most appropriate, and quantitative conclusions may be drawn. The early observations by Smoluchowski and Achter, already mentioned above, lead to the conclusion that below 20° orientation difference grain boundaries showed no preferred migration at all in copper and in silicon-iron. However, this negative result was probably due to insufficient accuracy. Later very careful experiments, especially those of YUKAWA and SINNOTT [1955] about the grainboundary penetration of radioactive nickel in oriented copper bicrystals, and also of HOFFMAN and TURNBULL [1951] about the grain boundary selfdiffusion of silver in oriented silver bicrystals, have shown that even at very low angles of disorientation the grain boundary penetration is higher than the volume diffusion. As to the exact dependence of the diffusivity on the orienration difference of the grains, no agreement exists between the different investigators. Turnbull and Hoffman assumed that the diffusion along the dislocation direction of a symmetric tilt boundary can be described in terms of a "dislocation diffusivity" D_d, in narrow cylinders around each dislocation line, of radius l and of an intermediate diffusivity D_g' in between these dislocation zones (but still within the width δ of the grain boundary, of course). They further took this quantity δ not constant like Fisher did, but equal to $l^2/d \approx l^2 \sin\theta/b$ where d is the spacing between the dislocations. Taking $D_d \gg D_g' \gg D_v$, at small disorientation angles θ the effective grain boundary penetration in this fairly artificial model becomes practically fully determined by the product $D_d \cdot \delta$, and thus is proportional to $\sin\theta$. The authors were able to interpret their observations numerically by taking:

$$D_d = 0.14 \exp(-0.86\,(\text{eV})/kT)\ \text{cm}^2/\text{sec}\,.$$

This energy of activation agrees very well with that of the diffusion in grain boundaries of *average* orientation, viz. 0.87 eV.

On the other hand, Yukawa and Sinnott established a definite dependence of the grain boundary diffusivity on the crystallographic orientation with respect to the lattices on both sides of the grain boundary itself; boundaries parallel to a ⟨100⟩-direction in copper showed by far the greatest penetration. These authors were *not* able

to interpret their results on small-angle grain boundary diffusion in terms of Fisher's model.

6. Enhanced Diffusion along Individual Dislocations, and Owing to Point-defects in Non-equilibrium

The low-angle boundary diffusion results seem to point to a preferential diffusion along dislocation channels in metals. Direct quantitative evidence on such a "dislocation-pipe" diffusion is lacking, however. The amount of material that can be transported (albeit at an enhanced rate) through dislocations is very small, especially so when the latter are present in such low densities that they can be counted individually. Therefore, the direct experimental verification of dislocation diffusion is very difficult. By letting the material that diffuses along dislocations accumulate somewhere for a long time, the transport may nevertheless become detectable. HENDRICKSON and MACHLIN [1954] tried to obtain this accumulation at the free surface. They bent and afterwards polygonized a thin silver strip which thus contained a known number of more or less parallel dislocations, running approximately perpendicular to the plane of the strip. After plating the strip on one side with radioactive silver, they heated it to such a temperature and for so long a time that, although volume diffusion from one side to the other was still negligible, the transport through dislocation "pipes" could be expected to have been appreciable. The radioactive silver material transported this way would have accumulated by spreading over the opposite surface, and it was therefore to be expected that the radioactivity of that face would have increased. This was indeed observed, and, what is more, by removing a thin surface layer of this opposite face, the activity decreased again considerably. This showed that material had indeed accumulated on the surface only, and that it had not reached that surface by volume diffusion. It might have attained the opposite face by surface migration around the strip, however. It has not yet been possible to express the results quantitatively, or to derive an effective cross-section of the dislocation pipes. The really crucial experiment to ascertain the very existence or otherwise of dislocation diffusion has not yet been carried out. Indirect indications of the effect exist, however, although they are not unambiguous either.

Many observations on the rate of precipitation and the kinetics of

phase changes in alloys can only be interpreted in terms of a diffusivity which is much higher (up to 10^7 times) than the ordinary volume diffusion in the material. Most spectacular in this respect is the formation of so-called Guinier-Preston zones (see section XXI–3.3) in quenched aluminium-copper, aluminium-silver and aluminium-zinc alloys with a few % Cu, Ag or Zn, when kept at the relatively low temperature of about 100 °C. Between 50° and 150 °C in Al-Cu alloys plate-like, coherent copper-rich precipitates form, a few atomic spacings thick and about 50 spacings wide. The rate of formation of these G.P.-zones is more than 10^7 times greater than can be calculated from the volume diffusivity of copper in aluminium, assuming the usual random precipitation process. On the contrary, the rate of ordinary precipitation to the $CuAl_2$ (θ) phase that occurs on heating to higher temperatures, is indeed determined by this diffusivity. The same discrepancy between precipitation rates is found in aluminium-silver alloys, and TURNBULL and HOFFMAN [1954] suggested that it might have its cause in enhanced diffusion of the copper atoms along dislocation channels ending at the zones. Even then, the dislocation density would have to be exceedingly high to explain the enormous rations observed. The dislocations were therefore moreover assumed to be mobile, "sweeping" their glide planes to catch the copper atoms present and transport these to the G.P.-zones. GUINIER [1958], GRAF [1958] and others have criticised this idea on the basis of its artificiality and inter alia, of the following evidence. They observed that the rate of formation of G.P.-zones in Ag-Al and other alloys depends on the quenching temperature, showing a maximum after quenching from 460 °C. After heating the quenched metal for a long time at 200 °C, the precipitation rate has nearly become unobservable. The explanation proposed by these authors was not in terms of dislocations at all, but rather of vacancies, quenched-in in supersaturated amounts, which also promote the diffusivity. Quenching from a very high temperature results in fewer vacancies because of coagulation, and annealing allows the excess vacancies to diffuse out of the metal. The quenched-in vacancies associate partly with the atoms of the alloying element and render these extremely mobile (Johnson effect, see above). They can also interact with screw dislocations, transforming these into helicoidal dislocations which provide nuclei for the precipitation of zones (THOMAS and WHELAN [1959]).

Graf was able to show that simultaneous plastic straining after

quenching increases the rate of formation of the normal θ-phase. This is in accord with the idea that not only dislocations can enhance the diffusivity but vacancies (or in some cases interstitials) also, as will be clear from the general theory of section 2. In terms of the discussion of that section, the presence of vacancies affects the coefficient α. A difference with dislocations is that, whereas the latter predominantly affect diffusion over long distances, excess vacancies have preponderantly an effect on short-range diffusion processes: often the vacancies, having aided an elementary diffusion jump, become themselves lost. Vacancy effects are of course most conspicuous in alloys containing an excess number of defects. Therefore, to detect their influence in metals, diffusion and phase changes must be studied in alloys that have been cold-worked, quenched or irradiated (cf. Chapters XIV and XV). It is often difficult to separate in such experiments the effects of increased diffusivity and of enhanced nucleation, respectively. To the latter topic we will return in Chapter XXI.

An interesting type of experiment is the measurement of the rate of self-diffusion during plastic straining. This has been carried out on silver by a number of investigators (DARBY *et al.* [1958], LEE and MADDIN [1959]), but the results are not in mutual agreement. Darby *et al.* found the diffusivity in compressed or extended silver at about 800 °C not to be enhanced by more than a factor 2, but Lee and Maddin, who used torsion and applied a more detailed analysis, measured an increase of the diffusion rate of no less than a factor 100, proportional to the *rate* of straining, thus to the instantaneous vacancy concencentration.

Very appropriate for studying the influence of excess vacancies on diffusion are investigations on the rate of *ordering* of alloys possessing an ordered state. Thus DUGDALE [1954] has studied the effect of γ-irradiation on the ordening of partially disordered Cu_3Au. He found this to be increased appreciably above 100 °C, where the vacancies become mobile: each vacancy produced by the radiation could aid in ordering at least 10^5 atoms of the alloy. The ordering is usually measured by means of the resistivity. Changes in order produced by cold-work in Cu_3Au have been found to anneal out quicker in proton-irradiated samples than in non-irradiated specimens, again pointing to an acceleration of the short-range diffusion by the vacancies formed. Much other experimental evidence exists on this topic.

Quenching, of course, is also a method well adapted to study the

effect of vacancies on diffusion. Only, the excess vacancies often run out quickly at the temperature of the diffusion experiments. There is a category of experiments by which one can measure diffusivities at very low temperatures, however, namely those on internal friction (Chapter XVII). NOWICK and coworkers [1955] have carried out a series of internal friction studies at low frequencies in alloys (e.g. Ag-Zn) by which they were able to show that quenching indeed appreciably enhances the diffusivity of both components. These observations have also been used to ascertain the fact that the diffusion does occur by a vacancy mechanism. The quantitative interpretation of changes in relaxation strength in terms of defect concentration etc. is very uncertain, however.

A beautiful example of the interpretation of enhanced diffusion is presented by the "dissociative" diffusion of copper and nickel in germanium containing dislocations, which will be discussed fully in Chapter XXXII.

Copper diffuses exceedingly rapidly through the diamond lattice of Ge and Si, at least if sufficient dislocations are present. In "dislocation free" germanium the diffusivity of copper is "normal", that is, very small. At first sight one has here a clear case of dislocation-promoted diffusion, the more so where it was found that the copper precipitates rapidly along the dislocations. However, calculation of the amount transported through the dislocations results in impossibly large values. According to FRANK and TURNBULL [1956], the influence of dislocations must rather be sought in their activity as *sources of vacancies*, in which the rapidly migrating *interstitial* copper atoms can be trapped. They then become substitutional and can subsequently form precipitates of Cu or Cu_3Ge or some other phase. Far away from the dislocation lines the available vacancy concentration is soon used up by the interstitial copper atoms and no precipitation as substitutional Cu can occur on a large scale, only the very small interstitial concentration is present. Near the dislocations, however, a continuous supply of vacancies is available. (See further section XXXII–2.)

A somewhat analogous precipitation process has been observed to occur in α-bombarded polycrystalline copper (BARNES [1958]). The α-particles all penetrate to a narrow layer at a certain depth inside the copper metal, and are there converted into helium atoms. These are practically insoluble in copper. They cannot precipitate (in the form of gas bubbles), however, since the establishment of equilibrium

of vacancies even after heating to 800 °C, is too slow to provide the necessary diffusivity and space. Heating nevertheless results in very pronounced formation of gas bubbles, but only near the grain-boundaries crossing the helium layer. These boundaries, like the dislocations in the previous example, can obviously provide rapidly enough sufficient new vacancies to maintain the equilibrium and compensate for the loss of vacancies captured by the helium atoms. These newly provided vacancies can in turn aid in the precipitation of helium atoms. The grain boundaries thus do not act as diffusion paths, they only act as vacancy *sources*.

No influence of this kind by single dislocations was detected in this case.

We shall discuss questions relating to nucleation and precipitation in Chapter XXI.

REFERENCES

ACHTER, M. R. and R. SMOLUCHOWSKI (1951); J. appl. Physics **22,** 1260.

BARDEEN, J. and C. HERRING (1951); Imperfections in nearly perfect crystals (Wiley, New York).

BARNES, R. S. (1959); Proc. Symposium on solid state diffusion, Saclay 1958 (C.E.N., Saclay).

BIRCHENALL, C. E. (1958); Metallurgical Reviews **3,** 235.

BARRER, R. M. (1941); Diffusion in and through solids (Cambridge Univ. Press).

BOLK, A. (1959); Thesis Delft Technical University.

BOSMAN, A. J., P. E. BROMMER, G. W. RATHENAU, C. J. SCHINKEL and L. C. H. EYKELENBOOM (1960); Physica **26,** 533.

BRINKMAN, J. A. (1954); Phys. Rev. **93,** 345.

BRINKMAN, J. A. (1955); Acta Metall. **3,** 140.

COMPAAN, K. and Y. HAVEN (1956); Transact. Faraday Soc. **52,** 786.

COULING, S. R. L. and R. SMOLUCHOWSKI (1954); J. appl. Physics **25,** 1538.

DARBY, J. N., T. TOMIZUKA and R. W. BALLUFFI (1958); Bull. Am. Phys. Soc. **3,** 124.

DOO, V. Y. and R. W. BALLUFFI (1958); Acta Metall. **6,** 428.

DUGDALE, H. (1954); Rep. Conf. Defects in Solids Bristol (The Physical Society, London).

FARA, H. and R. W. BALLUFFI (1959); J. appl. Physics **30,** 325.

FISHER, J. C. (1951); J. appl. Physics **22,** 74.

FRANK, F. C. and D. TURNBULL (1956); Phys. Rev. **104,** 617

GLASSTONE, S., K. J. LAIDLER and H. EYRING (1941); The Theory of Rate Processes (New York, McGraw-Hill).

GRAF, M. R. (1958); C. R. Acad. Paris **246,** 1544.

GUINIER, A. (1959); Proc. Symposium on solid state diffusion, Saclay 1958 (C.E.N., Saclay).

HAYNES, C. W. and R. SMOLUCHOWSKI (1955); Acta Metall. **3,** 130.

HENDRICKSON, A. A. and F. S. MACHLIN (1954); Transact. AIME **200,** 1035.

HEUMANN, TH. (1957); La Diffusion dans les Métaux, p. 59 (Philips Technical Library, Eindhoven).
HOFFMAN, R. E. and D. TURNBULL (1951); J. appl. Physics **22,** 634.
JOHNSON, R. P. (1939); Phys. Rev. **56,** 814.
JOST, W. (1952); Diffusion in solids, liquids and gases (Academic Press, New York).
KEYES, R. W. (1958); Ann. Rep. Inst. Study of Metals, p. 89.
KRAMERS, H. A. (1940); Physica **7,** 284.
LACOMBE, P. (1957); La Diffusion dans les Métaux, p. 23 (Philips Technical Library, Eindhoven).
LE CLAIRE, A. D. (1949); Progress in Metal Physics **1,** 306.
LE CLAIRE, A. D. (1953); Progress in Metal Physics **4,** 265.
LE CLAIRE, A. D. (1954); Phys. Rev. **93,** 344.
LE CLAIRE, A. D. (1957); La Diffusion dans les Métaux, p. 1. (Philips Technical Library Eindhoven).
LEE, and R. MADDIN (1959); Trans. Met. Soc. AIME **215,** 397.
NACHTRIEB, N. H. (1958); Ann. Rep. Inst. Study of Metals 1957–1958, p. 77.
NOWICK, A. S. (1955); see Progress in Metal Physics **3,** 1.
OKKERSE, B., T. J. TIEDEMA and W. G. BURGERS (1955); Acta Metall. **3,** 300.
SEITH, W. (1955); Diffusion in Metallen, 2nd ed. (Springer, Berlin).
SEITH, W. and A. KOTTMAN (1952); Naturwiss. **39,** 40.
SEITH, W., TH. HEUMANN and G. WALTHER (1955); Naturwiss. **42,** 532.
SEITZ, F. (1948); Phys. Rev. **74,** 1513.
SEITZ, F. (1951); Imperfections in nearly perfect crystals, p. 77 (Wiley, New York).
SHEWMON, P. (1958); J. chem. Physics **29,** 1032.
SHOCKLEY, W. (1953); Phys. Rev. **91,** 1563.
SMIGELSKAS, A. D. and E. O. KIRKENDALL (1947); Transact. AIME **171,** 130.
THOMAS, G. and M. J. WHELAN (1959); Phil. Mag. **4,** 511.
TICHELAAR, G. W. and D. LAZARUS (1959); Phys. Rev. **113,** 438.
TURNBULL, D. and R. E. HOFFMAN (1954); Rep. Conference Defects in Solids Bristol, (London, The Physical Society).
WAGNER, C. (1938); Z. physik. Chemie **B38,** 325.
YUKAWA, S. and M. SINNOTT (1955); Transact. AIME **203,** 996.

CHAPTER XIX

GRAIN BOUNDARIES

About this topic we shall be very brief, since the study of these imperfections in terms of the elementary lattice defects has been mainly in the hands of a few investigators (Frank, Shockley, Read, Amelinckx), and an admirable account of the structure and of the physical properties can be found in a paper by AMELINCKX and DEKEYSER [1959] and a book by MCLEAN [1957]. Moreover, a number of topics related to grain boundaries is discussed in other chapters of the present book.

1. Geometry of Grain Boundaries

Many properties of a polycrystalline aggregate are influenced by the presence of *grain boundaries* in it. Examples of this have already been mentioned here and there: creep (Chapter XII), diffusion (Chapter XVIII), internal friction (Chapter XVII). We shall now study the structure of these *planar* lattice faults more in detail.

A grain boundary constitutes the transition zone between two crystals in contact but differing in spatial orientation. To specify the nature of a particular plane grain boundary, *five* parameters are needed; three of these determine the relative orientation of the second crystal with respect to a coordinate system belonging to the first, and the remaining two define the orientation of the plane of the boundary in this same system.

To describe the atomic structure of a grain boundary two general models are in use. One of them (due originally to MOTT [1948]) has already occasionally been mentioned; it assumes "islands of fit and misfit" to be present in the transition layer between the two crystals. This transition layer tends to conform as much as possible to the

crystallographic structure, but a region of good fit cannot of course extend indefinitely, the strains becoming too large. Therefore, here and there narrow regions of the boundary zone occur in which no crystallographic symmetry at all is present any more. Such a (linearly extended) island of misfit is then again followed by a region of good fit (Fig. 1). The larger the orientation difference, the more frequent are the misfit regions, until at very large differences (above say 50°) the whole boundary can be regarded as one disordered zone, in which the state of the material can best be compared to the liquid configuration (SMOLUCHOWSKI [1951]). This model is, of course, not well suited to quantitative computations.

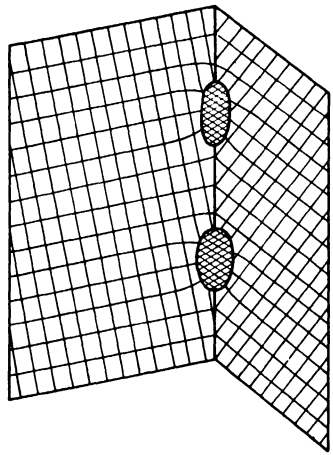


Fig. 1. Grain boundary model according to Mott. Reasonable fit between the crystallographic planes exists in part of the boundary, but these regions of good fit are separated by islands of misfit. These must be thought of as to extend as cylindrical regions perpendicular to the plane of drawing.

More accessible to quantitative attack is the *dislocation model* of grain boundaries, first proposed by W. G. and J. M. BURGERS [1939, 1940] and by BRAGG [1940]. This model constitutes in a sense a special case of the "island" model, in which the task of the regions of misfit is performed by groups of dislocations in the plane of the boundary. At first sight the dislocation model seems applicable for quantitative purposes only, if it is still possible to distinguish physically between the individual dislocations, that is, if these are spaced sufficiently wide apart (say more than 1 atomic spacing). This re-

striction would mean that the dislocation model is only suited for small-angle grain boundaries, an orientation difference of about 25° constituting the maximum for which it is theoretically applicable. There are exceptions to this rule, however; large-angle boundaries can also often be described by well-defined dislocation arrays, as has *inter alia* been pointed out by HORNSTRA [1959] (see Chapter XXIX for an example). Only, then partial dislocations etc. must also be included. This is equivalent to saying that practically every region

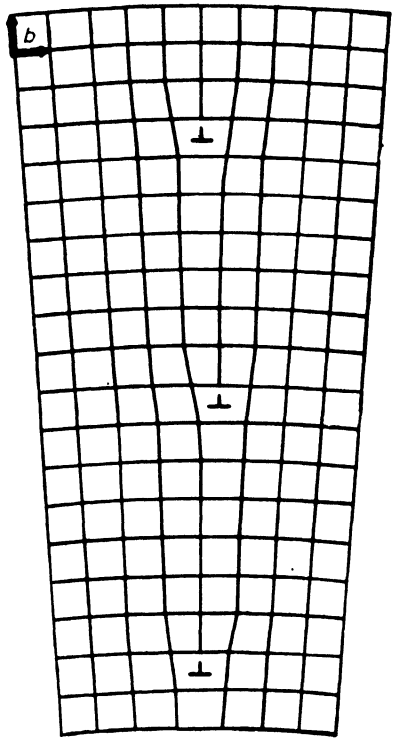


Fig. 2. Symmetrical tilt boundary in a cubic crystal built up out of one set of parallel edge dislocations (after READ [1953]).

of crystallographic misfit can be described in terms of a distribution of dislocations. Whether such a description is still of use to help in the most important problem of this field, namely the determination of the *energy* and *mobility* as a function of disorientation angle, is

questionable. Of the contributing dislocations in a large-angle boundary, it is only the *cores* and their topological properties (Burgers vector) that can be distinguished, since the elastic strain fields of the individual dislocations are largely cancelled by the influence of the other dislocations. To determine the energy of a dislocation core accurately is still very difficult, however.

Two very simple types of dislocation boundaries that yield to a quantitative description have been depicted in Figs. 2 and 3 (for the case of a simple cubic lattice). The first is a *tilt* boundary, oriented symmetrically with respect to both grains. It is entirely built up out of parallel pure edge dislocations with Burgers vectors b perpendicular to the joint (and with axes all parallel to a cubic direction). The orientation difference is a simple rotation over the angle θ. When the spacing between the dislocations is d, one has the already much quoted formula:

$$\frac{b}{d} = 2 \sin \frac{\theta}{2}. \tag{1}$$

Fig. 3 depicts a *twist* boundary. Whereas in the tilt boundary the two grains were rotated with respect to each other around an axis lying *in* the plane of the boundary, the twist boundary is characterized by a rotation around an axis *normal* to the boundary plane. In Fig. 3 moreover, this axis is a common cubic direction of the two simple cubic crystals. The twist boundary is built up by two mutually perpendicular sets of parallel screw dislocations.

There are two sets needed here, with the same dislocation spacing d, in order to ensure the fourfold symmetry in the boundary plane. Between spacing, Burgers vector and orientation angle the relation

$$\frac{b}{d} = \sin \theta \tag{2}$$

now holds, which for small θ is equivalent to (1).

In both examples, the misfit between the two grains is localized near the dislocation cores; these thus indeed play the role of "regions of misfit".

When the mutual orientation of the two grains is less symmetrical, or when there are no dislocations available with suitable Burgers vectors to form a symmetrical boundary along an arbitrary crystallographic orientation, the boundary is in general built up out of several

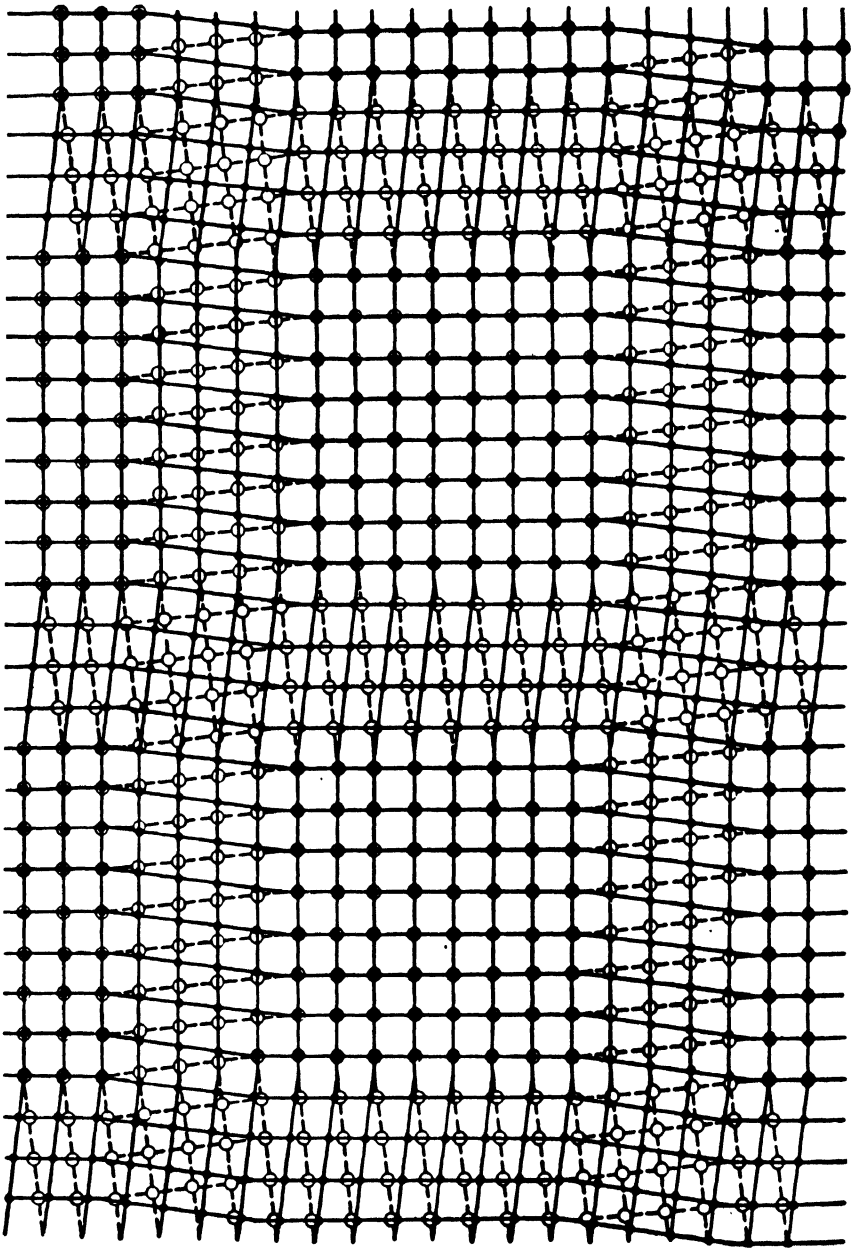


Fig. 3. Twist boundary consisting of two crossed systems of parallel screw dislocations (after READ [1953]).

sets of dislocations of mixed character. Formulae (1) and (2) are to be regarded as lower limits for the dislocation content of especially simple boundaries; usually many more, "less efficient" dislocations are needed to construct a general boundary. The properties of a grain boundary, for instance its energy, therefore depend strongly on the crystallographic boundary orientation, even when the disorientation angle between the grains remains constant. Specially symmetrical positions of the boundary plane (such as twin positions, for instance), correspond to prominent minima in the boundary energy.

It is not possible to construct a grain boundary of arbitrary orientation by simply superposing tilt and twist boundaries of the two types mentioned. FRANK [1950] has derived a general procedure for constructing the dislocation model of an arbitrary (low-angle) boundary in a given crystal structure, assuming this to be built up from straight dislocations with Burgers vectors which are minimum lattice vectors. READ and SHOCKLEY [1950] have extended the treatment to higher-angle boundaries. In the present discussion we follow READ's book [1953], in which an exact derivation of the formulae and conclusions can also be found. Suppose the crystals I and II have been rotated through an angle θ about the rotation axis u (Fig. 4). The rotation vector $u\theta$ will be called $\boldsymbol{\omega}$. The normal to the grain boundary be n. An arbitrary vector V attached to I and lying in the boundary plane, becomes, after rotation through $\boldsymbol{\omega}$, the vector V', the vector difference being $S = V' - V$. This vector difference must be equal to the vector sum of the Burgers vectors of all the boundary dislocations which are cut by the length of the vector V. Thus, when θ is small,

$$\theta(u \times V) = \boldsymbol{\omega} \times V = S = |V| \sum_i \nu_i b_i , \tag{3}$$

where ν_i is the number of dislocations of Burgers vector b_i present per unit length parallel to V in the boundary plane. This is the fundamental condition which the dislocation distribution must obey.

By making a number of suitable choices for V, dislocation configurations containing the minimum number of dislocations can be derived for each particular boundary. As an example, consider the tilt boundary of Fig. 2. Let the rotation (cube) axis be along [001], the boundary plane be [100]. First take V also along [001]. Then $S = V \times \boldsymbol{\omega} = 0$, no dislocations are cut, thus all boundary dislocations must be parallel to the rotation axis. Now take V parallel to [010]. Then $S = \theta V$, in

the [100] direction. The number of dislocations cut is thus proportional to θ and the vector sum of the Burgers vectors lies in the [100] direction.

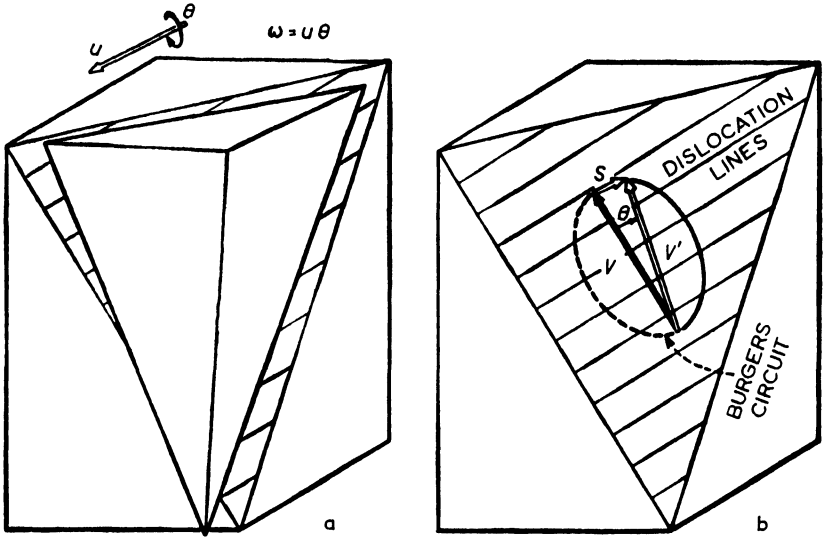


Fig. 4. To derive the general model of a dislocation grain boundary (after READ [1953]).

The simplest dislocation configuration that fulfills this condition is one set of pure edge dislocations parallel to [001] with Burgers vectors parallel to [100].

READ and SHOCKLEY [1952] have proved the validity of (3):

$$S(V) = \sum_i \nu_i (V) b_i \tag{3a}$$

for a general grain boundary (the brackets mean that S and ν_i both depend, of course, on V). For given rotation vector $\boldsymbol{\omega}$ and boundary normal n they derived *inter alia* the following general formula for the dislocation parameters ν_i and b_i. It is applicable to models containing three sets of dislocations with non-coplanar Burgers vectors $b_{1,2,3}$.

$$\nu_i = b_i^* \times \boldsymbol{\omega} - n(n \cdot b_i^* \times \boldsymbol{\omega}) , \tag{4}$$

where b_i^* is the reciprocal vector belonging to b_i in the space defined by the three b_i. For a further discussion the reader is referred to Read's book, and to two papers by BALL and HIRSCH [1955]. Recently AMELINCKX [1957] has still further extended the treatment to include also dislocation boundaries of the most arbitrary type.

The general problem how to find an unambiguous dislocation model for an arbitrary large-angle boundary, has not yet been solved. It might be noted here that a tilt boundary consisting of edge dislocations, which starts at the surface and ends somewhere in the interior of an otherwise perfect grain, forms an example of a Volterra dislocation of the rotational type mentioned in Chapter IV (KRÖNER [1958]).

HORNSTRA [1959] has approached the same problem not by an analytical but by a geometrical method, and has constructed dislocation models of grain boundaries in the diamond lattice that cover the range from 0° to 360° orientation difference in various boundary orientations. He made use of the core models of possible individual dislocations in the diamond lattice, to be described in Chapter XXIX.

2. Energy of a Dislocation Grain Boundary

The energy of a small-angle grain boundary can be mathematically derived as a function of its dislocation configuration and disorientation angle θ.

This problem has been extensively discussed by READ and SHOCKLEY [1950]. They divided the energy per constituent dislocation into three parts: 1) The elastic energy E_c of the dislocation core. 2) The elastic energy E_1 of the strain field contained within a distance R from the dislocation. This distance must be large compared to b but so small that interaction between the dislocations is still unimportant (R will be of the order of d, the spacing between the dislocations, in a simple tilt boundary). 3) The elastic energy E_2 contained in the material outside this distance, which is predominantly interaction energy.

When the disorientation angle of a boundary decreases, the core energy of the dislocation does not change, but E_1 increases, because d increases and hence R also. E_2 does not change, because it can be shown that the variations of R and d exactly cancel each other. It is therefore predominantly the strain energy E_1 of a single dislocation that changes with the orientation angle. Since dislocation stress and strain are both proportional to the inverse distance $1/r$, the energy density is proportional to $1/r^2$, and by integrating, the total energy change is found inversely proportional to R.

It is this property, the inverse proportionality with distance of the stress and strain associated with a dislocation, that determines fully the form of the angular dependence of the energy of a grain

boundary. As this property depends in turn wholly on the *topological* characteristics of dislocations (multivalued displacement, see Chapter IV), and not on the atomic structure of the core, the energy-orientation relationship of a grain boundary has a certain universal validity, as long as the boundary can be considered as an array of dislocations. Read and Shockley derived (see READ [1953]):

$$E = E_0\,\theta[A - \ln\theta] \qquad (5)$$

for the energy per unit area.

Here A is an integration constant, and E_0 depends on the type of dislocation. For a simple tilt boundary it is $Gb/4\pi(1-\nu)$; for the twist boundary it is $Gb/2\pi$ and for the general boundary containing several sets of dislocations the formula remains valid also, only E_0 is now a sum of terms, one corresponding to each set.

A tilt boundary of the type of Fig. 2, but not symmetric with respect to the two grains, must be built up from two sets of edge dislocations with perpendicular Burgers vectors of equal magnitude b. When φ denotes the angle inserted by the boundary and the symmetry plane between the crystals, one has

$$E = \frac{Gb}{4\pi(1-\nu)}(\cos\varphi + \sin\varphi)\,\theta\,[A - \ln\theta]\,. \qquad (6)$$

The derivation of (5) holds only at small disorientation angles, where a distance $R \gg d$ can still be defined within which interaction between individual dislocations is negligible. Read has shown that terms of order θ^2 should be added within the brackets of formula (5) when θ becomes larger than, say, 15°.

The constant A is an integration constant and must be fixed according to the case in hand. For the symmetrical tilt boundary in the simple cubic lattice the value of A follows from the argument that a "grain boundary" of $\theta = 0°$ can also be regarded as a "symmetrical" boundary with $\theta = 90°$. The latter must thus have zero energy. Substituting in (5) $\theta = 2\sin\theta/2$ for this "large-angle boundary", one finds $A = \ln(2\sin\pi/4) = 0.345$. In other cases the constant A cannot so easily be derived numerically. It is mostly approximately $\frac{1}{2}$, however.

The form of the E *vs.* θ dependence according to formula (5) has been depicted in Fig. 5. A maximum of the energy, E_m, occurs at $\theta = \theta_m = e^{-(1-A)}$. Apart from the invalidity of the derivation at higher

angles, near especially symmetric (twin) orientations the boundary energy has a sharp minimum (*cusp*). This complication has been indicated by the dotted curve. At larger angles, owing to the crystal symmetry, the curve often changes to a new branch.

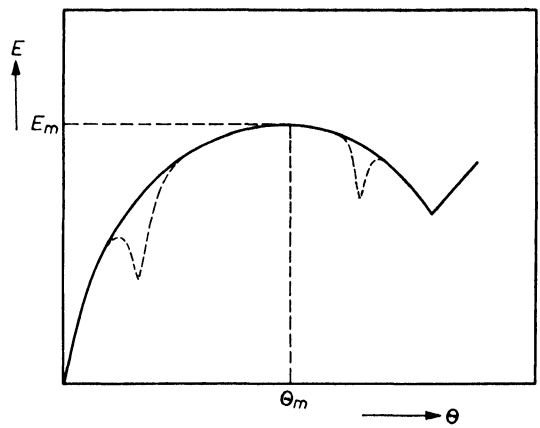


Fig. 5. E–θ relation for general dislocation boundary. The dotted cusps refer to twin and other symmetrical positions.

The first experimental verification of the form of formula (5) has been obtained in silicon-iron by Dunn and coworkers, while READ [1953] has compared measurements on different metals by plotting the *relative* grain boundary energy against the relative angle, according to the equivalent form of formula (5):

$$\frac{E}{E_m} = \frac{\theta}{\theta_m}\left\{1 - \ln\frac{\theta}{\theta_m}\right\}. \tag{7}$$

All observations (Fig. 6) fit remarkably well to this dimensionless relation, even at angles up to twice the optimal angle θ_m (which usually lies between 10° and 30°). This is not very suprising, in view of the arguments advanced above for the general applicability of (5) and in view of the rather liberal way in which the scale of the experimental curves can be fitted to the standard curve. Moreover, the results of different observers with respect to θ_m are rather at variance, however; whereas FISHER and DUNN [1952] found $\theta_m = 25°$ for tilt boundaries in Cu, Ag and Pb (all face-centred cubic metals), AUST and CHALMERS [1950] found $\theta_m = 36°.5$ in silver boundaries of accu-

rately symmetrical orientation. For a further discussion the reader is referred to a paper by H. BROOKS [1957] and to McLean's book.

The measurement of relative grain boundary energies is difficult and cannot be commented upon here; we refer to BURKE and TURNBULL [1952] and CHALMERS [1952] for a discussion and again to McLean's book

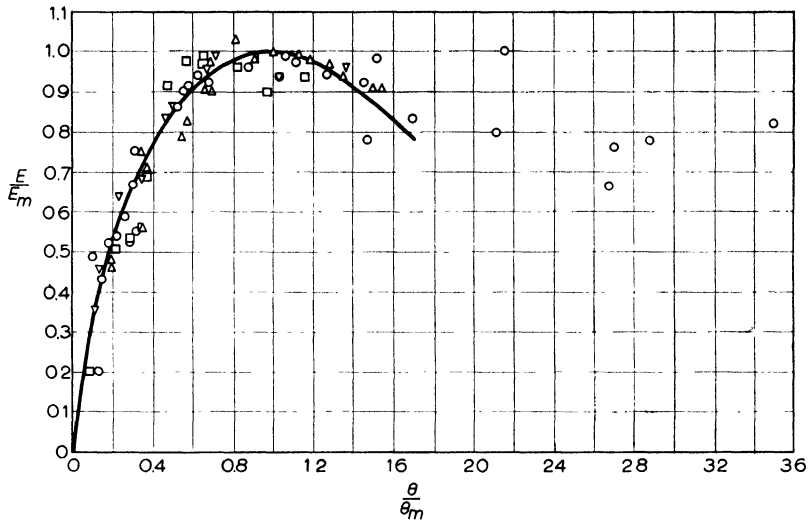


Fig. 6. Comparison of theoretical E–θ relationship (7) with observed relative grain boundary energies (after READ [1953]).

○ Dunn silicon iron (110) series, $\theta_m = 26.6°$;
△ Dunn silicon iron (100) series, $\theta_m = 29.8°$;
□ Aust and Chalmers tin, $\theta_m = 12.2°$;
▽ Aust and Chalmers lead, $\theta_m = 25.0°$.

for an extended review. The determination of the absolute boundary energy is even more laborious. The ratio of the energy of a grain-boundary to the surface energy (to be determined independently) can for instance be obtained from the groove-angle of a grain boundary groove produced by etching the surface. Another way is by making a statistical study of dihedral angles in multiphase systems, and determining independently one of the interfacial energies (e.g. of a solid-liquid interface). In this way E_0 from (5) for a tilt boundary was found to be about 1000 ergs/cm^2 in silver, 1390 ergs/cm^2 in copper and near to 2000 ergs/cm^2 in iron.

3. Motion of Dislocation Grain Boundaries

A dislocation grain boundary can move when all its dislocations move together. This motion will be particularly easy when all dislocations can glide. A boundary for which this is possible has been called *glissile* by Read. The simple tilt boundary of Fig. 2 is such a glissile boundary; Fig. 7 illustrates its glide motion as brought about by a shear stress. One of the most striking confirmations of the dislocation theory of grain boundaries was the observation by PARKER and WASHBURN [1952], that an artificially made 2° tilt boundary in a zinc bicrystal could indeed be moved by applying a shear stress parallel to it. The motion of the glissile boundary was reversible under the stress, and in general conformed exactly to the predictions of dislocation theory. BELL and CAHN [1957] found that even a 20° boundary in zinc showed some mobility under stress.

Usually, the distribution of dislocation orientations and types in an arbitrary grain boundary is such that no glissile motion is possible. Then, for grain boundary migration to occur, diffusive processes must come to the aid. For instance, vacancies are created at dislocations of one set, migrate to dislocations of another set and disappear there. The motion will then be thermally activated and cannot be produced by action of a shear stress alone. It is possible sometimes that the boundary splits into several components, each of which moves in pure glide, but of course into a different direction. In general, however, the stresses between the various dislocations that constitute the boundary oppose such a splitting. We shall not go into further details here. READ [1953] calculated that the stress needed for slip motion of a composite boundary is of the order of $\theta G/2\pi$; when $\theta \geqslant 1°$ this is usually much larger than the critical shear stress for plastic flow of the material.

4. Large-angle Grain Boundaries

For large-angle grain boundaries the dislocation model does not lend itself to numerical analysis, except in special symmetrical orientations. The model of Mott, mentioned in section 1, which is in a sense a generalization of the older "cement" theory, in which a layer of amorphous "molten" material was supposed to exist between the grains (*Beilby layer*), can now be used. This model has been applied with fair success to the problem of stress-induced diffusive grain boundary *migration*

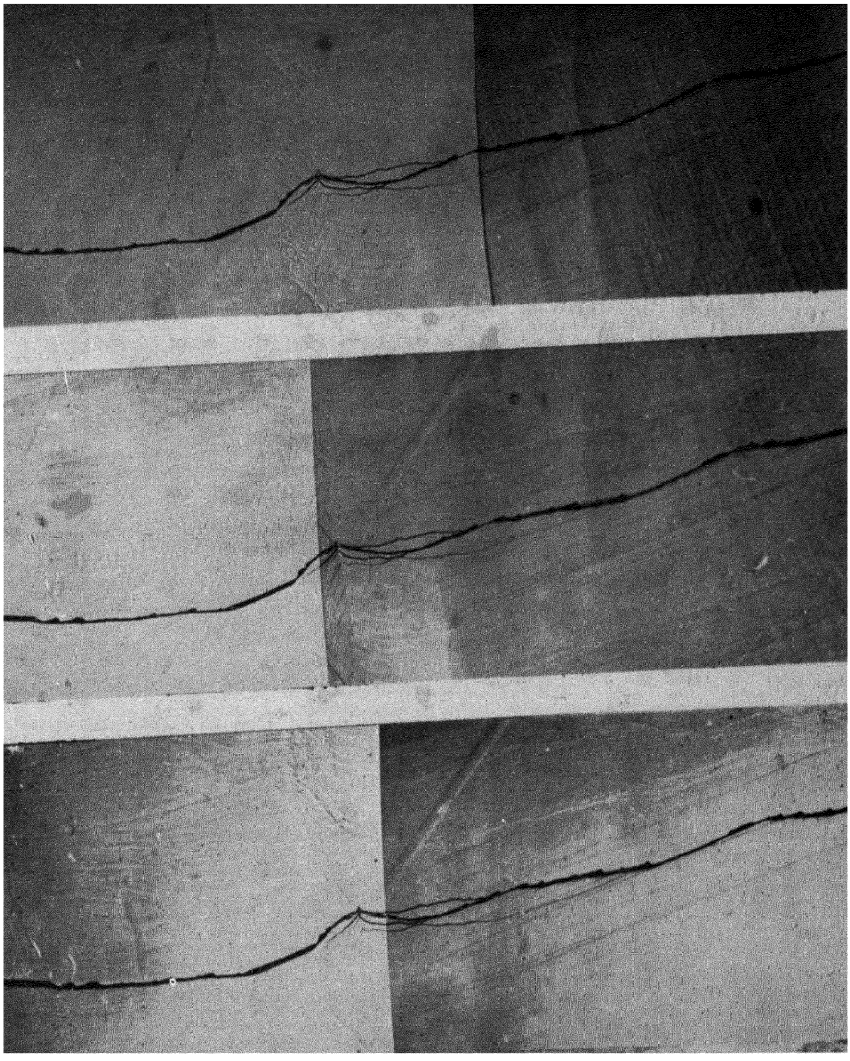


Fig. 7. Motion of a simple tilt boundary under an applied shear stress.
(After PARKER and WASHBURN [1952].)

(see also grain boundary creep, Chapter XI). The observed energy of activation for this process is interpreted as the energy necessary to disorder the atoms within each region of good fit; compare next chapter.

By interpreting the observed (rather large) values of the activation energy for grain boundary creep in this model it follows that the number of atoms n within such a region is usually larger than 1000. Another remarkable point is the simultaneous occurrence of a large pre-exponential factor (of the order of $10\,\sigma$ cm/sec, σ in dyn/cm^2). Although Mott's theory can indeed explain such a large factor, it has been suggested that the simultaneous occurrence of a large frequency factor and a large activation energy is the result of an erroneous interpretation of data, and that actually the activation energy for *grain boundary diffusion* should be rate-determining (McLEAN [1957]).

SMOLUCHOWSKI [1951] has extended Mott's theory to include the effects of (free) surface energy and of crystallographic structure. He added a term $nsa^2(T/T_m)$ to the free energy for disordering, where s is the surface energy of the grains: a large surface energy facilitates the migration of the boundary. WALTER and DUNN [1958] have advanced arguments to show that the so-called tertiary recrystallization of Si-Fe sheet is to a large extent determined by the energy of the crystal faces exposed to the *free* surface. For a review of these and other properties of grain boundaries see WEINBERG [1959].

5. Interaction between Grain Boundaries and Other Lattice Imperfections

In this section we want to mention briefly two additional features of the properties of a grain boundary that have great importance for the physics of metals. One of them is the interaction between grain boundaries and dislocations, the other between grain boundaries and impurities.

5.1. INTERACTION WITH IMPURITIES AND POINT DEFECTS

The interaction between grain boundaries and localized imperfections in the grains is (in metals) predominantly of an elastic nature. The stress field of a dislocation boundary (the only kind we will discuss in this section) can be regarded as the superposition of the stress fields of the constituent dislocations. At small distances from the boundary, viz. less than the mean dislocation spacing, the stress field

is essentially that of the nearest single dislocation. Therefore, once an impurity has reached this zone—which extends, however, over a few atomic spacings only from the boundary plane (5 spacings for $\theta = 10°$ and even less for higher angle boundaries)—the interaction can be treated by the same methods as were used in Chapter IX. The theory of Cottrell and Bilby applies, and the impurity will be attracted by a certain force and bound to the dislocation with a certain energy, which is of the order of a few tenths of an eV in most metals.

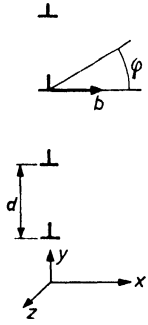


Fig. 8. On the derivation of the dilation field of a tilt boundary.

At larger distances, however, thus in most of the volume of each grain, it can be shown that the stresses due to a grain boundary decrease much faster, i.c. exponentially, with the distance. An edge dislocation at the position $(0,kd)$ in a tilt boundary lying in the (yz)-plane (Fig. 8), has associated with it a dilatation $\varepsilon_{xx} + \varepsilon_{yy}$ (compare III–(4)):

$$\varepsilon_{xx} + \varepsilon_{yy} = -\frac{b}{2\pi}\,\frac{1-2\nu}{1-\nu}\,\frac{y-kd}{x^2+(y-kd)^2}. \tag{8}$$

By summing this expression over all entire k, one obtains the field of dilatation of the whole boundary (see for the method of summing COTTRELL's book [1953]):

$$\Delta = -\frac{b}{4d}\,\frac{1-2\nu}{1-\nu}\,\frac{\sin(2\pi y/d)}{\sin^2(\pi y/d)+\sinh^2(\pi x/d)}. \tag{9}$$

When $x \gg d$, the sinh can be replaced by the exponential function and

$$\Delta \sim \mathrm{e}^{-\pi x/d}\,; \tag{10}$$

indeed an exponential decrease. The influence of the sine terms in (9)

is of course independent of the distance from the boundary. Therefore, the discrete dislocation structure of the boundary makes itself known as far as the stress field itself is notable. The energy of interaction between a solute atom of radius r, which is a factor $(1 + \varepsilon)$ too large for its available site, and a tilt boundary is now easily found as

$$U = \frac{2\pi G b \varepsilon r^3}{d} \frac{1 - 2\nu}{1 - \nu} \frac{\sin(2\pi y/d)}{\sin^2(\pi y/d) + \sinh^2(\pi x/d)} \tag{11}$$

(here as usual only the interaction due to hydrostatic stresses is considered).

By integrating (11), WEBB [1957] showed that the number of solute atoms needed to saturate all the dilatation of one sign associated with the boundary (that is, to compensate for all volume changes) is

$$N_0 = \frac{b}{32\pi r^3 \varepsilon} \frac{1 - 2\nu}{1 - \nu} \tag{12}$$

per unit area of the boundary. This is independent of the angle of disorientation, since the larger this angle, the smaller is d and the greater is the mutual compensation of the strains due to neighbouring dislocations, a factor which cancels the increase of dislocation length per unit area of the boundary. A typical value of N_0 is 10^{14} cm^{-2}; in a metal of grain size 0.1 mm the total area of grain boundary is 10^6 cm^{-1}, and a concentration of 10^{20} solute atoms per cm^3 or about 0.1% is thus needed to saturate all the boundaries. Since this is very large, the boundaries often remain unsaturated; the degree of saturation depends on concentration c and temperature T through the binding energy of the solute atom to a dislocation. Webb, on the basis of a rather artificial statistical model, obtained for the equilibrium concentration N the expression

$$\frac{N}{N_0} = c \frac{kT}{A} \frac{1}{\exp\left(-\frac{U^*}{kT}\right) + c\frac{kT}{A}}, \tag{13}$$

where A is a quantity of order 10^{-11} erg (dependent on the disorientation angle), and U^* is the maximum value of U from (11), i.e. the binding energy in the boundary; c is the bulk solute concentration.

Accurate observations of the segregation of impurities to grain boundaries have been forthcoming during the last few years. A typical

influence of the disorientation angle on the interaction is certainly present, and this is not explained by Webb's theory. Also the binding energy which is derived by a comparison of this theory with experiment is rather small.

LÜCKE and DETERT [1957] have studied the interaction between impurities and dislocations in a much simpler theoretical model, but they have compared the theoretical results with a large number of observations of the influence of impurities on the motion of grain boundaries during recrystallization. Satisfactory agreement could be obtained by assuming the presence of a simple Cottrell-type elastic interaction only.

J. W. CAHN [1958] has approached this problem in still a different way. We refer to his paper for particulars.

5.2. INTERACTION BETWEEN DISLOCATIONS AND GRAIN BOUNDARIES

One aspect of the interaction between grain boundaries and dislocations, namely the passing of slip bands through boundaries and the propagation of slip from grain to grain, has already been discussed in Chapter XIII. We only mention here Ogilvie's work on the influence of the orientation of the line of intersection of the two slip planes in the neighbouring grains. Only when this line is close to [110], [112] or [123] in either crystal (but not necessarily to the same direction in both), slip lines are observed to pass from one grain to the other, thereby, of course, changing their direction as observed on the surface of the metal.

TEISSIER DU CROS has pointed out [1955] that by modifying Read's model of a dislocation boundary in such a way that the boundary becomes asymmetric, sometimes a smaller number of dislocations per cm is needed, although the orientation difference is not altered. A structure called by him "leaved structure" (structure feuilletée) is formed, of which it can be expected that slip dislocations can more easily pass through it than through normal boundaries.

6. Special Types of Grain Boundaries; Subboundaries; Coherent and Semi-coherent Boundaries

Within the grains of a polycrystalline material or in a single crystal so-called *subboundaries* can often be observed by suitable microscopic or X-ray techniques. These subdivide the grain into a number of sub-

grains, with mutually but slightly different orientations. The separation interfaces are typical very-low-angle dislocation boundaries, the disorientation angle being mostly of the order of a few minutes of arc. This corresponds to a dislocation spacing by several hundreds of atomic distances or more, or to a dislocation content of the order of only 10^5 cm per cm^2 of the boundary. As the sizes of the subgrains vary between 10^{-5} and 10^{-3} cm, the total dislocation content in the subboundaries is thus 10^8–10^{10} dislocations per cm^2.

Subboundaries can be generated during various treatments. Thus, for instance, as has been mentioned in Chapter XII, a subgrain structure is observed to form in aluminium crystals which are slowly deformed in creep at a few hundreds of degrees centigrade. WOOD [1949] has called this kind of structure a *cell-structure*. Annealing of a cold-worked cubic crystal often results in the formation of subgrains by *polygonization* (see Chapter XI). Probably the cell formation during creep is the result of simultaneous deformation and polygonization in the crystal. Then, the phenomena of fragmentation and "recrystallization in situ" (Chapter XI) occurring during cold work give rise to a subgrain structure, and last but not least some types of mosaic structures in well-annealed metals can better be described as an actual subgrain structure than as a dislocation network (HIRSCH [1956]).

The prevalence of boundary-like dislocation structures over more arbitrary dislocation networks, especially when the material has undergone a high-temperature anneal, is connected with the decrease of the elastic energy content of the crystal on the arrangement of the dislocations in walls. The production of subboundaries during the deformation itself has been explained by McLean as the result of local rearrangement of the dislocations in slip bands, by which, of course, also a decrease in elastic energy is obtained. Many authors (see e.g. T. SUZUKI [1954]) have studied the relative stability of substructures under stresses and thermal vibrations. We refer to the original work for further particulars.

Sometimes no dislocations at all or only very few are needed to form a medium-angle boundary between two grains. An obvious case is that of a coherent twin boundary (Fig. 9 illustrates a coherent twin boundary in a face-centred cubic crystal along a (111)-plane; $\theta = 70°32'$). There are usually many possibilities for such coherent boundaries. In copper coherent (111)-twin boundaries were found to represent only about 1/30 of the energy of an average (111)-grain boundary.

Also in other metals similar reductions have been observed (HALL [1955]). Non-coherent twin boundaries represent a much higher energy although still less than the average boundary. In copper FULLMAN [1951] found about 0.8 for the ratio of the mean non-coherent twin boundary energy to the mean grain boundary energy.

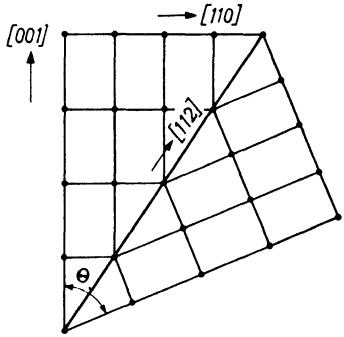


Fig. 9.
Symmetrical twin across (111)-plane in face-centred cubic lattice.

Another low-energy boundary occurs when two grains of *different* substances (or of *different phases* of the same substance) are oriented in such a way that along their common boundary a regular periodic transition structure somewhat resembling a dislocation wall is present (semi-coherent phase boundary, Widmanstätten structure, epitaxial deposit, see Chapters XX and XXIII). An example forms the boundary between martensite and austenite (FRANK [1953]). A final category of low-energy boundaries are the kinkbands discussed in Chapter XIII.

REFERENCES

AMELINCKX, S. (1957); Physica **23**, 663.
AMELINCKX, S. and W. DEKEYSER (1959); Solid State Physics **8**, 441.
AUST, K. T. and B. CHALMERS (1950); Proc. Roy. Soc. **A204**, 359.
BALL, C. J. and P. HIRSCH (1955), Phil. Mag. **46**, 1343; (1957) **2**, 977.
BELL, C. and R. W. CAHN (1957); Proc. Roy. Soc. **239A**, 494.
BRAGG, W. L. (1940); Proc. phys. Soc. **52**, 54.
BROOKS, H. (1957); Seminar on Metal Interfaces (A S M).
BURGERS, J. M. (1939); Proc. Kon. Ned. Akad. Wet. Amsterdam **42**, 293 and 378
BURGERS, W. G. (1940); Proc. phys. Soc. **52**, 23.
BURKE, J. E. and D. TURNBULL (1952); Progress in Metal Physics **3**, 220.
CAHN, J. W. (1958); Acta Metall. **6**, 772.

CHALMERS, B. (1952); Progress in Metal Physics **3, 293.**
COTTRELL, A. H. (1953); Dislocations and plastic flow of crystals (Oxford, Clarendon Press).
DUNN, C. G. and P. K. KOH (1958); Metall. Trans. AIME **212,** 80.
FISHER, J. C. and C. G. DUNN (1952); Imperfections in nearly perfect crystals (Wiley, New York).
FRANK, F. C. (1950); Pittsburgh Conference on plastic deformation, p. 150 (ONR, Washington).
FRANK, F. C. (1953); Acta Metall. **1,** 15.
FULLMAN, R. L. (1951); J. appl. Phys. **22, 448.**
HALL, E. O. (1955); Twinning (Butterworth Sci. Publ.).
HIRSCH, P. (1956); Progress Metal Physics **6, 236.**
HORNSTRA, J. (1959); J. Phys. Chem. Solids **4,** 129.
KRÖNER, E. (1958); Kontinuum Theorie der Versetzungen (Springer, Berlin).
LÜCKE, K. and K. DETERT (1957); Acta Metall. **5, 628.**
MCLEAN, D. (1957); Grain boundaries in metals (Oxford, Clarendon Press).
MOTT, N. F. (1948); Proc. phys. Soc. **60, 391.**
OGILVIE, J. (1953); Inst. Metals **81,** 491.
PARKER, E. R. and J. WASHBURN (1952); Transact. AIME **194, 1076.**
READ, W. T. (1953); Dislocations in crystals (McGraw-Hill, New York).
READ, W. T. and W. SHOCKLEY (1950); Phys. Rev. **78, 275.**
READ, W. T. and W. SHOCKLEY (1952); Imperfections in nearly perfect crystals, (Wiley, New York).
SMOLUCHOWSKI, R. (1951); Phys. Rev. **83,** 69.
SUZUKI, T. (1954); Rep. Conference Defects in Solids, Bristol (The Physical Society, London).
TEISSIER DE CROS, F. (1955); The plasticity of metals, Colloquium Madrid (Springer, Berlin).
WALTER, J. L. and C. G. DUNN (1958); J. Metals **10,** 509.
WEBB, W. W. (1957); Acta Metall. **5,** 89.
WEINBERG, F. (1959); Progress in Metal Physics **8,** 105.
WOOD, W. A. (1949); Series of papers in J. Inst. Metals **76.**

CHAPTER XX

RECRYSTALLIZATION AND GRAIN GROWTH

In the following four chapters we shall discuss a number of topics which, although they stand in the centre of attention of metallurgists because of their great technical importance, have not yet been understood physically with such a degree of reliability that a more or less definitive discussion of the underlying physical mechanisms could here be given. This, and the circumstance that for all five topics: recrystallization, precipitation, fatigue, fracture and crystal growth excellent and recent review papers and books are available, have induced the present writer to restrict himself to give only a brief outline of the salient aspects of these phenomena. It may well be that in a few years a complete and reliable physical theory of these processes will have been developed; until that time the comparison of the relative merits of the various ideas proposed and mechanisms put forward may better be left in the hands of specialists in the field.

1. Phenomenology

On annealing a cold-worked metal above a certain temperature, depending on the nature of the metal and on the degree of cold-work *primary recrystallization* usually occurs. That is, the old grains are replaced by new ones, the size, texture, etc. of which differ completely from those of the structure previously present. After recrystallization, all work-hardening has disappeared.

An attempt to discuss in the space of a few pages the numerous experimental details, speculations and theories that relate to the extended realm of recrystallization must fail. We have therefore selected only one or two details, that seemed of especial importance for other

parts of this book and refer for more complete accounts to review papers like those by BURGERS [1941], BURKE and TURNBULL [1952] and BECK [1954].

Recrystallization proper comprises two processes, a *nucleation* process and a *growth* process. Given the degree of deformation and the temperature of annealing, the fraction of recrystallized material and the mean size of the newly formed crystals both increase in time according to the (schematic) curves of Fig. 1. First an incubation period, during which nothing visible happens, must elapse before recrystallization starts. Afterwards the new grains grow at an approximately constant rate, until finally the stable end state is reached.

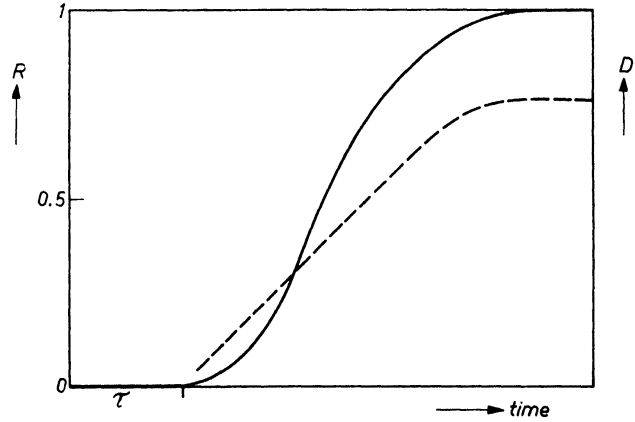


Fig. 1. Recrystallized fraction R (drawn curve) and mean diameter of newly formed grains D (dotted curve) as a function of time (schematic); τ denotes the incubation time.

Under certain conditions renewed grain growth occurs on further increasing the temperature. The kinetics of this *secondary recrystallization* resemble those of the primary recrystallization: again only after an incubation time has elapsed some grains begin to grow rapidly at the expense of others. Secondary recrystallization occurs most conspicuously when the primary recrystallization has been partially inhibited, for instance by the presence of inclusions obstructing the migration of the grain boundaries during this first process. Recently, WALTER and DUNN [1959] and COULOMB and LACOMBE [1959] have introduced the concept of *tertiary* recrystallization, a final growth of some grains left over after secondary recrystallization, the occurrence of which may be determined by the influence of the *free*-surface energy.

Sometimes, on annealing an *un*worked or slightly cold-worked polycrystalline metal, the grains begin to grow spontaneously, without an incubation time etc. being present. This is called *grain growth*, it is essentially a process of grain boundary migration and not of nucleation (BECK *et al.* [1950]).

The observed parallelism between recrystallized fraction and new grain size (Fig. 1) makes it obvious that the observed incubation time is indeed a *nucleation* time. As has been said, the grain size D is generally found to be proportional to the time t that has elapsed since the incubation time ended. Hence for the growth process it is possible to define a characteristic rate of growth G according to the formula:

$$D = G\,(t - \tau)\,. \tag{1}$$

The nucleation parameter τ indicates the inverse frequency with which nuclei of new grains are formed per unit volume of unrecrystallized material:

$$\mathrm{d}n = \frac{\mathrm{d}t}{\tau} = \dot{N}\mathrm{d}t\,, \tag{2}$$

where $\dot{N} = 1/\tau$ is called the nucleation frequency. When after a time t the volume fraction R of a sample has already been recrystallized, the number of new nuclei formed per second per unit volume of all the material (recrystallized and unrecrystallized together, *homogeneous* nucleation) is

$$\mathrm{d}n = (1 - R)\,\frac{\mathrm{d}t}{\tau}\,. \tag{3}$$

According to (1), the recrystallized volume of a given grain at the time t is given by

$$v = fG^3\,(t - \tau)^3\,, \tag{4}$$

where f is a form factor (of order unity). Since now evidently

$$R(t) = \int_0^t v\mathrm{d}n\,,$$

one obtains the following, so-called *von Göler-Sachs integral equation* for the kinetics of recrystallization:

$$R = fG^3 \int_0^t (t - \tau)^3 (1 - R) \frac{dt}{\tau} . \tag{5}$$

The equation holds only when the new grains are few and small and do not yet inhibit each other's growth. In the more exact theory developed by AVRAMI [1939] the interaction between grains is also considered, and on the left hand side of (5) R is replaced by $\ln [1/(1 - R)]$. By this alteration the formula retains its validity also for $R \to 1$, where the interaction between grains is certainly important. In two- and one-dimensional specimens or when the new grains occur in the form of platelets or needles, the third power in (4) and (5) must of course be replaced by a correspondingly smaller exponent (2 or 1).

JOHNSON and MEHL [1939] assumed the nucleation frequency $\dot{N}$ to be independent of the time; then equation (5) can be solved, leading to

$$R = 1 - \exp (-\tfrac{1}{4} fG^3 \dot{N} t^4) . \tag{6}$$

On the other hand, Avrami assumed that the number of nucleation sites is gradually used up and the nucleation frequency decreases according to:

$$\dot{N} = \dot{N}_0 \exp (-\nu t) . \tag{7}$$

This leads to a more intricate time dependence of R, in which, in the exponential, the time occurs to a power between 3 and 4 (for the three-dimensional case). It will be clear that recrystallization, in this model, cannot be regarded as a simple low-order rate process; this is also apparent from the "sigmoid" form of the growth curve. The typical time dependence of the nucleation rate has been verified more or less by several experiments, a description of which can be found in the review paper by Burke and Turnbull.

It has been proposed by others that recrystallization might nevertheless be a first order rate process, which, however, occurs only in the part of the material that has already undergone a certain previous *conversion*. This conversion also follows first order kinetics:

$$\omega = \text{fraction converted} = 1 - e^{-\beta t} . \tag{8}$$

The conversion is identified simply as the mechanical recovery discussed in Chapter X. Since the rate of recrystallization is in turn proportional to

$$\frac{dR}{dt} = k\omega (1 - R) , \tag{9}$$

one obtains, in the approximation that ω is small:

$$R = 1 - \exp (-\tfrac{1}{2}k\beta t^2) , \tag{10}$$

which may be compared to formula (6). The exponent of t is smaller than in that formula (for the three-dimensional case). One might, in addition, assume a quadratic exponential time dependence of ω to remove even this difference. In the recovery theory proper nucleation is no longer present, it is replaced by a sort of *pre-formation* of the material.

A great number of "external" factors influence the recrystallization kinetics, the most important of which is the amount of previous cold work. A minimum deformation is necessary to produce recrystallization at all; the larger the deformation is, the more rapid recrystallization takes place, but the smaller are the newly formed grains. Recrystallization is clearly a temperature-activated process (or a sequence of processes), the activation energy of which depends on the amount of cold work: the temperature where recrystallization becomes noticeable decreases with increasing cold-work, as does the duration of the process at a given temperature. Also the structure of the original material is of importance; coarse-grained material is more difficult to recrystallize and needs larger amounts of pre-deformation than fine-grained material. Finally, the presence of impurities usually retards recrystallization.

The effect of impurities on recrystallization has recently been demonstrated dramatically by Chaudron and his coworkers. They prepared aluminium of exceptional purity (<0.0001% imp.) by the *zone refining* process, and found that this material recrystallized at room temperature already after very slight handling, and even at −60 °C after moderate cold work. A functional relation between purity, amount of pre-deformation and recrystallization temperature was obtained by ALBERT and DIMITROV [1957] and by BLADE *et al.* [1959]. This observation seems to rule out the recovery theory, as the latter process occurs only at 300 °C or higher. MASING *et al.* [1958] showed that addition of 0.01% Ni or Mn to 99.997% Al decreased the recrystallization rate at a certain temperature by a factor 10^{11}. More examples of the influence of impurities are given in the Burke-

Turnbull paper; a theoretical explanation has been attempted by LÜCKE and DETERT [1957].

Quite generally, the recrystallization process appears to need a certain degree of *structural disorder* of the crystallographic lattice. This disorder can be described in terms of dislocations, point defects, grain boundaries, etc. These have an effect equally well on the nucleation rate $\dot{N}$ as on the rate of growth G. Both quantities depend on temperature according to

$$(G,\dot{N}) = A\mathrm{e}^{-Q/kT}\,, \tag{11}$$

that is, through an activation energy (which may be different for the two quantities).

It is very difficult to ascertain experimentally the quantitative dependences of the parameters A and Q on the pre-strain, the temperature, the impurity content, the time, etc.. We have tried to give in Table 1 at least the general trends (basing ourselves on the Burke and Turnbull paper). It must be emphasized that the relations mentioned there are approximations to the average of the actually observed, rather indefinite, relationships, and that numerous exceptions to this scheme occur.

TABLE 1

Approximate trends of the variation of the parameters A and Q of the rate of growth and of nucleation, respectively, upon pre-strain ε, temperature T, impurity content c_1, orientation θ and time t.

		ε	T	c_1	θ	t
Rate of growth	A	decreases with increasing ε	often decreases with increasing T	decreases rapidly with increasing c_1	preferred directions of growth; a minimum misorientation is needed to make growth possible	none until grains meet
	Q	decreases with increasing ε	decreases sometimes with increasing T	none? (except in very pure material)		none
Rate of nucleation	Q	increases expon. with increasing ε		none?	—	decreases at large t

2. Nucleation and Growth of Grains—Formal Theory

The nucleation of new grains (or of new phases for that matter, see Chapter XXI) takes place since in forming a new grain the free energy is lowered (for instance becauees intrnal stresses are removed). In *homogeneous* nucleation theory the free energy of a nucleus can be written as the sum of a surface free energy, proportional to the surface area of the nucleus, and a volume free energy, proportional to its volume. Therefore, when i is the number of atoms in a spherical nucleus, its free energy is

$$\Delta G = ai^{2/3} + bi \,. \tag{12}$$

This is the starting point of the classical nucleation theory, a description of which has been given recently by HOLLOMON and TURNBULL [1953].

The quantity ΔG reaches its maximum when

$$i = i^* = (2a/3b)^3 \,; \tag{13}$$

this maximum is

$$\Delta G^* = \frac{4}{27}\,\frac{a^3}{b^2} \,. \tag{14}$$

Nuclei containing more than i^* atoms will tend to grow, those containing less will tend to disappear. The nucleation period τ discussed above can now be regarded in this model as the mean time needed to form a nucleus of critical size i^* under the action of thermal fluctuations in the matrix.

This time is proportional to

$$\exp\left(\frac{\Delta G^*}{kT}\right),$$

and ΔG^* thus forms part of the activation energy, previously called Q, which determines the temperature dependence of the nucleation rate. The rate of nucleation is further necessarily proportional to the number of atoms per unit volume of the matrix, B, to the surface area of the critical nucleus, thus to $i^{*2/3}$, and to the frequency with which the atoms move through the boundary (interface) between the matrix and the recrystallized nucleus. This transfer is also characterized by a free energy of activation, ΔG_A (TURNBULL and FISHER [1949]). Hence

we obtain formally for the nucleation rate the expression

$$\dot{N} \sim \nu_0 \, i^{*2/3} \, B \exp - \left\{ \frac{\Delta G_A + \Delta G^*}{kT} \right\}, \tag{15}$$

where ν_0 is, as usual, the atomic vibration frequency (10^{13} sec^{-1}) or a quantity directly related to it. Expression (15) holds when the nucleation rate is independent of time. Often this condition is not fulfilled (see above) and a time dependent factor must still be added. A plausible form of this factor is for instance (see (7)):

$$\exp - \left\{ \nu t - \frac{\theta}{t} \right\}, \tag{16}$$

where the first term between brackets describes the exhaustion of nucleation sites, the second term a possible initial transient in the nucleation rate.

The rate of grain growth, G, is determined by the free energy ΔG_A alone. ΔG_A is usually smaller than the activation energy for volume self-diffusion, and is more in accord with the energy for *grain boundary self-diffusion* (see Chapter XVIII). However, the *entropy* of activation ΔS_A for grain growth is unusually large, thus in the quantity A of formula (11) a large entropy contribution is contained. This has been explained by MOTT [1948] along the lines of his "local melting theory" mentioned in the previous chapter, in which the simultaneous movement of groups of a large number of atoms is considered.

Another theory (BURKE [1949]) connects the large entropy factor in the rate of grain growth with the effect of inclusions or other stress centres on the strain energy in the matrix. In general the retarding effect will be less at higher temperatures and therefore the observed temperature dependence of the rate of grain growth will be much steeper than is due to the exponential in (11) alone. This again results in an apparently larger entropy factor.

Also when the activation energy for grain growth, ΔG_A, itself depends on temperature, e.g. when it is smaller at higher temperatures owing to the effects of recovery, a similar apparent increase of the activation entropy would follow.

The *driving force* for recrystallization can be identified with the energy S per unit volume stored after the pre-deformation. The constants a and b in (12) can thus be expressed in terms of S and the

surface energy σ of a stress-free nucleus in the deformed matrix. For a spherical nucleus of radius r, one finds:

$$\Delta G^* = \frac{16\pi}{3}\frac{\sigma^2}{S^2}; \tag{14a}$$

$$i^* \sim \frac{\sigma}{S}; \tag{13a}$$

and equation (15) can be written as

$$\dot{N} = \nu_0 B \left(\frac{\sigma}{S}\right)^{2/3} \exp - \left(\Delta G_A + \frac{16\pi}{3}\frac{\sigma^2}{S^2}\right) \Big/ kT . \tag{17}$$

Here σ and S both depend on the strain and thus vary throughout the matrix. Since the surface energy σ may be compared to the mean grain boundary energy in the metal, it varies much more slowly with the strain than does S. (The dependence of σ on strain can be compared to the dependence of σ on the disorientation angle; this angle in its turn can be taken as a measure of the strain: $\varepsilon \approx \tan \theta$.) Therefore, in highly strained regions, σ/S, thus ΔG^* will be small and the nucleation rate increases. This is the quantitative interpretation of the occurrence of *heterogeneous nucleation*: nucleation which takes place preferentially at sites of local strain, and the rate of which increases more or less exponentially with the strain.

It has been mentioned before that the rate of *grain growth* also depends fairly sensitively on the strain. This must be essentially due to a decrease of ΔG_A with increasing strain (the pre-exponential factor also decreases, but this has less effect): grain growth depends only on the activation energy for migration of the boundary, not on the energy difference between the grains. However, ΔG_A varies much less sensitively with strain than the second term in the exponential of (17). To explain that it varies at all, one must assume that the presence of lattice imperfections formed during the pre-strain, accelerates the atomic migration (diffusion, grain boundary migration).

Grain growth proper, as opposite to recrystallization, is governed *only* by the interfacial energy, that is, by the surface tension of the grain boundaries (EWING and ROSENHAIN [1900]). This is confirmed by many independent observations. The driving force is given by

$$\Delta F = \frac{\sigma V}{2r}, \tag{18}$$

where σ is the surface tension, V the atomic volume and r the radius of curvature of the boundary. The rate of migration is proportional to

$$\frac{\Delta F}{kT} e^{-\Delta G_A/kT}.$$

Assuming $r \sim D$, the dimension of the growing grain, one has

$$\frac{dD}{dT} = \frac{\sigma V}{2kT} e^{-\Delta G_A/kT} \cdot \frac{1}{D}, \tag{19}$$

the solution of which is:

$$D = D_0 + \sqrt{\frac{\sigma V}{kT}} e^{-\Delta G_A/kT} \cdot t^{1/2} \tag{20}$$

when the original grain size D_0 can be considered as small compared to D. Experimentally, the exponent of t is usually found to be less than $\frac{1}{2}$ (the observed values range between 0.1 and 0.5) but then, D is usually *not* much larger than D_0, the original grain size in the first stages of grain growth, whereas the ultimate grain size is limited by impurities or inclusions. On approaching this limit, the rate of grain growth certainly slackens.

3. Recrystallization and Recovery

Since recrystallization usually occurs at such high temperatures that only dislocations and dislocation-like structures (subboundaries) have remained in the lattice after the previous deformation, it must be concluded that it are predominantly these imperfections that influence the rate of nucleation (at least in not very pure materials). It must be emphasized here that any structural theory of the process must remain premature as long as the excessive influence of traces of chemical impurities mentioned in section 1 is not explained by it.

CAHN [1950] has proposed a dislocation theory of nucleation which is essentially based on the process of polygonization (see section XI–2). He assumed that in highly strained regions of the lattice local polygonization starts within a time short compared to the observed nucleation time. The polygonized region extends gradually by *dislocation climb*, until at last a polygonized, strain-free block of sufficient size has developed, which can grow by ordinary grain boundary motion at the cost of the already present strained grains. The nucleation

period τ would be the "coarsening time" that must elapse before this real grain growth can start. It is the shorter, the higher is the local strain; Cahn takes τ proportional to ϱ, the local radius of elastic curvature.

Against any theory that regards recovery or polygonization as necessarily preceding recrystallization, we adduce the observations that recrystallization often precedes observable recovery, especially in pure materials, and that it has clearly been shown by BURGERS *et al.* [1951], that, once complete recovery has taken place, grain growth has become virtually impossible. A distinction should be made, however, between *local* recovery in heavily strained regions, leading to recrystallization, and *complete* recovery, resulting in the removal of all strained regions together.

Burgers has also drawn attention to the observation that, if two grains of nearly the same orientation meet, the probability that one of them consumes the other by recrystallization is low: small angle boundaries are very stable against grain growth, and have only a slight tendency to migrate. This does not hold for polygonization walls, however, although these form a special case of small-angle boundaries: according to LACOMBE and BEAUJARD [1948] polygonization boundaries (tilt boundaries, and thus glissile) are very mobile.

As alternatives to the recovery theory of recrystallization there are various theories which try to explain the phenomenon in terms of the influence of impurities and self-diffusion. Although to the present author this seems a promising line of attack, these theories are still in a primitive stage and will not be discussed here (LÜCKE and DETERT [1957]).

The apparent similarity of the activation energies for recrystallization and for self-diffusion, as observed in many instances, has led VAN BUEREN and HORNSTRA [1960] to a description of the mechanism of recrystallization in terms of the motion of grain boundaries containing a number of *different sets of dislocations*, as induced by the diffusion of vacancies from sites of high local strain towards these boundaries. At the former sites the vacancy concentration is enhanced over the thermal equilibrium value, at the boundaries it is equal to that value, hence a concentration gradient exists. The preferential absorption of vacancies by one set of dislocations leads to *climb* of these dislocations, and in order to keep the boundary complete, the other sets move too, namely by a combined mechanism of *glide* and

climb. Hence, grain boundary motion is induced by the adsorption of the vacancies. This theory was originally advanced to explain the vanishing of isolated pores during the *sintering process*. Many experiments have made it clear that in this case certainly grain boundaries play an important role, rather than individual dislocations.

4. Recrystallization Textures

Often a heavily cold-worked polycrystalline metal develops a *texture*, that is, the grains show a preferred orientation. This is a consequence of the restricted possibilities for deformation of the individual grains; these must deform without leaving voids at their boundaries. Five independent slip systems are at least needed to account for the most general deformation of each grain (see Chapter XIII) and, since each slip system gives rise to a certain amount of rotation, the rotations of the various slip systems add up to the resulting preferred orientation of the grains. In metals of the same crystal structure textures are thus

TABLE 2

Deformation textures of metal aggregates

Metal structure	*Cold-drawn*	*Cold-rolled*	*Compressed*
face-centred cubic	double fiber: either ⟨111⟩ or ⟨100⟩ parallel to axis of extension.	{110}-plane parallel to rolling plane, ⟨112⟩ parallel to rolling direction. Also {112} ⟨111⟩ combination occurs; then, on the average, {123} ⟨121⟩ combination is found. Other interpretations have been proposed.	single fiber: ⟨110⟩ parallel to axis of compression.
body-centred cubic	single fiber: ⟨110⟩ parallel to axis of extension.	{100} parallel to rolling plane; ⟨110⟩ parallel to rolling direction.	double fiber: ⟨111⟩ and ⟨100⟩ parallel to axis of compression.
hexagonal close-packed	single fiber: ⟨10$\bar{1}$0⟩ or different direction parallel to axis of extension.	{0001} parallel to rolling plane; ⟨1101⟩ parallel to rolling direction.	single fiber: ⟨0001⟩ parallel to axis of compression.

closely related; they differ according to the mode of deformation and further particulars. Some observed deformation textures are given in Table 2. This table represents only a very crude classification; many exceptions and deviations occur from the "ideal" textures cited.

On annealing a cold-worked metal so that recrystallization occurs, a new texture is often observed, which is called the *recrystallization texture*. Generally this texture is most pronounced when a deformation texture was present before; it need not be simply related to the latter, however. In cold-drawn aluminium, the ⟨111⟩-fiber texture is unchanged on recrystallization below 500 °C; it is replaced by a ⟨112⟩ or ⟨210⟩-fiber texture when recrystallized above 600 °C, however, or when secondary recrystallization has taken place above this temperature. In copper, a similar change from the rolling texture to a different recrystallization texture (⟨100⟩ [001], cube texture) occurs above 350 °C.

Silicon-iron, a very important material for technical applications (as transformer cores) shows a recrystallization texture which differs much according to the treatment received. A {110} ⟨001⟩ ("cube on edge") texture of recrystallized rolled sheet is obtained when a certain amount of impurity is present (e.g. nitrogen); it is absent in very pure material (FAST [1956]). Recently, a cube texture {100}⟨001⟩ has been obtained in very thin sheet (ASSMUS *et al.* [1957]). We refer for more particulars about this very intensively studied material to the numerous publications of Turnbull, Fisher, Dunn and coworkers.

A simple scheme like that in Table 2 cannot be given for recrystallization textures. The kind of metal, temperature of recrystallization, initial deformation texture, impurity content, all influence the texture in a not very systematic manner. A relatively simple example is provided by the b.c.c. metal tungsten. The effect of certain impurities bordering the initial fiber elements in cold-drawn tungsten wires is particularly outspoken. There, crystal growth can be impeded in all directions, except in the length direction of the wire, and this results in a remarkable recrystallization texture (⟨531⟩, RIECK [1958]) which is specific for "doped" tungsten wires. Also the form of the final crystals is typical ("swallow tails"), and depends strongly on minute amounts of impurities.

An explanation for *this* peculiar recrystallization texture is the following. Before the final stages of drawing the material has been annealed and a fiber deformation structure consisting of small axially

elongated grains of all possible orientations has developed. These are surrounded by thin walls of impurity, according to an idea put forward by MEYERING [1955]. During the final stages of deformation the grains are strongly cold-worked, but owing to the presence of a large number of possible glide systems in each (b.c.c.) grain, the deformation can in general easily be followed by intragranular slip and the impurity walls remain intact.

However, some grains are exceptionally unfavourably oriented with respect to the active stresses, the resolved shear stress being of appreciable magnitude on one or two slip systems only. Therefore, these grains must deform *also* by grain boundary sliding and the impurity walls around *them* are *broken up* during the deformation. During the subsequent recrystallization treatment, these grains are now in a favourable position for growth, and it is their orientation that dominates in the final aggregate. Analysis shows that it is precisely when the fiber axis lies between ⟨421⟩ and ⟨531⟩ that such crystals are found.

Recrystallization textures are usually more pronounced than deformation textures. This is connected with the mechanism of their formation; whereas deformation textures are the result of simultaneous slip on many slip systems, recrystallization textures arise either since the *nucleation* of grains with certain spatial orientations is preferred (W. G. BURGERS [1953]), or because the *growth* of specifically (relatively) oriented grains is favoured (BECK [1953]). The second "oriented-growth" theory seems at the moment to be in accord with most, but not all observations. The reason for the orientation dependence of grain growth is probably to be sought in an orientation dependence of the quantity ΔG_A of section 2, the free energy of activation for boundary migration, or, what is the same, for the transport of atoms across the boundary. The transfer of atoms from the matrix to the growing grain seems easiest when the new grain deviates in orientation from its neighbours by a not too small amount, whereas grains differing but slightly in orientation from the matrix do not grow (see previous section). Directly connected with this is the often noted occurrence of inclusions of nearly the same orientation or of twin orientation in a completely recrystallized material; these inclusions represent unrecrystallized grains of the initial material that have not been consumed by a larger growing grain, because the latter possessed nearly the same or twin orientation.

Those orientation differences are especially suited for grain growth,

where *cooperative* atom motions can be invoked to obtain the necessary transfer. Such motions can be described by the movement of one or more dislocations (of the partial type, of course).

Recently, BURGERS and VERBRAAK [1957] proposed a theory to explain the formation of a cube recrystallization texture from the deformation texture of cold-rolled copper sheet in terms of the inverse Rowland transformation. The Rowland transformation consists in the

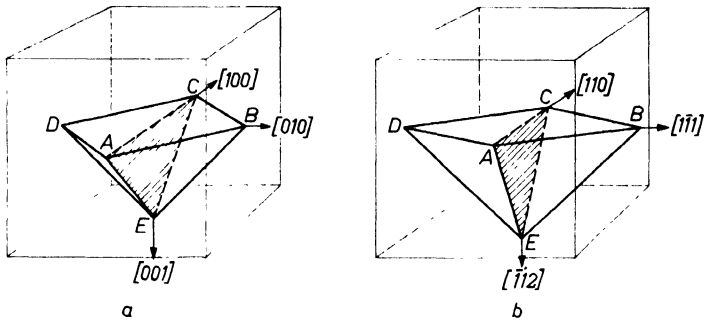


Fig. 2. Rowland transformation of cubic lattice. *a*) By compressing the half-octaeder ABCDE along AC, leaving the hatched plane ACE fixed, two tetraeders belonging to two cubic lattices standing mutually in a twin relation (but not to the original cube) are formed. *b*) The ⟨100⟩-directions of the original lattice have been transformed as indicated into [110], [1$\bar{1}$1] and [$\bar{1}$12]-directions in the partial lattices. The twin plane is the octahedral (1$\bar{1}$1)-plane. (After BURGERS and VERBRAAK [1957].)

compression of a cube along one of its cube axes, leaving one of the cube planes containing this axis fixed. The other cube axes both extend in length and the original lattice splits up into two other cubic lattices that stand in a twin position relative to each other (Fig. 2). It can be shown that, if the two rolling textures mentioned in Table 1 occur together, or the second one alone, configurations bearing this twin relation to each other will undoubtedly be formed locally. In terms of the oriented-nucleation theory, such regions are now supposed to act as favoured nucleation regions, since by merely carrying out an inverse Rowland transformation such as can for instance be produced by the successive movement of dislocations on {112}-planes (compare section XIII–3 on twinning)—the single cubic recrystallization orientation is produced.

A particular kind of *twinning* is sometimes induced by heating a deformed metal. The formation of such *annealing twins* can be com-

pared directly to recrystallization. Here, the influence of definite orientation differences is particularly pronounced, since twinning occurs specifically by dislocation motion (see Chapter XIII). We shall return to this and similar lattice transformations in the next chapter.

REFERENCES

ALBERT, P. and O. DIMITROV (1957); C. R. Acad. Paris **245,** 2275.

ASSMUS, F., K. DETERT and G. IBE (1957); Zs. Metallk. **48,** 344.

AVRAMI, M. (1939); J. chem. Physics **7,** 1103; (1940): J. chem. Physics **8,** 212; (1941): J. chem. Physics **9,** 177.

BARRETT, C. S. (1952); The Structure of Metals (New York, McGraw-Hill).

BECK, P. A. and P. R. SPERRY (1950); J. appl. Phys. **21,** 150.

BECK, P. A. (1953); Acta Metall. **1,** 230.

BECK, P. A. (1954); Adv. in Physics **3,** 245.

BLADE, J. C., J. W. H. CLARE and H. J. LAMB (1959); Acta Metall. **7,** 136.

BUEREN, H. G. VAN and J. HORNSTRA (1960); Proc. 4th Int. Symp. Reactivity of Solids (Elsevier, Amsterdam).

BURGERS, W. G. (1941); Handbuch der Metallphysik **3,** Part 2.

BURGERS, W. G. (1953); Acta Metall. **1,** 234.

BURGERS, W. G., Y. H. LIU and T. J. TIEDEMA (1951); Proc. Kon. Ned. Acad. Wet. **54,** 459.

BURGERS, W. G. and C. A. VERBRAAK (1957); Acta Metall. **5,** 765; see also C. A. VERBRAAK, Thesis Delft 1959.

BURKE, J. E. (1949); Trans. AIME **180,** 73.

BURKE, J. E. and D. TURNBULL (1952); Progress in Metal Physics **3,** 220.

CAHN, R. W. (1950); Proc. phys. Soc. **A364,** 323.

COULOMB, P. and P. LACOMBE (1959); C. R. Acad. Paris **248,** 964.

EWING, J. A. and W. ROSENHAIN (1900); Phil. Trans. Roy. Soc. **A193,** 353.

FAST, J. D. (1956); Philips Research Reports **11,** 490.

HOLLOMON, J. and D. TURNBULL (1953); Progress in Metal Physics **4,** 337.

JOHNSON, W. A. and R. F. MEHL (1939); Trans. AIME **135,** 416.

LACOMBE, P. and G. BEAUJARD (1948); J. Inst. Metals **74,** 1.

LÜCKE, K. and K. DETERT (1957); Acta Metall. **5,** 628.

MASING, G. (1958); Zs. Metallk. **47,** 64.

MEYERING, J. L. (1955); 2. Plansee Seminar Reutte, p. 305.

MOTT, N. F. (1948); Proc. phys. Soc. **60,** 391.

RIECK, G. D. (1958); Acta Metall. **6,** 360.

TURNBULL, D. and J. C. FISHER (1949); J. chem. Phys. **17,** 71.

WALTER, J. L. and C. G. DUNN (1959); J. Metals Trans. Metall. Soc. AIME **215,** 465.

CHAPTER XXI

PHASE TRANSFORMATION AND PRECIPITATION

1. Transformation

Recrystallization is a special kind of phase transformation; it leaves the composition and structure unaltered, but changes the orientation. In general phase transformations include changes in structure and composition as well, but they are still characterized by the processes nucleation and growth. Another transformation which changes orientation only, but in a well-defined way, is the formation of *annealing twins*. A phase change that leaves the composition unaltered but changes the structure is the *martensitic transformation* (an "allotropic" transformation). We shall discuss these in section 2.

The most common class of transformation consists in the formation of another phase and is controlled by a *diffusion* process. Processes like pre-precipitation, formation of Guinier-Preston zones, precipitation of a new phase, changes of order, etc., can be distinguished. These will form the subject of a brief discussion in the following sections. A systematic description of the vast field of transformation processes will not be attempted, only the role played by imperfections will be indicated.

2. Formation of Annealing Twins and Martensitic Transformations

Microscopic investigation of face-centred cubic polycrystalline metals, and of some body-centred cubic metals or alloys as well, as shown recently by Cahn and by Leshi (unpublished), reveals the existence of so-called *annealing twins* after an annealing treatment has been carried out. These twins occur especially often in the later stages of recrystallization. (Occasionally they are already present in

a crystal as grown from the melt, in which case one speaks of *growth twins*.) Twinning planes in f.c.c. metals are the octahedral {111}-planes, and the geometry of this type of twinning is depicted in Fig. 1. It is characterized by the absence of any change in the relative position and number of nearest neighbours; only next-nearest and further

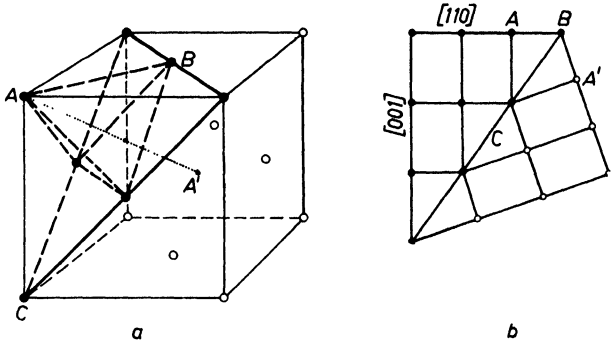


Fig. 1. Geometry of twinning of face-centred cubic metal with (111)-plane as habit plane (after MASING [1952]).

neighbours across the twin boundary undergo a relatively slight position change. The stacking order of cubic octahedral planes which reads .....*abcabc*..... in a perfect crystal (cf. Part I) has now been changed to .....*abcacbac*..... The boundaries of annealing twins are mostly very straight and stable. The exact mechanism of annealing twin formation is as yet unknown, it is closely connected to the formation of stacking faults during the migration of grain boundaries. A (111)-boundary in an f.c.c. crystal migrates by the sequential deposition of new (111)-layers, and the accidental deposit of a wrongly stacked layer initiates a twin. The occurrence of such an accident can for instance be promoted by the presence of extended dislocations. For a detailed discussion we refer to a series of papers by FULLMAN [1950, 1951].

Mechanical twinning (see Chapter XIII) occurs seldom in face-centred cubic metals, presumably because glide is there very easy. It is, instead, very common in body-centred cubic and hexagonal metals. The twinning planes are usually {112} like those of annealing twins in b.c.c. metals and $\{10\bar{1}2\}$ in h.c.p. metals. We refer the reader for more particulars to the chapter mentioned.

Of great importance in metal physics are the *martensitic* transformations. These have been called after the prototype of the body-

centred tetragonal (or cubic) phase *martensite*, an Fe-C alloy occurring at low temperatures. Martensitic transformations, together with twinning and other allotropic transformations are often called *diffusionless* phase changes, as they are apparently not thermally activated. Ordinary grain growth or recrystallization is not a diffusionless transformation, of course.

Martensite proper is formed as a metastable product of the γ-α-transition of iron containing some carbon. As is well-known, austenite or γ-iron is stable at high temperatures, and α-iron or ferrite at low temperatures (the temperature of the transition $\gamma \rightarrow \alpha$ depends on the carbon content). When the material is rapidly cooled through this transition, so that no appreciable diffusion can take place, the austenite is first retained in a metastable form. However, below a certain temperature M_s, a new phase martensite develops spontaneously (instead of the stable ferrite), apparently in the form of needles, but actually as platelets. M_s depends sensitively on the carbon content; for 1% C it is about 200 °C, for 1.8% C about 100 °C. The martensitic temperature is altered by pre-straining. On reheating the material, the structure ultimately transforms back to the stable α-structure and cementite at a temperature $> M_s$, after a series of metastable states has first been passed.

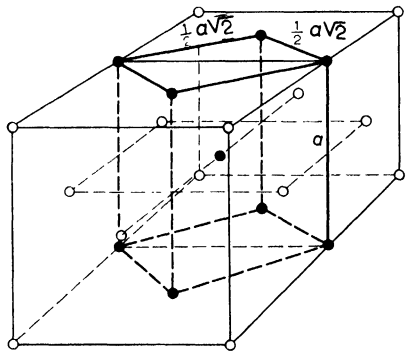


Fig. 2. Body-centred tetragonal lattice with axial ratio $\sqrt{2} : 1$, regarded as equivalent to a face-centred (cubic) lattice. A tetragonal lattice with a different ratio can be obtained from this by uniaxial compression or extension along [001].

The tetragonal Fe-C martensite lattice can be considered formally as the product of a uniaxial compression along a ⟨100⟩-direction of a face-centred cubic lattice (Fig. 2). However, it follows from X-ray

observations that the actual martensitic transformation is not equivalent to such a simple uniaxial compression. Face-centred cubic Fe-Ni alloys up to 40% Ni also show a martensitic transformation after rapid cooling; here the product is body-centred cubic. Many other metals and alloys like lithium, cobalt, gold-cadmium, indium-thallium show analogous diffusionless transitions resulting in allotropic transformations of the type described. Table 1 (partly after Barrett) gives a selection of the best known cases.

TABLE 1

Diffusionless transformations

Alloy		*Transformation on cooling*	*Orientation relation*	*Habit plane (parent phase)*	M_s
α-Fe-C	0 –0.4% C	f.c.c. → b.c.tetr.	$\{110\}\,//\,\{111\}\gamma$, $\langle 1\bar{1}1\rangle\,//\,\langle 1\bar{1}0\rangle\gamma$ (K.S.)	{111}	> 300 °C
id.	0.5–1.4% C	id.		{225}	300–120 °C
	1.5–1.8% C	id.	?	{259}	120–80 °C
Fe–Ni	≈ 30% Ni	f.c.c. → b.c.c.	$\{110\}\,//\,\{111\}\gamma$, $\langle 1\bar{1}0\rangle\,//\,\langle\bar{2}11\rangle\gamma$ (N)	≈ {259}	—30 °C
Fe–Mn	1–15% Mn	f.c.c. → b.c.c.	?	?	
	15–25% Mn	f.c.c. → c.p.h.	?	?	
Au–Cd	47.5% Cd	b.c.c. → orthorhombic	$\{001\}\,//\,\{011\}\beta$, $\langle\bar{1}\bar{1}0\rangle\,//\,\langle 1\bar{1}1\rangle\beta$	{133}	60 °C
Cu–Al	11–13% Al	b.c.c. → c.p.h. (distorted)	$\langle 11\bar{2}0\rangle\,//\,\langle 1\bar{1}1\rangle\alpha$	≈ {133}	
	13–15% Al	b.c.c. → c.p.h.		≈ {122}	
Cu–Zn	40% Zn	b.c.c. → f.c.c.	$\{110\}\,//\,\{111\}\alpha$, $\langle 111\rangle\,//\,\langle 110\rangle\alpha$ K.S.	≈ {155}	
Li		b.c.c. → c.p.h.	$\{0001\}\,//\,\{110\}\alpha$	{144}	—195 °C
Na		b.c.c. → c.p.h.	$\{0001\}\,//\,\{110\}\alpha$		—236 °C
Zr		b.c.c. → c.p.h.	$\{0001\}\,//\,\{110\}\beta$	?	865 °C
Co		f.c.c. → c.p.h.	$\{0001\}\,//\,\{111\}\alpha$, $\langle 11\bar{2}0\rangle\,//\,\langle 1\bar{1}0\rangle\gamma$	{111}?	470 °C
In–Tl	20.75% Tl	f.c.c. → f.c.tetr.		{110}	
Ti–(Mo, etc.)	4–20% Mo	b.c.c. → c.p.h.	$\{0001\}\,//\,\{110\}\beta$	{334}	400–800 °C

The orientation relations entered in this table are often very uncertain. The relation mentioned for iron-carbon alloys is the so-called Kurdjumow-Sachs relation; that belonging to iron-nickel alloys the so-called Nishiyama relation. Both have only approximate validity, they have been proposed because the corresponding transformations can then be considered as a sequence of simple shears. For a detailed

description we refer to papers by COHEN [1951], to Barrett's book and to a Symposium Report on Phase Transformations in Metals [1956].

Sodium, lithium, zirconium and titanium form a distinct group, which transforms by the so-called lithium transformation from a body-centred cubic to a close-packed hexagonal structure. The occurrence of such a transformation in the alkali metals at very low temperatures had been predicted by Zener and was first observed by BARRETT and TRAUTZ [1948] (in lithium).

Most of the martensitic transformations can be regarded crystallographically as the superposition of a shear and an extension. In the Kurdjumow-Sachs relation the shear amounts to $\frac{1}{3}$ along [110] (in the austenite lattice), the extension to 6% in the corresponding [112]-direction. Both processes are presumably effected by the motion of dislocations. A number of mechanisms have been proposed by which this could take place in the various cases, and it would lead us to far to discuss any of them here in detail. In only one, especially simple case a definite description in terms of dislocations seems to have been established. This concerns the tranformation in cobalt (SEEGER [1953]). The transformation (from the f.c.c. lattice into the c.p.h. lattice) can be brought about by the motion of a number of Shockley half-dislocations of the type dicussed in Chapter III. When such half-dislocations pass over every second parallel (111)-plane of the cubic lattice the cubic layer stacking ...*abcabc*... reverts to the hexagonal stacking ...*ababab*.... Far above the transformation temperature, the dislocations present in the cubic cobalt lattice are but little extended, corresponding to a large difference in free energy between the hexagonal low-temperature phase and the cubic phase and thus to a high energy of a stacking fault. On lowering the temperature, this free energy difference and the stacking-fault energy decrease, and the dislocations along ⟨110⟩, which were originally nearly perfect, split up into wider and wider extended dislocations. Once the transformation temperature is reached, the repulsive force between the two half-dislocations of each pair drives them so far apart, that a large fraction of each glide plane is transformation into a stacking fault.

To explain the spontaneous transformation on *alternating* lattice planes of a large volume of the crystal, as is observed at the transformation temperature, Seeger assumed that migrating half-dislocations pass regularly from one (111)-plane to another parallel to it but two interplanar distances away, owing to the effect of *intersecting* dislo-

cations. These might lie along the [111]-direction perpendicular to the (111)-planes under consideration (but this is not a necessary condition). The intersecting dislocation causes jogs in the original dislocations. Near these jogs the partial splitting has been removed and the situation as

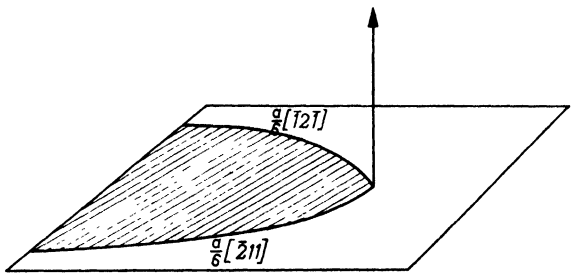


Fig. 3. Intersection between two dislocations in the cubic phase of cobalt. Only the situation in one single (111)-plane is shown. The dislocation lying in this plane is split up into two Shockley half-dislocations with Burgers vectors $\frac{1}{6}a\ \langle 112\rangle$, that have come together at the point of intersection. When the two partials rotate around the jog in opposite directions, they do in general not meet again after a complete revolution. Thus the stacking fault is transformed in a helicoidal surface.

depicted in Fig. 3 occurs. If now the intersecting perpendicular dislocation, whether also extended or not, has a screw component $\frac{2}{3}a[111]$, it transforms the parallel (111)-layer stacking into a continuous spiral surface with pitch $\frac{2}{3}a\sqrt{3}$ along [111], thus equal to two interplanar spacings. Expansion of the stacking fault in the (111)-plane means that one of the half-dislocations sweeps the spiral surface, say, upwards, whereas the other half-dislocation migrates downwards along that surface. This rotation is favoured by the mutual repulsion, and hence the whole spiral surface is transformed successively into one huge spiral stacking fault. This is precisely the cobalt transformation as it is observed to occur over a large number of lattice planes. The transformation is activated by the aid of one "pole" dislocation only, which must possess the required screw component along [111]. The simplest pole dislocation imaginable has the Burgers vector $\frac{1}{2}a[211]$, it is a combination of a Shockley partial and a Frank sessile dislocation.

The mechanism described here much resembles that of *deformation twinning* by a dislocation mill mechanism as discussed in Chapter XIII for the body-centred cubic lattice. From the observed very high speed of the cobalt transformation (it is often accompanied by an audible click) also in this case a lower limit to the velocity of the rotating dislo-

cations in the lattice can be inferred; this limit has always been found to be smaller than the velocity of sound by several orders of magnitude.

A corresponding detailed dislocation mechanism for the martensitic transformation in iron has not yet been devised. Various ingenious mechanisms have been proposed (FRANK [1953]), but the verification of these is rather difficult, since the crystallography of the transformation is intricate and often known only approximately (BOWLES and MACKENZIE [1953, 1954]). To take up the shear in the Kurdjumow-Sachs relation, for instance each sixth plane parallel to the observed habit plane (225) must glide over a certain distance along a [110]-direction; to take up the volume change and ensure coherence, on the surface of each martensite plate a certain distribution of dislocations must be present. How these are formed is still unknown. We approach here the problem of martensitic *nucleation*. All diffusionless transformations start from predisposed nuclei, as twinning and other transformations and phase changes do. Presumably in martensitic transformations heterogeneous nucleation at dislocationlike structures is of great importance (VERBRAAK [1959]). Since the transformation can be induced by previous cold-work, the nuclei presumably consist of dislocation arrays with a geometry suitable for the generation of dislocations for the transformation process. Cold-work further promotes the transformation, even above the M_s-temperature (but not too far above), by helping the necessary shears to develop. Below the transformation temperature cold-work sometimes introduces new metastable phases in several systems. For a description of possible nucleation mechanisms we refer to COHEN [1958].

A discussion of the *thermodynamics* of martensitic transformations has been given recently by KAUFMAN and COHEN [1959].

3. Diffusion-controlled Transformations

3.1. GENERAL

In diffusionless transformations lattice imperfections play an essential role in the actual crystallographic mechanisms of nucleation and of growth, and especially dislocations are of great importance—the situation is in some aspects even comparable to plastic glide, which could also be considered as a transformation. In the much wider category of diffusion-controlled transformations lattice imperfections play only a "secondary" role in the growth process, in that they can

promote the necessary *diffusion*. Therefore point defects have now the dominating influence. Of course, an influence of dislocations on the *nucleation* process remains present. In diffusion-controlled transformations the latter influence is due primarily to the overall stress release that can be obtained near dislocations (and other defects) when a new phase is nucleated there; in diffusionless transformations the preference for nucleation near disordered regions is primarily associated with the presence of peculiar shears etc., that may help the transformation on its way (preferred nucleation theory).

Several categories of diffusion-controlled transformations can be distinguished, such as changes of chemical phase, precipitation phenomena (often resulting in hardening, see Chapter X), and order-disorder transitions. We shall discuss only a few apparently unrelated aspects of the extended subject, which are connected with lattice imperfections. A beautiful compilation of the whole field is available in the symposium report Phase Transformations in Solids [1951]; apart from this an extended review of more limited scope dealing with precipitation only, has been written by HARDY and HEAL [1954], a shorter paper by GUINIER [1956]. Any attempt to improve on these here would be futile.

3.2. INFLUENCE OF PLASTIC DEFORMATION AND IRRADIATION ON PRECIPITATION AND PHASE TRANSFORMATION

It is well known that many phase transformations, and especially precipitation, are appreciably accelerated by plastic deformation or by irradiation. If, as is often possible, the process can be described as a first-order rate process, it has been observed that usually the activation energy of this remains unchanged, but that the rate of the reaction is multiplied by an approximately temperature-independent factor. Many observations have been made on the *age-hardening* (a process due to precipitation, see Chapter X) of cold-worked alloys; this can be accelerated by factors of the order 10 to 10^7 by plastic deformation or irradiation; again the relative temperature dependence is unaltered. The ageing time for precipitation of Cu_2Al in Cu-Al alloys was found to depend exponentially on the degree of cold-work, for instance. The cause of all this must be sought in the facilitation of *substitutional* diffusion by vacancies, formed during the irradiation or the deformation process. The number of vacancies is considerably

enhanced above the thermal equilibrium value by this treatment (by factors ranging from 10 to 10^8, see Chapters VII and XIV), and the fraction of time spent by a given atom as neighbour to a vacancy is correspondingly greater than in the annealed state. Therefore, the jumping frequency is increased, but the energy needed to make the jump remains the same. It is not to be ruled out a priori that the dislocations also formed during cold-work accelerate the ageing process too. No unambiguous evidence for this exists at the present moment, except in as far as typical *nucleation* processes are concerned (Chapter XX).

In this respect the following observation may be mentioned. DE SORBO and TURNBULL [1956] observed the kinetics of the precipitation of tin in lead-tin solid solutions, present as small spheres. They found the precipitation rate to be practically zero in spheres of diameter less than $40\,\mu$, which contained practically no structural imperfections. Either by cold-working or by artificial coagulation of the globules, the rate of precipitation could be brought back to the normal value pertaining to larger blocks of material.

Irradiation with fast particles, especially with fast neutrons, often accelerates precipitation processes to an even larger extent than deformation. A great many examples are known, such as the acceleration of the ageing of supersaturated Cu-Be alloys by neutron irradiation, which can be explained either in terms of a superconcentration of vacancies or of enhanced nucleation of the precipitate. Sometimes *phase changes* are even observed to occur under irradiation, which normally could not take place at the temperature at which the metal is held. A 9%-uranium-molybdenum alloy irradiated at about 100 °C transforms partly from the rhombic α-ε-phase mixture, stable at room temperature, to the cubic γ-phase, although the latter is only stable under normal circumstances above 600 °C. COTTRELL [1956] explained this not in terms of enhanced diffusion, but in terms of the *stirring up* of the molybdenum lattice in the fission spikes caused by the neutrons. At the low temperature of the irradiation the displaced atoms in the spikes cannot diffuse and after the "stirring" fall back into the position of maximum entropy, that is the regular cubic phase.

A change of the magnetic properties of a 2% Fe-Cu alloy under proton irradiation has been measured. DENNEY [1956] has interpreted it as due to a transformation of the precipitated α-Fe phase into the high-temperature stable γ-phase. This is probably another

example of the stirring effect, which now occurs in displacement spikes of the Brinkman type (see Chapter XIV). OKKERSE and GONSER [1957] found "molten" zones of higher density in a gallium-antimony alloy after irradiation with deuterons at −180 °C (in this material the liquid phase is denser than the solid). They explained this also as an indication of the formation of displacement spikes of the Brinkman type. The material outside the spike exerts a pressure of about 10^5 atmospheres on the matter inside, and this causes a lowering of the melting temperature of the latter by several hundreds of degrees, sufficient to cause melting at room temperature.

3.3. GUINIER-PRESTON ZONES, AGE-HARDENING

The formation of Guinier-Preston zones is of the greatest technical importance; the hardening of aluminium alloys of the duralumin type is based on it. We shall therefore devote a few words to this phenomenon here. The aluminium-rich side of the constitution diagram of Al-Cu is schematically presented in Fig. 4. Below 400 °C

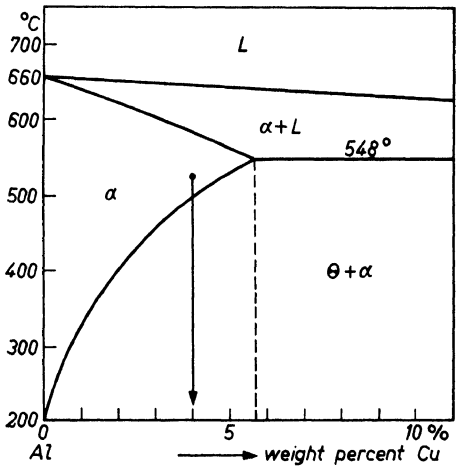


Fig. 4. Aluminium-rich side of the Al-Cu constitution diagram.

hexagonal $CuAl_2$ (called θ) should be in equilibrium with the solid solution α. However, on quenching to below 150 °C, and subsequent heating between 50 and 150 °C, instead of θ *coherent* copper-rich *zones* (*G.P. 1 zones*) form from the supersaturated solid solution, and these result in appreciable hardening. From the typical shape of the Laue

diffraction pattern PRESTON [1938] and GUINIER [1938] deduced independently that these zones are platelets with cube faces as habit planes and one or a few atomic spacings thick. Their lateral dimension is of the order of 100 atomic spacings or less. Accurate observations have shown a distinction to exist between not less than three kinds of platelike precipitates that can form in the quenched alloy (at different ageing temperatures). These phases have been called G.P. 1, G.P. 2 or θ'' and θ' respectively, in the sequence of progressive precipitation. The third phase forms plates of much larger dimensions than the two other kinds, which form only "*zones*". In the aluminium-silver system these additional complexities are absent and only one type of *spherical* zones is formed under comparable circumstances. The copper content of G.P. 1 zones in Cu-Al is still a matter of doubt. Guinier and Preston estimated them to be nearly pure copper, JAGODZINSKI and LAVES [1949] argued (on the basis of the maximum available diffusion current) that in the ageing time under consideration only zones containing at most 10% of Cu can be formed. The deformations of the lattice would then be expected to be rather small. More probable is Guinier's original estimate, when the additional supposition is made that during precipitation the diffusion rate of the copper atoms is considerably enhanced. Whereas the low-temperature extrapolated diffusion data describe satisfactorily the formation rate of the stable θ phase, they are too small by not less than a factor 10^7 or higher when used to explain the rate of formation of G.P. 1. zones. No local depletion of copper in the matrix close to the zones has been observed, thus *locally* enhanced diffusion can not explain the extreme rapidity of zone formation. (On the contrary, in Al-Ag alloys the precipitates corresponding to G.P.-zones are surrounded by regions which are indeed to some extent depleted of silver atoms.) For a further discussion we refer to GUINIER's lecture [1956] and, inter alia, to a series of papers by KÖSTER c.s. [1954].

A review of the possible causes for the enhanced diffusion rate in the formation of G.P.-zones has already been given in Chapter XVIII.

FRANZ and KRÖNER [1955] have pointed out that dislocations might play a role in the transformation of G.P. 1 into G.P. 2. The second type of zonelike precipitate can probably be regarded as a "superstructure" of G.P. 1 zones, that is, several copper-rich zones arranged on top of each other. The formation of this "superstructure" is probably due to elastic interaction. Franz and Kröner stressed the



Fig. 5. Dislocation loops of Frank-Read source in Cu-Al alloy, made visible by precipitation of θ-phase on the dislocations, which run nearly parallel to the surface (WILSDORF [1954]).

similarity between the stress field of a single copper platelet in the aluminium matrix and a circular edge dislocation with a Burgers vector perpendicular to its (100)-plane. When six copper plates are on top of each other, the deviation of the lattice parameter is about equal to one [100]-lattice vector, and the structure could indeed be surrounded by a real edge dislocation loop, to compensate for the long-range stress field. Existing edge-dislocations loops could therefore, in principle, provide especially suitable nuclei for the transformation of G.P. 1 into G.P. 2. Precipitation of θ on dislocations has been used to "decorate" dislocations in Cu-Al alloys. A beautiful example is shown in Fig. 5.

The *coherent* precipitation mentioned here is but a special case of a relation between the structures and orientations of a precipitated second phase and the parent lattice respectively. The existence of a fixed orientation relationship often reveals itself in the formation of a *Widmanstätten* structure, first found in Fe-Ni alloys occurring in meteorites. Here, it was the consequence of a martensitic transformation; however, in age-hardened alloys such relations are also observed. BARRETT [1951] advanced an explanation of this in terms of the tendency for platelike precipitates to surround themselves with coherent interfaces. For a further discussion we refer to the review papers cited.

3.4. ORDER-DISORDER TRANSFORMATIONS

Many alloys can exist in an ordered and a disordered state. The latter may never be completely disordered as far as short-range effects are considered (Guinier has once stated that no alloy should exist whose structure is completely disordered), but generally above a certain transition temperature T_c long-range order is absent, whereas under suitable conditions it may be present below T_c. Well-known intermetallic phases showing ordering are AuCu, $AuCu_3$, β-brass (CuZn), FeCo, Ni_3Mn, Fe_3Al, and many others. As the transformation is never sharp from completely disordered to completely ordered, it is useful to define a long-range order *parameter* S. In a binary alloy AB this can be done as follows:

$$S = \frac{n_A - p_A}{1 - p_A}, \qquad (1)$$

where n_A is the fraction of atoms of species A that occupy sites indeed

destined to be occupied by A in the completely ordered arrangement, and p_A is the atomic concentration of A. When complete long-range order exists, $n_A = 1$ and $S = 1$; when complete disorder is present, $n_A = p_A$ and $S = 0$.

The degree of order is usually observed by noting the presence of *super-structure* lines in X-ray diffraction diagrams. It can be shown that the structure factor for these lines, determining their intensity, is proportional to the order parameter:

$$\text{Intensity} \sim F^2 = S^2 (f_A - f_B)^2 , \tag{2}$$

where f_A and f_B are the atomic scattering factors of the atoms A and B respectively. It is thus a relatively simple matter to measure the degree of long-range order. This is not so for short-range order, however.

The order parameter S in an equilibrium alloy decreases with increasing temperature either gradually (second order) or in a discontinuous manner (first order). In the alloys Cu_3Au, Ni_3Fe, a few tens of per cent of order have disappeared already 50 degrees below the critical temperature. This points to a co-operative phenomenon being present: the tendency to disorder increases with increasing disorder until it leads to a catastrophic change at a certain temperature. However, in CuZn, Fe_3Al the order vanishes quite suddenly at T_c (there is a slight decrease at lower temperatures).

The ordering is accompanied by a considerable uptake of energy, and the state of order therefore affects the specific heat of the alloy. Moreover, this state influences sensitively the electric, magnetic and mechanical properties, which in general do not vary linearly with S, however. This dependence is very useful, because it makes possible the study of the ordering *kinetics* with great accuracy.

The rate of transformation on cooling differs greatly from case to case: β-brass (CuZn) orders extremely rapidly, in fact unobservably rapidly, when cooled down through the critical temperature; the alloy Ni_3Mn, on the contrary, requires many days to establish the equilibrium amount of order just below the critical temperature.

To describe the ordering kinetics, one introduces the half-value time τ, the time in which the first half of the change of the observed physical properties occurs. τ is generally found to depend on the temperature through an activation energy:

$$\tau = \tau_0 \exp W/kT . \tag{3}$$

This activation energy sometimes compares very closely with the activation energy for diffusion of one of the components, indicating that a simple diffusion process underlies the change of state. This is for instance the case in CuAu. In Cu_3Au, however, it is found from resistivity observations that the ordering occurs in two steps, which are apparently not related to simple diffusion processes. Owing to the strong temperature dependence of τ, it is usually possible to obtain a disordered alloy below the critical temperature by sufficiently rapidly quenching it from above that temperature.

The ordering kinetics, as well as the state of order itself, can be sensitively influenced by quenching, plastic deformation and irradiation. This is, of course, again connected with the formation of defects that enhance the transformation rate and which themselves already produce some disorder. We have already encountered some examples of this (Chapter XV); another is Cu_3Au ($T_c = 380$ °C) that after quenching from 600 °C to 150 °C shows an accelerated ordering rate, which is due to the effect of frozen-in vacancies. The same was found in γ-irradiated Cu_3Au samples (DUGDALE [1954]).

On the other hand, in ordered alloys the kinetic properties of dislocations are sometimes markedly different from those in random alloys or pure metals. This is due to the effect which a moving imperfection has on the local state of order (see e.g. Chapter X about the Fisher mechanism of alloy-hardening).

The migration of vacancies can increase the degree of order of a disordered alloy below the critical temperature. Therefore, in alloys that show long-range ordering, the effects of point defects and dislocations on properties such as the electrical resistance may differ markedly from those in pure metals. Although their influence on the resistivity by electron scattering may be comparable in alloys and pure metals, the effect on the state of order may completely mask the scattering distribution. Thus has it been found, for example, that plastic deformation of a 75–25 Ag-Pd alloy *decreases* the electrical resistance at −180 °C (AARTS and HOUSTON-MCMILLAN [1956]). The annealing of a quenched or irradiated alloy can result in very large resistance changes, because together with the disappearance of vacancies, order is partially restored. Irradiation below the transition temperature may as well increase or decrease the electrical resistance, depending on the original state of order of the alloy.

As the resistivity of an ordered alloy is usually very sensitive to

the state of order, and the latter is disturbed by the introduction of point defects, ordered alloys are exceptionally sensitive media to detect the presence of point defects in them, such as are produced e.g. by irradiation. Many experiments are therefore being performed in this field. We refer to a paper by GLEN [1955] for a discussion of these.

SIEGEL [1949] has found ordered Cu_3Au to become extremely rapidly *disordered* under low-temperature pile-neutron irradiation. In fact, the rate of disordering becomes so great that it could not be explained by taking into account the disorder created by the atomic displacements only. Presumably, temperature spikes and possibly displacement spikes also, contribute considerably to the degree of disorder. These observations have occasionally been considered as indications for the validity of Brinkman's theory of displacement spikes.

By comparing resistivity changes in ordered and in disordered alloys of the same composition, and irradiated by the same flux, an estimate can be made of the degree of disorder produced in the former case. SEITZ and KOEHLER [1956] concluded in this way from observations by BRINKMAN *et al.* [1954] that the number of atoms disordered per single displaced atom is not of order unity as one would expect, but slightly greater than 100 (9 MeV proton bombardment). On the basis of the theory given by Seitz of the enhancement of diffusion in and near thermal spikes (section XIV–2) one could only explain a multiplication factor of about 10 (40 at the utmost). A way out of this difficulty is shown by the concept of *replacement* collisions (section XIV–2.2.5), which might take place when the energy of the colliding atom has become lower than the displacement energy but is still sufficiently great to push an atom out of its place, and occupy that place itself. Such a "replacement" would have to travel 10 atomic distances to explain the observed discrepancy in disordering rates. This seems a reasonable number indeed. In another case, that of Ni_3Mn studied by ARONIN [1954], the required path length of the "replacing" atom is of the order of a few atomic distances only. Seitz proposed the "plasticity spike" as another possible explanation; in such a spike dislocations are present which cause disordering on a large scale. Of course, it is not to be excluded that the resistivity contributed to by a point defect in a partially ordered alloy differs considerably from that in the pure metal, even in the apparently simple case of Cu_3Au. This would then influence the number quoted above.

For a further discussion of irradiation effects in alloys we refer to the SEITZ and KOEHLER [1956] review paper. The *mechanical* influences due to imperfections in alloys have already been described in Chapters IX and X. A theoretical discussion of the geometry of (imperfect) dislocations in alloys can be found in FRIEDEL's book [1956]. Short-range order effects have been completely left out of discussion (see however the work of Damask quoted in Chapter XVIII), since they are too complicated to be dealt with summarily.

REFERENCES

AARTS, W. H. and A. S. HOUSTON-MCMILLAN (1956); South African J. Science **53**, 88.
ALBERT, P. et O. DIMITROV (1957); C. R. Acad. Sci. Paris **275**, 681.
ARONIN, L. R. (1954); J. appl. Phys. **25**, 344.
BARRETT, C. S. (1951); Phase Transformations in Solids, p. 343 (Wiley, New York).
BARRETT, C. S. (1953); The Structure of Metals (McGraw-Hill, New York).
BARRETT, C. S. (1958); Acta Metall. **6**, 156.
BARRETT, C. S., H. F. KAISER and R. F. MEHL (1935); Transact. AIME **117**, 39.
BARRETT, C. S. and O. R. TRAUTZ (1948); Trans. AIME **175**, 579.
BLADE, J. C., J. W. H. CLARE and H. J. LAMB (1959); Acta Metall. **7**, 136.
BOWLES, J. S. and J. K. MACKENZIE (1953); Acta Metall. **2**, 129; (1954) Acta Metall. **3**, 224.
BRINKMAN, J. A., C. S. DIXON and C. J. MEECHAN (1954); Acta Metall. **2**, 38.
COHEN, M. (1951); Phase Transformations in Solids, p. 588 (Wiley, New York).
COHEN, M. (1958); Trans. Metall. Soc. AIME **212**, 171.
COTTRELL, A. H. (1956); Brit. J. appl. Physics Supplement, see also De Ingenieur (1957), nr. 49 (in English).
CHAUDRON, G. (1954); Nature **197**, 923.
DENNEY, J. M. (1956); Acta Metall. **4**, 586.
DUGDALE, K. (1954); Report Conf. Defects in Solids Bristol (London, The Physical Society).
FRANK, F. C. (1953); Acta Metall. **1**, 15.
FRANZ, E. and E. KRÖNER (1955); Zs. Metallk. **46**, 639.
FRIEDEL, J. (1956); Les Dislocations (Gauthier-Villars, Paris).
FULLMAN, R. (1950); J. appl. Physics **21**, 1069; (1951) J. appl. Physics **22**, 448 and 456.
GLEN, W. (1955); Advances in Physics **4**, 381.
GUINIER, A. (1938); Nature **142**, 569; C. R. Acad. Paris **206**, 1641.
GUINIER, A. (1956); Trans. AIME **206**, 621.
HARDY, H. K. and T. J. HEAL (1954); Progress in Metal Physics **4**, 143.
Institute of Metals (1955) Symposium on Mechanism of Phase Transformations in Metals (London, The Institute of Metals).
JAGODZINSKI, H. and F. LAVES (1949); Zs. Metallk. **40**, 296.
KAUFMAN, L. and M. COHEN (1959); Progress in Metal Physics **7**, 165.
KÖSTER, W. and others (1954); Series of papers in Zs. Metallk.
MASING, G. (1952); Lehrbuch der Metallphysik (Springer, Berlin).
OKKERSE B. and U. GONSER (1957); Phys. Rev. **105**, 756; (1958) J. Phys. Chem. Solids **7**, 55.

Phase Transformations in Metals (1956); (The Institute of Metals, London).
Phase Transformations in Solids (1951); (Wiley, New York).
PRESTON, J. (1938); Proc. Roy. Soc. London **A167,** 526; Phil. Mag. **26,** 855; Nature **142,** 570.
SEEGER, A. (1953); Zs. Metallk. **44,** 247.
SEITZ, F. and J. S. KOEHLER (1956); Solid State Physics **2,** 307.
SIEGEL, S. (1949); Phys. Rev. **75,** 1823.
SIEGEL, S. (1951); Phase Transformations in Solids, p. 366 (Wiley, New York).
SORBO, W. DE and D. TURNBULL (1956); Acta Metall. **4,** 495.
VERBRAAK, C. A. (1959); Thesis Delft Techn. Univ.
WILSDORF, H. (1954); Zs. Metallk. **45,** 14.

CHAPTER XXII

FATIGUE AND FRACTURE

Most of the afore-mentioned influences of lattice imperfections left the overall structure of the material intact. We shall now briefly turn our attention to more disastrous effects that are also caused by imperfections, viz. fracture (brittle or ductile) and failure by fatigue. The theory of these phenomena in terms of lattice imperfections is still in its embryonic stage and the tremendous amount of available experimental data are largely of a technical nature. Therefore, an extensive treatment in this book dealing with the fundamental properties of lattice defects will not be attempted.

1. Fatigue

Fatigue is the name given to the failure of metal structures by fracture when these are loaded cyclically during a certain (large) number of periods and at a stress amplitude which is less than the yield stress or flow stress of the metal in static loading. Rotating and vibrating machine parts are especially liable to fatigue, and this is therefore one of the most frequently encountered, though least understood, causes of accidental failure of aircraft and motor vehicles.

The strength of a metal with regard to fatigue is commonly displayed by the socalled SN curve. There, the number of stress cycles N endured before failure is plotted against the stress amplitude S, as schematically shown in Fig. 1. Such a curve must be regarded as an average of a large number of tests, and individual results scatter widely. Instead of the smooth variation displayed by the drawn curve in Fig. 1, a more discontinuous SN curve is sometimes found (for instance in many types of steel, see dotted curve). Then, a *fatigue or*

endurance limit can be defined, below which an indefinite number of cycles can apparently be endured without cracking or fracture.

During a fatigue test (carried out above the fatigue limit, if present) a certain amount of plastic deformation takes place, notwithstanding

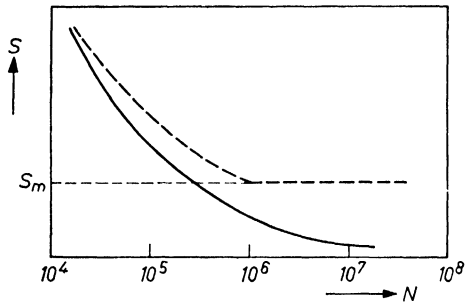


Fig. 1. Schematical *SN* curves for fatigue. The curves show the relation between the maximum stress amplitude *S* which can be endured and the number of cycles (endurance) *N*. The dotted curve applies to a commercial steel showing an endurance limit S_m.

the low stress amplitude. This can be judged from the observed formation of sliplines and other surface features like the remarkable *extrusion* phenomena (see below). Even below the fatigue limit occasional sliplines are generated. The slipline pattern differs markedly from that obtained in unidirectional slip, the lines are more widely spaced and coarser. If the stress amplitude is comparable to the initial flow stress of the material, the material quickly hardens during the first 1000 or 10 000 cycles of the fatigue test. After this, the stress-strain curve no longer changes much, but the slipline pattern becomes increasingly dense, and slipbands increase continuously in width, until finally a crack develops from one or a few of them. The plastic flow is apparently very localized, and it is difficult to follow the progress of fatigue damage by measuring bulk properties such as electrical resistivity. A promising experimental method of obtaining insight into the *internal* processes that take place, is the measurement of the dislocation-induced internal friction of the specimen (which is actually a somewhat related phenomenon).

The end stage of the fatigue test consists in the propagation of the fatigue crack through the specimen. Generally the direction of the crack is along the slip planes on which it has been formed. These are the same on which slip has started in the initial period, and on which

the resolved shear stress is largest. The propagation stage often requires a very large proportion of the fatigue test; in some cases the beginning of crack formation (incipient cracks) has already been observed to take place during the first 10 or 20% of the total time to rupture. At the end of a test hundreds of cracks may be seen, spread through the specimen.

For a thorough discussion of the experimental features of fatigue, we refer inter alia to a recent paper by THOMPSON and WADSWORTH [1958]. A number of hypotheses exist as to the *origin* of fatigue cracks (we shall postpone the discussion of crack *propagation* to the next section). One theory tries to explain the observed phenomena in terms of the stress fields connected with piled-up groups of dislocations, and resembles in outline the fracture theory of STROH [1954, 1955, 1958], which we shall discuss in the next section. MOTT [1955] has criticized the application of the pile-up concept to the fatigue problem. He argued that such pile-ups can only be formed at the *ends* of slip lines, where the dislocations are obstructed in their motion by grain boundaries. It is observed, however, that in fatigue a crack forms by the gradual broadening of a coarse slip line. It can be imagined that the dislocations in such a slip line move continuously backwards and forwards, driven by the alternating stress. In doing this, they will produce vacancies and interstitials in gradually increasing numbers, in the same manner as during ordinary plastic deformation. A dislocation containing j jogs per unit length creates of the order of Lj/b defects per unit length in moving the distance L. Taking $j = 10^4$/cm, in general a dislocation must only carry out about 10^4 cycles before all atoms on its slip plane have been pushed out of place at least once. Of course, a large amount of redistribution will occur, but a high degree of disorder may yet be expected to be present near the slip plane after a number of cycles of the order mentioned. Such a region will then necessarily expand to diminish the strains set up in the surrounding lattice. Mott proposed that, when the disordered region has become wide enough, spontaneous recrystallization will occur (even at low temperatures). This will be accompanied by the generation of strong tensile stresses perpendicular to the slip band which may in turn initiate a crack. It is also possible that, of the point defects formed, the interstitials can diffuse away (by reason of their smaller migration energy, see Chapter XV) and only the vacancies are left, which then would coalesce to a void, with practically the same result.

FUJITA [1954] reconciled to a certain extent the ideas of pile-up stresses and defect accumulation, by pointing out that the short-range interaction between two pile-ups (or groups of dislocations) on neighbouring planes may give rise to a micro crack a few Å wide (Fig. 2).

An interesting aspect of the problem is that some fatigue cracks seem to form preferentially on the surface or at grain boundaries,

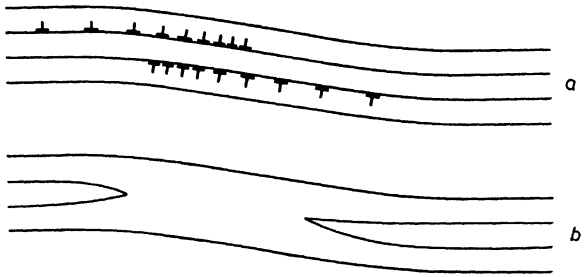


Fig. 2. Schematic model for the formation of a crack out of two groups of dislocations of opposite sign on neighbouring glide planes.

especially so at high temperature and high stress. Here, it should be especially easy for the stresses to relax and the vacancies to diffuse, and consequently surface cracks would be expected to *close up* as soon as they are formed. Mott tried to explain the occurrence of persistent surface cracks as due to the influence of solute gas, by which cracks are kept open. The preference for the surface might be related to the lower critical shear stress of surface dislocation sources compared to sources in the interior (Chapter III).

A very peculiar phenomenon observed in fatigued specimens is *extrusion*, by which is meant the rising of regions of the surface bounded within a coarse slip band above the material on either side of it. Also the reverse has been noted, viz. the formation of narrow crevices in slipbands, which have been called *intrusions*. COTTRELL and HULL [1957] showed that extrusions and intrusions occur about equally frequently (in fatigued copper), and that for their occurrence no thermal activation is needed: they form readily even at 20 °K. These authors proposed the *sequential* operation of two intersecting slip bands at the surface to be the cause of the phenomenon. The two systems are active during alternating half-cycles. Owing to the interference of the other system, the position of the successive slip steps at the surface due to the first system are continuously displaced, and

vice versa. Thus wide surface irregularities are formed, which may either have extrusion or intrusion character. The formation of intrusions, which are actually to be regarded as incipient cracks, is considered as one of the origins of the final fatigue failure.

The close connection between the formation of extrusions and the ultimate fatigue failure has also been stressed by MOTT [1958], who proposed an explanation in terms of cross-slip of the periodically moving screw dislocations as an alternative to the previous model.

The extent to which vacancy formation, dislocation pile-up eventually aided by cross-slip, or mechanical displacement of material can be held responsible for the actual fatigue fracture is still unknown at present. It might well be that the relative importance varies greatly with frequency, temperature and type of metal. That vacancies and interstitials play a role of some kind at least in some cases seems rather certain. It has been shown by MCCAMMON and ROSENBERG [1957] that copper partly fatigued at 20 °C shows a higher flow stress than the same metal fatigued at 100 °C, but that after annealing the former at 100 °C and thereby eliminating the vacancies, the two metals showed comparable stress-strain curves. More about the effect of point defects on fatigue can be found in the recent Symposium Report on Vacancies and other Point Defects [1958].

To mention still one other hypothesis, Machlin has related the fatigue phenomenon to the presence of micro-cracks in the specimen already before the test. At the ends of such cracks, large stress magnification factors will occur, and when loaded, dislocations may form near the tips and cause the cracks to extend through the material. Thermal fluctuations must surely aid this process, and this is in disaccord with the observed relative indifference to temperature of the fatigue phenomenon.

The essential feature of fatigue is that plastic deformation occurs in very localized regions, and there becomes successively more intense without notable extension of the region itself. The hypothesis of preexisting microcracks, however vague, could indeed provide an explanation for this. The intensification of the deformation progresses extremely slowly; the cracks extend (on the average) by only one atomic distance per cycle (HEAD [1952]). It is probable, however, that the deformation is an intermittent process, in that new slip lines are formed as a whole at randomly spaced intervals, which propagate rapidly, but discontinuously. The dislocation arrangement in fatigued

aluminium has recently been investigated by electron-transmission-microscopy (SEGALL *et al.* [1959]). Very inhomogeneous dislocation distributions were observed, in agreement with the previous statement.

A satisfactory theory of fatigue in terms of lattice imperfections and their properties is not yet available. We shall proceed with the discussion of the possible mechanisms of "normal" fracture, where the situation is only slightly better.

2. Fracture

Two types of fracture can be distinguished in crystalline solids. *Brittle* fracture without any or only very little previous plastic deformation is often observed in non-metals and in many metals at very low temperatures (but not in face-centred cubic metals). In the latter case one speaks of low-temperature brittleness. *Ductile* fracture on the other hand, which is common to most metallic solids at normal and elevated temperatures, takes place only at the end of considerable plastic deformation. Whereas brittle fracture is generally not characterized by a definite fracture stress and even occurs occasionally at very low applied stresses, ductile fracture usually takes place only when the material is (heavily) loaded above a certain limit (measures of quality such as the ultimate tensile stress of steel, etc., are based on this fact).

The first attempt to produce a theory of brittle fracture goes back to GRIFFITH [1920]. To explain the low value of the observed fracture strength of *glass* (less than 1% of the theoretical cohesive strength),

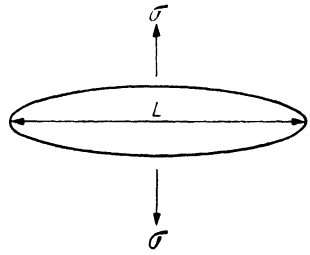


Fig. 3. Griffith crack.

he postulated the existence of *microcracks* of submicroscopic dimensions which could act as sites of considerable stress concentration. A flat ellipsoidal crack of diameter L (Fig. 3) in a field of tensile stress σ

normal to the mid plane of the crack, is accompanied at its tips by stresses of order of magnitude

$$\sigma_m = \sigma\sqrt{L/r}\,. \tag{1}$$

where r is the distance from the tip. If L is not *too* small, σ_m can easily exceed the applied stress by a factor of 100, within a few atomic distances from the tip and thus reach the theoretical fracture stress (about $G/10$, see Chapter V). In that case the crack *spreads*. A critical dimension of the crack, above which it will propagate, and below which it will not, can thus be defined as

$$L_{cr} = (G/10\sigma)^2\, b \tag{2}$$

(where b stands, as usual, for the atomic distance). This length is normally of the order of 1 micron or less.

In terms of lattice imperfections, the hypothetical crack must be replaced by an *array of dislocations*. Although fracture through the presence of voids is common in fatigue, and has been observed to occur sometimes also in minerals (and occasionally even in metals, for instance when solute gas is present: hydrogen brittleness of welds), such voids must nevertheless be regarded as exceptions to the rule.

The dislocation array most suited to initiate a crack is the piled-up group of dislocations. The stress magnification at the tips of the crack must now be replaced by the stress due to the dislocations themselves. As has been discussed in Chapter VIII, the number n of dislocations in a single slip plane that can be piled up against a barrier of some kind or another (e.g. a Lomer-Cottrell barrier) under the action of an applied shear stress σ is given by

$$n = \frac{2L\sigma}{Gb}\,, \tag{3}$$

where L is now the length of the pile-up. The stress at the head of such a system is of the order of n times the stress of the leading dislocation. When it reaches the theoretical cohesive strength, the pile-up will again develop into a crack (Fig. 4). The condition for this was found by STROH [1954, 1958a] and by PETCH [1954] as

$$\tau \approx \frac{G}{3}\sqrt{\frac{b}{L}}\,, \tag{4}$$

a formula which can directly be compared to (2), and is found to differ but slightly from it in numerical magnitude.

FRIEDEL [1956] describes the development of a crack from the pile-up in detail as the result of a transformation of the piled-up dislocations into socalled "cleavage" dislocations, linear steps along which the two crystal surfaces are separated. When the cleavage dislocation "climbs" along the fissure and extends it, the other piled-up dislocations can glide nearer to the barrier and produce subsequent cleavage dislocations.

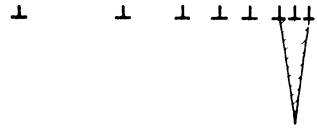


Fig. 4. A crack formed at the tip of a pile-up of dislocations.

FISHER [1955] has drawn attention to a different mechanism by which plastic deformation could lead to the formation of microcracks. This is closely related to the process by which moving intersecting screw dislocations produce either a row of vacancies or of interstitial atoms; in fact, instead of a row of discrete vacancies a continuous crack of atomic dimensions might be formed, when a moving screw dislocation crosses a *stationary* screw perpendicular to its slip plane. The crack will usually be stable, but in the presence of a normal stress of sufficient magnitude across the slip plane of the moving dislocation, it could propagate. This stress is computed to be indeed again of the order of magnitude of the observed fracture stress of metal single crystals. Fisher's mechanism can only operate at relatively low temperatures, where the "atomic" cracks formed do not close up immediately simply by diffusion of the vacancies.

In order to explain the effect of previous cold-work on the special kind of brittle fracture of single crystals which is called *cleavage*, FRIEDEL [1956] has pointed out another, opposite influence of intersecting screw dislocations, viz. their activity as obstructions for the gliding pile-up dislocations. Cleavage consists of brittle fracture along *one* (or a few) close-packed lattice planes over measurable distances (in metals it is especially easy to observe when only few slip systems can operate, such as in zinc crystals).

Heavy cold-work produces so many intersecting screws, that the formation of large pile-ups is more or less prevented. When nevertheless

cleavage occurs, the cleavage faces are no longer smooth, but contain steps, so-called *rivers*, as direct consequences of the presence of the screws perpendicular to the cleavage faces. Experimental support for this idea has been supplied by the work of GILMAN [1957].

A theory of cleavage which is *not* based on the presence of dislocation pile-ups has been developed by STROH [1958b]. Stroh considered a low-angle tilt boundary that ends somewhere within the crystal. As already mentioned in Chapters IV and XIX, this configuration constitutes the crystallographic analogon of a Volterra dislocation of the rotational type (according to the example labelled *g* of Fig. IV–7). It is accompanied by a high stress concentration at the end of the boundary, the magnitude of which has been estimated by FRIEDEL [1956]. The stress depends on the angle of misorientation, and becomes equal to the theoretical cohesive strength when this angle approaches about 5°. When this happens, the dislocation wall spreads into a crack.

A terminating wall might be formed in practice from a complete wall if the latter is induced to move and obstacles prevent part of it from doing so. GILMAN [1957] has shown that cleavage cracks can indeed be initiated in this way in zinc single crystals.

In all theories of brittle fracture, it is the *spreading* of the crack that requires most attention. Conditions for this have been set up by many and we must refer to the literature (STROH [1958a]) for a discussion.

A slight amount of cold-work is always necessary to produce pile-ups of sufficient dimensions to cause brittle fracture. When these dimensions are such that the intervening effects of the deformation can prevent cracking (according to the above screw dislocation mechanism or any other mechanism), the metal deforms plastically without fracturing until the dislocation density has become so large that the internal structure loses its consistency. One then has ductile fracture.

Mostly, during the later parts of the deformation, local strain inhomogeneities accelerate the final rupture. This is the basis of the well-known phenomenon of necking encountered in tensile tests.

Whether a crystalline material will fracture in a ductile or brittle manner depends on a number of factors, such as the grain size, the impurity content etc. For fracture to occur, pile-ups of sizeable dimensions must be formed, containing between a few hundreds and a thousand of dislocations in representative cases (this follows directly from the Stroh-Petch-Friedel theory). In pure single crystals this is

not very well possible, given the lack of obstructing objects, and fracture will occur only after large numbers of other dislocations have been generated throughout the crystal (ductile behaviour). Only when plastic yielding is in itself a very difficult process, such as is the case in many mineral crystals and impure metals or when intervening effects like twinning occur, which indeed give rise to obstacles for dislocation motion, can cleavage fracture take place without visible pre-deformation. In general, brittle fracture is to be expected when dislocations can either move or be generated only with difficulty. In this respect impurity atoms play an important role, in as far as they can pin dislocation sources.

When this occurs and there are suitable obstructions such as grain boundaries which can act as stopping sites for those dislocations that have nevertheless been generated, the conditions are ideal for brittle fracture to occur. In this way Friedel explained the well-known low-temperature brittleness of polycrystalline iron and ferritic steels containing a small amount of carbon (below 50° to 100 °K). The carbon blocks most of the dislocation sources at low temperatures, the grain boundaries aid the formation of piled-up groups. Whereas Friedel accentuated the importance of *impurities* in this respect, PETCH [1954] pointed out that the *grain size* is an equally controlling factor, and not only the presence of boundaries as such. From formulae (3) and (4) it follows that the cleavage strength σ_c is related to the size of the pile-ups, that is, also to the mean grain size in the polycrystalline material discussed here, according to an inverse square-root relation. There is a third factor that must not be underrated either, namely the *crystal structure* itself. Low-temperature brittleness at −196 °C has been observed only in body-centred cubic metals (with some exceptions such as vanadium and tantalum), in some hexagonal close-packed metals such as zinc, but not in face-centred cubic metals. This has occasionally been explained as follows: in the body-centred lattice impurity atoms interact elastically with edge dislocations as well as with screw dislocations, whereas in face-centred cubic lattices there is no elastic interaction with the latter (see Chapter IX). Therefore, the blocking of dislocations by impurity atoms should be most effective in the body-centred structure. However, the interaction is certainly not purely elastical and *pure* screw dislocations will be relatively scarce in any crystal.

COTTRELL [1958] has recently developed a more convincing theory

of low-temperature brittleness in iron and steel in which the effect of crystal structure on the geometry of dislocations was considered in detail. The nucleation of cracks was not regarded as a result of the piling up of dislocations, but rather of the coalescence of pairs of suitably oriented edge dislocations which together form *new* edge dislocations, which then act as wedges along the (001)-plane, thus favouring cleavage (this mechanism resembles that of the formation of a cleavage dislocation proposed by Friedel). In Fig. 5 Cottrell's mechanism is

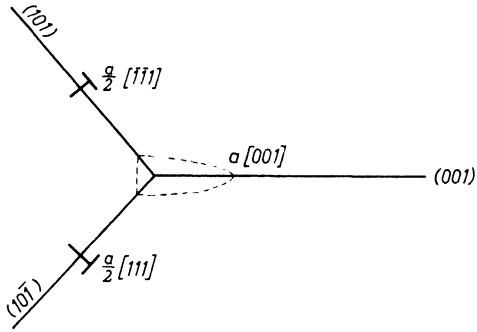


Fig. 5. Coalescence of two slip dislocations to form a cleavage dislocation with Burgers vector along [001] in the body-centred cubic lattice (after COTTRELL [1958]). More slip dislocations arriving at the node enlarge the Burgers vector of the cleavage dislocation and ultimately form a crack.

illustrated; it holds only in body-centred cubic metals because there a *decrease* of elastic energy can be obtained by the dislocation reaction:

$$\frac{a}{2}[\bar{1}\bar{1}1] + \frac{a}{2}[111] \rightarrow a\,[001]\,.$$

(The screw components of the two normal slip dislocations on the left hand side, which glide along the (101)- and (10$\bar{1}$)-planes respectively, eliminate each other. Such an elimination is not possible in the face-centred cubic lattice.) Cottrell actually believes a small crack to be present at the intersection between the slip planes, because of the large Burgers vector of the [001] dislocations. This crack requires a certain stress to spread, which can be shown again to be comparable to the Griffith stress. When this stress is smaller than the yield stress in polycrystalline material, brittle fracture occurs; if it is greater, the material deforms plastically. Since the yield stress is found to depend on the grain size (see PETCH [1953]; section XIII-3) and on temper-

ature, the occurrence of a brittle-to-ductile transition at a certain temperature, which depends on the grain size, as has been observed by many authors, can now be explained quantitatively.

It is not possible here to discuss the large body of experimental data pertaining to the influence of grain size and purity on the ductile-to-brittle transition. It seems at present as if Cottrell's semi-empirical theory combines to the right extent the effect of impurities with that of grain size (through the dependence of the *yield stress* on the latter!).

JOHNSTON and GILMAN [1959] stressed the importance of the *velocity* of dislocations and their rate of generation for the brittle or ductile behaviour of a crystalline solid. This idea pertains presumably best to mineral crystals of the heteropolar or homopolar type (see also Chapter XXX). The relative proportion between the velocity of propagation of a crack and the rate of initiation of plastic glide in a crystal determines the ductility: when the first is small compared to the second, ductile behaviour may be expected. The plastic deformation initiated in front of the crack will tend to produce more obstructions for the further propagation of the latter (such as intersecting screw dislocations), and thus reduces the propagation speed still further. Gilman has applied his theory to observations of the crack propagation in LiF crystals.

Low-temperature brittleness in steel could, in principle, be related to the existence of a delay time for yield in this material (section IX–4). When this becomes so large at low temperatures (of the order of 10^{-2} sec at -196 °C) that the stresses around a propagating crack build up in less time than is needed for yield to take place, the material shows brittle behaviour. Only experiment can decide on the validity of the various theories.

3. Intercrystalline Cracking

Cracks can, and generally will be formed at grain boundaries or subboundaries rather than in the interior of the crystal. When such planar imperfections are present, they can, however, also spread into cracks themselves; one then speaks of *intercrystalline* cracking (in contrast to *intracrystalline* fracture). The former occurs preferentially at high temperatures in most metals. CHEN and MACHLIN [1956] have proposed an interesting possible mechanism for intercrystalline fracture. A *step* in a grain boundary will act as an obstruction to grain boundary sliding, because it must either widen ("down-type step")

or it gives rise to compressive stresses ("up-type step"). If dilatational stresses act on the former a microcrack is formed, which subsequently expands. Up-type steps do of course not act in this way. According to this view, only the *shear* stress components along the grain boundary should affect the tendency for cracking, tensile stresses normal to the boundary should have no effect in as far as the *nucleation* of the cracks is concerned. Once the cracks are there, however, the presence of a normal stress component can of course accelerate its propagation into a cleavage step. This has been strikingly confirmed by Chen and Machlin in bicrystals of copper. CHANG and GRANT [1956], on the other hand, have observed heavy deformations (and even recrystallization) in the regions immediately adjacent to the intercrystalline fracture surfaces (of Zn-Al alloys deformed at 280 °C), which can not easily be explained by the picture given above. The conditions (creep test) were, very different from those in the bicrystal tests, however, which were, performed on very pure copper.

An extensive discussion of fracture theory and experiments is contained in a recent conference report (AVERBACH *et al.* [1960]).

REFERENCES

AVERBACH, B. L., D. K. FELBECK, G. T. HAHN and D. A. THOMAS (1960); Fracture (M.I.T. Technology Press and Wiley, New York).
BULLEN, F. P., A. K. HEAD and W. A. WOOD (1953); Proc. Roy. Soc. **A216, 332.**
CHANG, H. C. and N. J. GRANT (1956); J. Metals **8,** 544 and 1241.
CHEN, C. W. and E. S. MACHLIN (1956); Acta Metall. **4,** 655.
COTTRELL, A. H. (1958); Transact. Metall. Soc. AIME **212,** 192.
COTTRELL, A. H. and D. HULL (1957); Proc. Roy. Soc. **A242,** 211.
FISHER, J. C. (1955); Acta Metall. **3,** 109.
FRIEDEL, J. (1956); Les Dislocations (Gauthier-Villars, Paris).
FUJITA, T. (1954); Sci. Res. Rep. Tohoku Univ. **6,** 565.
GILMAN, J. J. (1957); Trans. AIME **209,** 449.
GRIFFITH, A. A. (1920); Phil. Transact. Roy. Soc. **A221, 163.**
HEAD, A. K. (1952); J. Mech. Phys. Solids **1,** 134.
JOHNSTON, W. and J. J. GILMAN (1959); J. appl. Physics **30,** 129.
MCCAMMON, R. D. and H. M. ROSENBERG (1957); Proc. Roy. Soc. **A242,** 203.
MOTT, N. F. (1955); Deformation and flow of solids, Madrid (Springer, Berlin).
MOTT, N. F. (1958); Acta Metall. **6,** 195.
PETCH, N. J. (1953); J. Iron Steel Inst. **173,** 25.
PETCH, N. J. (1954); Progress in Metal Physics **5,** 1.
SEGALL, R. L. and P. G. PARTRIDGE (1959); Phil. Mag. **4,** 912.
STROH, A. N. (1954): Proc. Roy. Soc. **A223, 404**; (1955): Proc. Roy. Soc. **A232, 548**
STROH, A. N. (1958a); Advances in Physics **6,** 418.
STROH, A. N. (1958b); Phil. Mag. **3,** 597.
THOMPSON, N. and N. J. WADSWORTH (1958); Advances in Physics **7,** 72.
Vacancies and other Point Defects in Metals and Alloys (Institute of Metals, 1958).

CHAPTER XXIII

CRYSTAL GROWTH AND IMPERFECTIONS

1. Growth of Crystals as a Dislocation Mechanism

Crystal growth from the vapour takes place at a much lower vapour pressure than would be expected from the rate of formation of growth nuclei on the surface of a perfect crystal. Such nuclei are schematically depicted in Fig. 1; they consist of *kinks* in *steps* on the crystal surface.

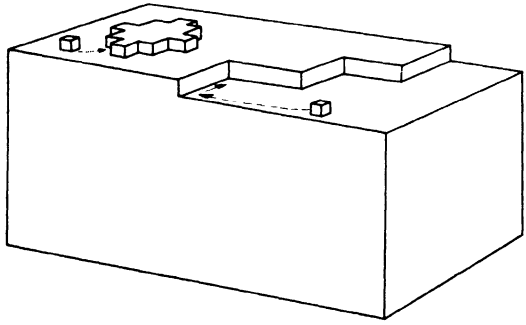


Fig. 1. Steps with kinks on the surface of a crystal, acting as possible transient nuclei for growth from a supersaturated vapour.

Conventional growth theory starts from the observation that usually the low-index crystallographic planes of a crystal exhibit a much lower surface energy than the other planes, owing to the high density of atoms in the former. At the absolute zero of temperature, therefore, a crystal tends to be bordered by low-index planes only, as smoothly as possible since this represents a state of lowest energy. At higher temperatures, atoms are continuously emitted and absorbed at the surface, and can diffuse out of the bulk unto that surface, thereby leaving a vacancy in the interior. Steps and isolated atoms will thus

form and disappear continuously on the surface in great numbers, and all kinds of lattice planes will be exhibited as bounding planes. If now the vapour pressure above the crystal surface is increased above the equilibrium pressure at the temperature of the system, more atoms will, on the average, arrive at the surface than are emitted from it, and the crystal *grows*. However, the addition of a new atom somewhere arbitrarily at the surface, where there is no kink or step, is energetically a very unfavourable process: the free energy of attachment is low, many unfavourable crystal planes are exhibited and the atom is easily evaporated off again. Only when the atom, once arrived, can by migration over the surface reach a kink before it is evaporated, does it have a chance to become permanently attached to the crystal: the number of free faces is then again reduced to what it was before. In this way all kinks will successively be filled up, and all steps will propagate, until the surface has become relatively smooth again. The only way in which growth can continue is by the accidental formation of new islands of absorbed atoms on the now perfect crystal faces, which are of such size that they will tend to grow rather than to dissolve again under the action of the surface tension (Fig. 2). The *critical size*

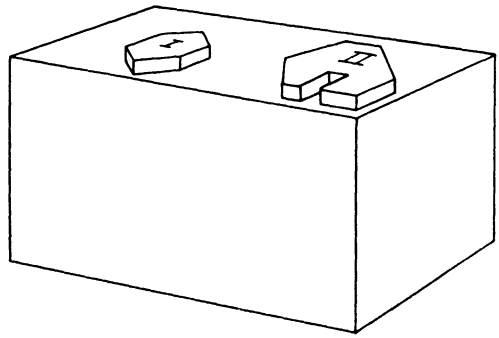


Fig. 2. Isolated "islands" on the surface of a growing crystal. I is smaller than the critical size and will ultimately dissolve, II is larger than the critical size and will grow out to form a new atom layer.

of such a nucleus is easily found to depend on the vapour pressure p as

$$R_c = \frac{\gamma a^3}{kT \log p/p_0} \tag{1}$$

where γ is the surface tension, a the lattice parameter and p_0 the equilibrium vapour pressure. A critical nucleus represents an increase

in surface energy of $2\pi\gamma a R_c$ and accordingly their formation under the action of thermal fluctuations alone can be expected to occur at a rate proportional to exp $(-2\pi\gamma a R_c/kT)$, which is usually an extremely small factor. In fact, observable growth rates cannot be expected to occur according to this theory below excess vapour pressures of 25%. Actually, however, growth is often observed at very low degrees of supersaturation, of the order of a percent (FRANK [1952]).

To explain this discrepancy, Frank considered the influence on the nucleation of the presence of *dislocations* in the growing crystal. As will be clear from Chapter III, a *self-perpetuating step* is present on the crystal surface where a dislocation emerges with an appreciable *screw* component more or less perpendicular to the surface (Fig. 3).

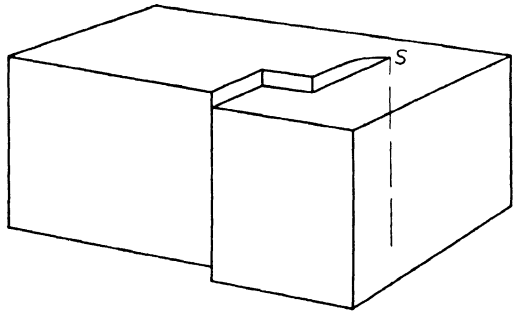


Fig. 3. Self-perpetuating growth step ending at the point of emergence of a screw dislocation (or dislocation with screw component perpendicular to the surface). The upper surface is now a spiral surface and the step cannot be eliminated by adding atoms to it.

Such a step can *never* be completely eliminated by adding atoms to it: since the end point of the step is fixed, addition of new atoms results only in a *rotation*. A straight step then assumes a curved form, because at first the angular velocity depends on the distance from the fixed point; far from this point, more atoms must be added to produce a certain angular rotation than near to the end point. In the course of time a stable shape of the rotating step develops, however. According to the nucleation theory mentioned above, parts of the step having acquired a radius of curvature less than R_c will on the average not absorb more atoms, only those parts with a radius of curvature larger than R_c will do this. The velocity of propagation of this remaining part will vary with curvature, because the equilibrium vapour pressure does so, and one can prove in this way quite generally that

the parts curved most tend to move most slowly. What finally results is a stationary spiral form of the step, which as a whole rotates with a *constant* angular velocity around the dislocation (Fig. 4).

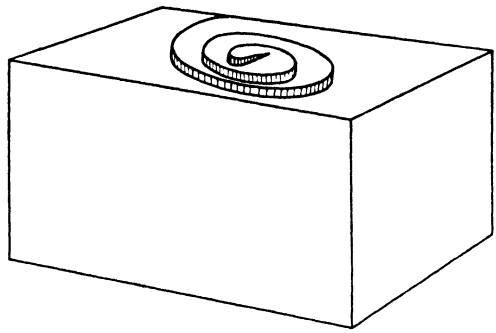


Fig. 4. The stable form of a rotating growth step is a spiral with equidistant arms (on an isotropic surface). This results in a spiral growth terrace with step equal to the screw component of the Burgers vector of the (perpendicularly emerging) dislocation.

From Frank's theory it follows that the form of this spiral can be approximated, at least when the surface tension is isotropic, by the equation (in polar coordinates r,θ):

$$r = 2R_c\theta - v_0 t \tag{2}$$

where v_0 is the velocity of advancement of a *straight* step according to the classical nucleation theory. This depends, of course, as R_c, on the degree of supersaturation $\Delta p/p_0$ of the vapour. If this is small, v_0 is proportional to $\Delta p/p_0$ and R_c is inversely proportional to this quantity (formula (1)). Since the rate of crystal growth is proportional to the number of times per second the spiral makes a complete turn and as this number is simply $v_0/4\pi R_c$, the rate of growth according to this *rotating step mechanism* is approximately proportional to the *square* of the degree of supersaturation. This result is in good accord with observations (such as those by VOLMER and SCHULTZE [1931] on the growth rate of iodine crystals from the vapour).

The linear distance between successive turns of the winding spiral (Fig. 4) is $4\pi R_c$, which is of the order of a micron in usual cases. The result of the growth process here described will thus be the formation of growth pyramids or terraces, with a step height of the order of one

atomic distance and a pitch of the order of 1 micron. Such shallow steps are not easily detected. However, as is well known, growth spirals of precisely the predicted shape but with a much larger step height (100 Å or more) have been observed in a large number of crystals in the last 15 years (several examples have already been known for 50 years, however). For a review of many observations we refer to the books of VERMA [1955] and of DEKEYSER and AMELINCKX [1956]. The reason for the large step height is that in some crystals "growth" dislocations occur with very large Burgers vectors ($\approx$ 100 Å), so that microscopic observation of the growth steps is possible. These growth dislocations occur either because of the large dimensions of the unit cell (organic crystals like paraffin), or because of the occurrence of *multiple* dislocations, having a combined Burgers vector of many times a lattice spacing (SiC, CdI_2). These can be considered as narrow bundles of many individual dislocations that have coalesced under the influence of surface tension. If the total dislocation core energy per unit area of such a multiple dislocation exceeds the surface tension per unit length of a hollow cylinder with radius large enough to contain all the dislocations, such a *hollow* core indeed exists in the centre of the growth spiral. However, this can occur in practice only when the combined Burgers vector exceeds at least 4 times the lattice vector.

The shape of a growth spiral generally varies with the type and structure of the crystal on which it occurs. Owing to the variation of growth rate with crystallographic direction, the turns often assume a polygonal shape. The emergence of two or more dislocations near to each other can give rise to various combined spiral structures or even closed loops of growth steps (due to two dislocations of opposite sign), of which some have a configuration that greatly resembles that of the dislocation arrangement around a Frank-Read source. Others are double-pitch spirals due to two very close dislocations of equal sign.

Although the dislocation mechanism can fully explain the discrepancy between observed and (classical) theoretical growth rates, the question remains how the dislocations come into existence in the nuclei from which the crystals subsequently grow. Some aspects of this question have already received attention in Chapters III and VI.

The problem varies according to the growth phenomenon considered. The Frank mechanism applies best to growth from a slightly supersaturated vapour or liquid. Another, technically very important growth process of crystals is the solidification from the melt. (This has also

been discussed in Chapter III.) Here the rotating spiral mechanism loses its meaning and must be replaced by another.

It has been found from many experiments that the circumstances under which the growth takes place, whatever be its nature, determine to a large extent the dislocation content of the new crystal. A very extensive discussion of recently obtained evidence is contained in a Symposium Report on the Growth and Perfection of Crystals [1958], to which the interested reader is referred. In this Report is contained inter alia the outline of a theory by Frank on growth from the melt, which we shall not discuss here, and a paper in which Dash reviews a number of possible mechanisms by which dislocations could form in a solidifying crystal. These are

1. Propagation by growth of dislocations already present in the solid crystal;
2. Multiplication of dislocation loops when they reach the solid-liquid interface, under the influence of thermal stresses;
3. Generation of dislocations by thermal stress from surface irregularities. This is probably a very important mechanism in practice;
4. Condensation of excess vacancies in the form of dislocation loops (cf. Chapters III, VI and XV);
5. Segregation of impurities, leading to regions with a different expansion coefficient;
6. Growth accidents, such as poor epitaxy.

The subject *epitaxial growth* has received much attention in the last few years. When a crystal forms on the surface of another crystal with a narrowly related but not exactly the same structure, often the orientation of the newly formed crystal stands in a rigid relation to that of the substratum. One then speaks of epitaxial growth. This can extend over macroscopic distances only when at the interface between the two crystals (joint) an occasional dislocation of suitable Burgers vector appears, to take up the integrated amount of relative misfit between the two crystal lattices, thereby compensating the internal stresses. The interface with its dislocations can be directly compared to a grain boundary (FRANK and VAN DER MERWE [1949]). The Burgers vector of the epitaxial dislocation is usually not a primitive lattice vector, but a defining vector of the coincidence plane which need not be a very low index plane (although in many instances it is).

2. Special Growth Phenomena

2.1. POLYTYPISM

A phenomenon which is closely related to the formation of crystals by the dislocation mechanism is the so-called *polytypism* of crystals. By this, we understand the occurrence of a regular periodic variation in the order of stacking of lattice planes, extending over macroscopic distances without any appreciable change of regularity. Carborundum and cadmium iodide, the same substances on which conspicuous growth pyramids are often present, display polytypism in a very pronounced manner. It is not so much the variation in stacking order that poses a question—similar phenomena are encountered in many substances, and even in close-packed metals—it is the strict regularity of the variation which is remarkable, which can for instance cause that, on every fifteenth lattice plane, a change of structure occurs. The existence of dislocations with large Burgers vectors as indicated by the occurrence of growth spirals of large step height, points to the explanation. Evidently, polytypism is directly related to the dislocation growth process, in the sense that the growth dislocation (or rather the combined multiple growth dislocations) produces a stacking of layers with a periodicity of deposition which is many times a lattice vector. Stacking disorders contained accidentally in the rotating growth step repeat in every complete revolution of the step, thus with a periodicity equal to the height of this step.

2.2. WHISKER GROWTH

A very remarkable type of crystal growth has already received attention previously (Chapter VI), namely the *filamentary growth* or *whisker* formation that occurs in many metals and other substances, spontaneously or only under especially prepared (physical or chemical) conditions. From the mechanical properties (high tensile strength and large elastic limit) of very thin whiskers (less than a few μ in diameter), it has been concluded that those whiskers are crystals containing none at all or only a few dislocations. Recent evidence shows that the strength of whiskers depends very sensitively on their size: large-sized whiskers are "weak" like ordinary crystals, medium sized (1–10 μ) whiskers show a very high yield strength, but once they *have* yielded, the flow stress is again comparable to that of ordinary crystals, whereas very thin whiskers do not deform at all but fracture at approximately

the theoretical fracture strength (BRENNER *et al.* [1958]). Whether a whisker contains none, one or a few dislocations seems not to have much influence on this, probably because the size determines primarily what the dislocations can *do*, not whether they can *exist*.

The growth of whiskers has been shown to occur in the length direction only, although it differs according to the way of preparation whether they grow at the tip or at the base.

Several mechanisms have been proposed for whisker formation. Most of them are based on the theory of dislocation growth, and a fairly complete critical review of all of them has been given recently by NABARRO and JACKSON [1958], to which we refer. The theory which seems most suited to explain the observed properties and growth rate is based on an original idea by ESHELBY [1953] and FRANK [1954], in which the growth mechanism is considered as essentially a form of dislocation climb. It bears some resemblance to the pole mechanism of twin growth discussed in Chapter XIII. A screw dislocation ends very near to the surface of a crystal and is pinned there, for example by a foreign atom. To reach the surface, the dislocation must turn through a right angle, and then becomes edge type. If the edge portion now turns around the anchoring point as a pole, it describes a spiral ramp, since after each turn the segment is displaced upwards by one lattice vector. A whisker is then extruded, provided there is a suitable driving force present, and the (surface ?) diffusion is sufficiently rapid to ensure the transport of material to the rotating dislocation (formation of an extra half plane that continuously extends in length). The maximum amount of material that can be supplied determines thus the width of the whisker and the speed of growth. That, as observations show, contamination of the surface (oxydation) or the application of a hydrostatic pressure to the solid accelerates the rate of growth considerably, is in principle explainable on the basis of a *basal* growth mechanism such as this. Other related mechanisms of base-whisker growth by a dislocation mechanism have been proposed, which were all unified by AMELINCKX *et al.* [1957]. The latter authors pointed out the close relation between the mechanism of whisker growth and the concept of a *spiral prismatic* dislocation and the *helicoidal* dislocation, as explained in Chapter III. It should be noted that a whisker formed according to this theory need not contain a screw dislocation along its axis, a question that has been much commented upon. The Frank-Eshelby-Amelinckx theory can explain the outlines of whisker growth,

but as soon as details, such as pressure dependence and finite length of the whiskers are considered, quantitative difficulties arise. To solve these, from case to case other factors should be considered.

Several instances of whisker growth of ionic crystals are known which can be regarded simply as special cases of normal crystal formation, leading to very small crystals (for instance the chemical deposition of CdS crystals from a vapour phase).

It is not necessary to postulate the existence of a particular dislocation mechanism here. The mechanical and structural properties of such "natural whiskers" are entirely comparable to those of the metal whiskers discussed above. They are determined only by the small dimensions of the crystal, as has been shown quite convincingly by PEARSON, READ and FELDMAN [1957], who compared silicon whiskers grown on a substrate with thin etched-down silicon crystals and found exactly similar behaviour with regard to the mechanical properties.

To close this chapter we may mention here the very imperfect single crystals which have been prepared and studied inter alia by Hilsch and his collaborators since a number of years. These are obtained by vapour deposition on a very cold substrate. Such crystals show highly interesting physical properties, connected with their large degree of structural disorder (e.g. the superconductive transition temperature is increased). It would carry us too far to include this interesting work in our discussion.

REFERENCES

AMELINCKX, S., W. BONTINCK, W. DEKEYSER and F. SEITZ (1957); Phil. Mag. **2,** 355.
BRENNER, S. S. (1958); in: Growth and Perfection of Crystals (Wiley, New York).
DEKEYSER, W. and S. AMELINCKX (1956); La Croissance des Cristaux.
ESHELBY, J. D. (1953); Phys. Rev. **91,** 755.
FRANK, F. C. (1952); Advances in Physics **1,** 91.
FRANK, F. C. (1954); Phil. Mag. **44,** 857.
FRANK, F. C. (1958); in: Growth and Perfection of Crystals (Wiley, New York) p. 1.
FRANK, F. C. and J. H. VAN DER MERWE (1949); Proc. Roy. Soc. **A198,** 205.
Growth and Perfection of Crystals (1958); (Wiley, New York).
NABARRO, F. R. N. and P. J. JACKSON (1958); in: Growth and Perfection of Crystals (Wiley, New York).
PEACH, M. O. (1952); J. appl. Phys. **23,** 401.
PEARSON, G., W. T. READ and W. L. FELDMAN (1957); Acta Metall. **5,** 18.
SEARS, G. W. (1955); Acta Metall. **3,** 361 and 367.
VERMA, R. (1955); Dislocations and crystal growth (London, Butterworth).
VOLMER, M. and W. SCHULTZE (1931); Z. physik. Chemie **A156,** 1.

PART III

IMPERFECTIONS IN NON-METALLIC CRYSTALS

CHAPTER XXIV

FUNDAMENTAL PROPERTIES OF IMPERFECTIONS IN POLAR CRYSTALS

1. General Distinctions between Imperfections in Metals and Non-metals

There are several reasons why metals and non-metals should be dealt with separately, the strongest of which concern the differences in structure and typical properties of lattice imperfections. In metals, the geometrical structure of imperfections is but slightly influenced by the interatomic bonding forces. The electron gas heavily screens off the irregularity from its surroundings, and the influence of imperfections on the *electrical* and *optical* properties of a metal is but small. On the other hand, the rigidity of the lattice structure is severely reduced in an imperfect lattice, and the *mechanical* properties of a metal are therefore seriously affected.

Non-metallic solids can be roughly divided into three classes: polar or ionic solids, homopolar or valence crystals and van der Waals solids; the latter group we shall pass by in this book, since very little is known about their mechanical properties. Point defects in ionic crystals are characterized by an electric charge which is much less effectively screened off than in metals. Therefore, they have a strong electrostatic interaction with the ions and electrons (which results for instance in the formation of *centres*). Dislocations can act as sources of point defects and influence the equilibrium concentrations locally. Dislocations thus affect for instance the liability of a crystal to attain centres under certain treatments. On the other hand, their influence on the mechanical properties is less paramount than in metals, although certainly not negligible.

Valence crystals, among which are numbered the technically important homopolar semiconducting elements and compounds such as germanium, silicon, indium antimonide and silicon carbide, are completely brittle at low temperatures and only become plastic above

about $\frac{2}{3}$ of their absolute melting temperature. The mobility of dislocations in these substances depends strongly upon the temperature; this is intimately connected with the presence of directed valence bonds between the atoms, which render the structure very rigid. The electronic configuration in these materials is such that even extremely small defect concentrations have a large effect on the electrical properties, either through their influence on the number of charge carriers, or through their effect on the mobility, recombination rate (lifetime), etc. Therefore, whereas at normal temperature in homopolar materials the mechanical properties are scarcely affected by the presence of imperfections, these make their presence known very markedly in the electrical properties, such as resistivity and Hall effect. Several general and fundamental aspects of dislocation motion in crystals can, on the other hand, be studied with advantage by investigating the mechanical properties of valence crystals at elevated temperatures; among these are the Peierls force and the interaction of dislocations with impurities.

2. Point Defects in Polar Crystals

Part of the contents of this section has already been touched upon in Chapters I and II; some repetition cannot be avoided here.

Polar crystals are built up out of differently charged ions situated at well-determined lattice sites. These interact with each other electrostatically, every ion occupying an equilibrium position (and of course, vibrating about it) in the electrostatic field of the other ions. The simplest compounds of this type contain two kinds of ions which are oppositely charged. The alkali halides NaCl, KCl, LiF, etc. and the silver halides AgCl, AgBr, etc. represent the binary compounds of this kind which have been most extensively studied in recent times.

It is common usage to distinguish between several kinds of *pairs* of lattice defects in polar compounds. This distinction, which has little sense in metals, arises from the extremely stringent condition of electroneutrality that must always be fulfilled, in perfect as well as in imperfect crystals. One distinguishes primarily between Schottky defects, which consist of a positive-ion vacancy and a negative-ion vacancy, not necessarily in each others immediate neighbourhood, and Frenkel defects, pairs of vacancies and interstitial ions of the same sign. Of course, also pairs of interstitials of opposite charge can exist

together (anti-Schottky defects). The condition of neutrality ensures that in pure materials defects occur always in pairs of either type. This means that also the generation of defects in thermal equilibrium occurs in pairs. The simple formula (II–3) for the thermal equilibrium concentration of point defects in metals must then be replaced by the slightly more complicated expressions:

$$c_1 c_2 = \text{const. } e^{-U_{S,F}/kT}, \tag{1a}$$

$$c_1 = c_2; \tag{1b}$$

where c_1 and c_2 stand for the concentrations of both constituents of the defects and U_S or U_F is the energy of formation of the pair. Combining the two expressions one has

$$c_1 = c_2 = \text{const. } e^{-U/2kT}. \tag{2}$$

The pre-exponential factor can easily reach values of the order of 10^2 to 10^4 as was shown by MOTT and GURNEY [1946]. This reflects the large dependence of the free energy of formation of a vacancy or an interstitial ion on temperature (a large thermal expansion coefficient usually pertains to ionic crystals). Since the lattice vibrations are oppositely affected by interstitial atoms and by vacancies, in formula (2) the pre-exponential factor will usually be considerably smaller when Frenkel defects are concerned than when it applies to Schottky defects. A further discussion of this can be found in Chapter II.

The theoretical calculation of the energy of formation of point defects in polar crystals is somewhat easier than in metals, since practically this whole energy arises from the electrostatic forces in the lattice. A well-known method to describe the magnitude of these forces in a perfect polar lattice is by means of the Madelung constant α which measures the electrostatic lattice energy in units of e/a. For the formation of a Schottky defect this energy is required in the first place. However, part of it is regained owing to the relaxation of the lattice around the vacant lattice site, and part owing to the polarization of the ion shells. This polarization is the larger, the greater is the dielectric constant of the material. MOTT and LITTLETON [1938] were the first to obtain theoretically a numerical estimate for the formation and migration energies of the various defects in alkali- and silverhalides. FUMI and TOSI [1957] and KUROSAWA [1958] have recently revised these computations. Table 1 shows some of the theoretical data and also some experimental values (between brackets)

most of them drawn from SEITZ's [1946, 1954] extensive review papers, in which also other references can be found (see further SEEGER [1956] and PRATT [1958]).

TABLE 1

Theoretical energies of formation, U_S and U_F, of Schottky and Frenkel defects respectively, and diffusion energies, Q, of the various point defects in some simple ionic crystals. Experimental values, where known, are placed in brackets.

Substance	U_S (eV)	U_F (eV)	Q_V^-(eV)	Q_V^+(eV)	Q_I^-(eV)	Q_I^+(eV)
NaCl	1.86 (2.02[a])		(1.28)[m]	(0.72–0.85)[a,k,l]		
NaBr	2.01 (1.71[a])		(1.18)[l]	(0.80)[e]	(≈1.2)[e]	
KCl	2.08 (2.1[b]–2.4[c])		(1.4)[m]	(0.68)[f]–(0.77)[n]	(≈2.0)	
KBr	1.92			(0.65)[n]		
LiF	2.34 (2.7[d])			(0.65)[d]		
AgCl	1.98	1.4–1.8		(0.36)[o]		(0.1)[o]
AgBr		1.1		(0.36)[p]		(0.15)[p]

[a] ETZEL and MAURER [1950]; [b] WAGNER and HANTELMANN [1950]; [c] KELTING and WITT [1949]; [d] HAVEN [1950]; [e] SCHAMP and KATZ [1954]; [f] ASCHNER [1954]; [h] FRIAUF [1957]; [k] MAPOTHER, CROOKS and MAURER [1950]; [l] HAVEN [1958]; [m] HARRISON [1958]; [n] GRÜNDIG [1960]; [o] COMPTON and MAURER [1956]; [p] TELTOW [1949]. V+ means positive-ion vacancy, etc.

To form a Frenkel defect, a positive or negative ion must be removed, and another ion of the *same* charge must be made interstitial. The contribution to the energy of formation by the first process can be computed fully along the lines set above, but that due to the second process requires a very difficult theoretical investigation, with uncertain results. Therefore, the energies of formation of Frenkel defects are not very well known theoretically, not even in the alkali halides.

As mentioned before, around point defects a certain relaxation (polarization) of the surrounding ions occurs, which in terms of volume expansion amounts theoretically to 6.5% for a positive-ion vacancy, 7.2% for a negative-ion vacancy in NaCl (MOTT and GURNEY [1946]). The nearest-neighbour ions of a vacancy are displaced outwards by this amount, in accordance with the fact that a negative ion vacancy can be regarded as carrying a positive charge, and thus repulses its nearest positively charged neighbours, and vice versa. Similarly, a lattice expansion is associated with an interstitial ion, although no

reliable computations of its magnitude exist. (Frenkel defects play no important role in alkali halides, the only group studied extensively theoretically. Moreover, the short range repulsive energy between ions, which is important in this case, is not known very well.)

A Schottky defect can also occur in the *associated* form, thus as a *divacancy*, the resultant charge of which is zero. There is a certain association energy for such a pair that was estimated by SEITZ [1954] (taking only Coulomb interaction into account) to be about 0.9 eV in NaCl and KCl. More recently Fumi and Tosi recomputed this energy and found only 0.6 eV. Such a pair should be extremely mobile, the migration energy being estimated as only 0.5 eV or less. PRATT [1954] pointed out that such a divacancy has a typically elongated lattice strain pattern around it; along the direction of the axis of the pair the displacements of the neighbouring ions are larger than perpendicularly to this axis (Fig. 1).

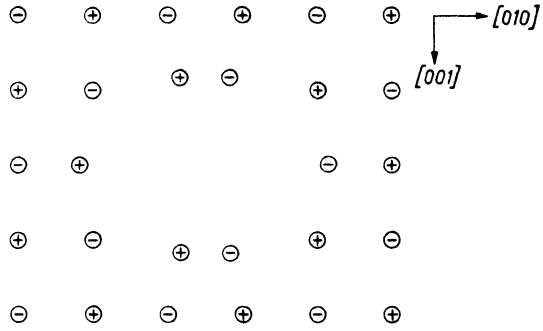


Fig. 1. Lattice deformation in (100)-plane around a neutral vacancy pair in the sodium chloride lattice (after PRATT [1954]).

Association (characterized by an energy of the same order) will also occur between negative ion vacancies and divalent impurity ions when present in the lattice. This will be discussed in section XXV–3.

3. Equilibrium of Point Defects in Polar Semiconductors

3.1. POLAR SEMICONDUCTORS

As has already briefly been indicated in the previous section, numerous groups of compound materials exist which are more or less ionic in nature, but which apart from ionic conduction, under normal circumstances also show appreciable conductivity by electronic motion.

They form the transition between the purely heteropolar and the homopolar substances, to be discussed later in this book. Many oxydes belong to this class (NiO, sapphire) as well as sulphides (ZnS, CdS, PbS) and selenides. (Of course, *all* purely ionic crystals show *some* electronic conductivity at elevated temperatures.) In the band picture, the "transition" compounds are characterized by a band gap which is not very wide; in other words, the energy needed to transfer an electron from a negative ion to the conduction band is smaller than in most pure polar crystals and can already be provided to a certain extent by thermal vibrations at room temperature (see section 6).

The existence of many kinds of charged particles simultaneously, influences the occurrence and concentration of carriers and of lattice defects in a typical manner. The theory of imperfections (point defects) in these crystals (primarily in diatomic compounds) has been extensively developed by KRÖGER and VINK [1955, see also 1958]. Although the subject lies a little outside the purview of this book, we nevertheless thought it useful to devote a few words to it, as its importance for solid state physics is paramount.

Kröger and Vink treated the matter largely from a physico-chemical point of view. They approached the problem of equilibrium concentrations by the theory of chemical reaction rates (law of mass action), and introduced a specifical nomenclature to be used for two-atomic compounds. These compounds are denoted by the formula MX where M denotes the metal ion, and X an element of a more electronegative character.

Three kinds of disorder by imperfections are distinguished. As explained in section 2, the all-prevailing condition of charge neutrality (except near the surface or near dislocations, see section 5) necessitates the presence together of pairs of oppositely charged defects in pure material. The components of such a pair can be either distributed at random, or be present at adjacent sites (associated pairs). One distinguishes between:

a) *Frenkel disorder*, in which metal vacancies (V_M) exist together with interstitial metal ions (M_i), or X-vacancies (V_X) together with interstitial X-ions (X_i);

b) *Schottky disorder*: vacant M-sites (V_M) together with vacant X-sites (V_X).

c) *Antistructure disorder*: M-atoms occupying X-sites (M_X) and X-atoms occupying M-sites (X_M). This kind of disorder can be

expected to occur only when the components of the crystal have approximately the same electronegativity and atomic radius.

d) *Anti-Schottky disorder*: Interstitial M-atoms (M_i) together with interstitial X-atoms (X_i). This is rare and will not be considered further.

M and X without suffix denote atoms occupying normal sites.

3.2. EQUILIBRIUM REACTIONS

The formation of defects can now be described by quasi-chemical reactions. For instance, the formation of δ (molar concentration) isolated (metallic) Frenkel defects can be written as:

$$MX \rightleftharpoons M_{1-\delta} + (V_M)_\delta + (M_i)_\delta + X\,, \tag{3}$$

whereas the formation of δ associated Schottky defects is denoted by

$$(1-\delta)\,MX \rightleftharpoons M_{1-\delta} + X_{1-\delta} + (V_M V_X)_\delta\,. \tag{4}$$

According to the law of mass action, the product of the concentrations of the various reaction partners depends on the temperature through the reaction constant:

$$K = e^{-\mu/kT} \tag{5}$$

(μ is a chemical potential).

Generally, $\delta \ll 1$, and thus the concentration of atoms occupying the normal sites does not change appreciably. Then, the law of mass action pertains to the concentrations of the imperfections only. Rules of the same type govern the dissociation equilibria of pairs of defects. In addition to the law of mass action, the condition of charge neutrality plays a role. To find, for instance, the various equilibrium concentrations when simultaneously single Frenkel defects and single Schottky defects are present, one has to obtain the solution of the following three simultaneous equations (brackets denoting concentrations):

$$\begin{aligned} [V_M]\cdot[M_i] &= C_F e^{-W_F/kT}\,,\\ [V_M]\cdot[V_X] &= C_S e^{-W_S/kT}\,,\\ [V_M] &= [V_X] + [M_i]\,. \end{aligned} \tag{6}$$

The concentrations of electrons and *holes* (defect electrons which behave as positive charge carriers, see section 6), can be treated in complete analogy.

In material containing defects, apart from direct excitation from

band to band, excitation of an electron associated with an imperfection (*centre*) to the conduction band and vice versa, and excitation of an electron from the valence band (normal valence electron) to a localized state on an imperfection and vice versa, are feasible. This is indicated schematically in the energy band picture of Fig. 2. Here, energy levels within the forbidden energy band have been assigned to various imperfections, and the transitions mentioned have been denoted by arrows.

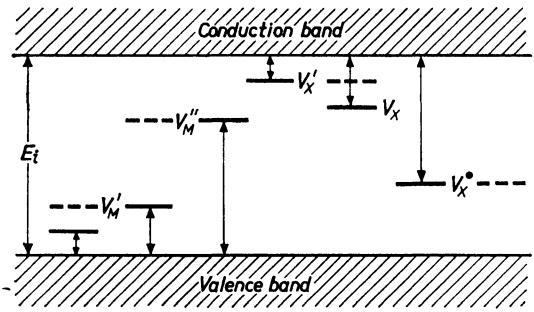


Fig. 2. Electronic imperfection levels (centres) in the energy band scheme of a binary semiconducting compound containing Schottky defects. Occupied levels are denoted by a solid line, unoccupied levels are dashed.

An example of a transition of the first kind is given by the ionization of a neutral interstitial metal ion (*donor* activity of a metal ion):

$$M_i \rightleftarrows M_i^{\cdot} + \ominus , \tag{7}$$

where the symbol $\ominus$ denotes an electron in the conduction band, and the dot indicates a positive charge on the defect. In the Kröger-Vink notation the symbols + and — indicate the *nature* of the defects rather than their charge.

An example of a transition of the second kind is, for instance, the ionization of a negative interstitial X atom into a neutral interstitial and an electron:

$$X_i' \rightleftarrows X_i + \ominus . \tag{8}$$

Another way in which the negative charge (indicated by the ′) on the interstitial could be neutralized is by absorption of a hole $\oplus$ from the valence band:

$$\oplus + X_i' \rightleftarrows X_i . \tag{9}$$

A large variety of reactions between the various kinds of defects

and electrons and holes can be imagined, of which many occur in practice. The rates of all of these are governed in thermodynamic equilibrium by the law of mass action. Thus, for reaction (7) one has for instance

$$\frac{n \cdot [\mathrm{M}_\mathrm{i}^{\cdot}]}{[\mathrm{M}_\mathrm{i}]} = C e^{-E_\mathrm{D}/kT}, \tag{10}$$

where n denotes the concentration of free electrons and E_D the dissociation (ionization) energy of the reaction. The concentration of atomic defects, electrons and holes, the latter either within the energy bands or localized at the centres are mutually interdependent, since the various relations (6), (10), etc. have factors in common. This makes the computation of actual equilibrium concentrations a fairly intricate problem. Apart from these equilibrium relations and the electroneutrality condition, a relation must hold as a consequence of the fact that the total number of M-atoms in the various possible states is always equal to the total number of X-atoms. This relation no longer holds when the compound is not purely stoichiometric (as is for instance the case in additively coloured crystals, compare Chapter XXVI); then it is replaced by one or more other relationships, which express the equilibrium with the surrounding atmosphere. We shall come back to this below.

M_i-ions can usually be easily ionized, that is, the corresponding occupied levels often lie close to the bottom of the conduction band. X_i-ions, on the contrary, usually attract an electron easily, and *occupied* levels often lie close to the bottom of the conduction band. Similar arguments point to V_M'- and $\mathrm{M}_\mathrm{X}^{\cdot}$-levels close to the valence band and $\mathrm{V}_\mathrm{X}^{\cdot}$- and X_M'-levels close to the conduction band. No quantitative weight should be attached to these conclusions, however.

Departures from the equilibrium concentration of carriers and imperfections can be produced by rapid cooling (quenching). When the imperfection centres possess energy levels far from the edges of the energy bands, the system can in this manner be obtained in a *metastable* excited state. Transition to the equilibrium state only occurs after a long time interval or on heating. In specific cases this transition may be accompanied by the emission of radiation (luminescence, phosphorescence).

In general, the presence of imperfection levels within the forbidden energy band will not give rise to an appreciable "free" electron or

hole concentration and thus to an increase of conductivity, since the levels due to the oppositely charged component defects are usually situated approximately symmetrically relative to the band edges. The full levels are then always below the middle of the band gap and the empty levels above it. Consequently they are not ionized. The situation may, however, be different in compounds with a non-symmetrical level distribution (non-binary compounds).

3.3. COMPUTATION OF CONCENTRATIONS

To obtain numerical information about the equilibrium concentrations and occupation of the various centres, BROUWER [1954] has developed an approximate method of calculation. The various product relations of the type (6), (10), etc., transform into a set of simultaneous linear equations by taking the logarithm of both sides. For a solid containing only simple Frenkel disorder of metal ions, with a band scheme like that shown in Fig. 3, one obtains the following set of four equations:

$$\begin{aligned} \log [\mathrm{M_i}] &+ \log [\mathrm{V_M}] = \log K_\mathrm{F} \, ; \\ \log [\mathrm{M_i^\cdot}] &- \log [\mathrm{M_i}] + \log n = \log K_\mathrm{M} \, ; \\ \log [\mathrm{V_M'}] &- \log [\mathrm{V_M}] + \log p = \log K_\mathrm{V} \, ; \\ \log n &+ \log p = \log K_\mathrm{i} \, ; \end{aligned} \tag{11}$$

(p is the concentration of holes).

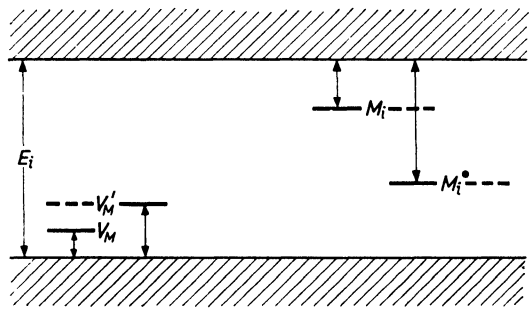


Fig. 3. Electronic energy scheme corresponding to a simple case of Frenkel disorder.

To this set must be added the stoichiometry condition (in a pure stoichiometric compound), or an equivalent equilibrium relation in non-stoichiometric compounds, and the neutrality condition. The latter

two are not product relations but contain *sums* of the concentrations. However, in most cases on both sides of these equations one term dominates over all the others, and therefore the equations may be approximated by simple equalities between two concentrations only. These can again be written as linear equations between logarithms. What the dominating quantities are, depends inter alia on the values of the various activation energies occurring in the reaction constants on the right hand sides of the equations (11). Therefore in *various regions of temperature*, the dominating terms are usually *different*.

By solving the set of linear equations, one obtains the logarithmic concentrations of the various imperfections and charge carriers, in the form of analytical expressions in the logarithms of the reaction constants. It is then possible to check the correctness of the approximations in the various regions of temperature, and change the treatment accordingly.

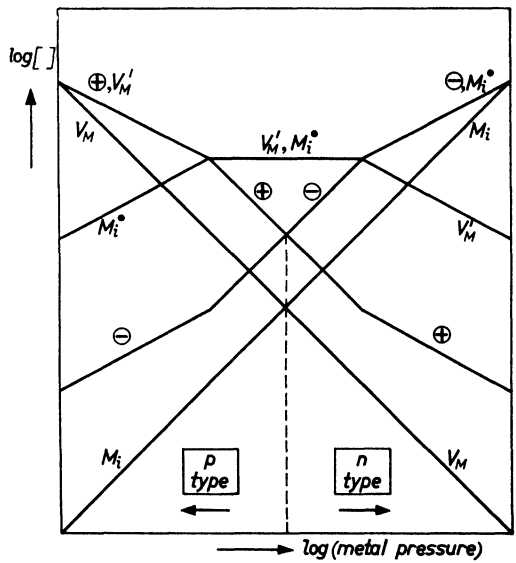


Fig. 4. Concentration-pressure chart for a crystal MX with Frenkel disorder according to the level scheme of Fig. 3, under the condition that $[V_M'] \cdot [M_i] \gg n \cdot p$ (after KRÖGER and VINK [1955]).

The method becomes really powerful when applied to a compound in equilibrium with its vapour, in order to study the variation of the concentration of the various centres with the vapour pressure. Then,

in different ranges of temperature, the neutrality condition changes, which leads to a variation of the concentrations with pressure. By making different approximations to the neutrality condition but now in various regions of *pressure*, this dependence can be approximated by a "piece-wise linear" relation, similarly to the approximation of the temperature dependence as explained above. Once this relation is known, the exact solution can be obtained without much trouble. An example of a "concentration-pressure chart" thus obtained is shown in Fig. 4. It pertains to the case of simple Frenkel disorder discussed above, for a compound MX in equilibrium with its metal vapour. We refer to the Kröger-Vink paper for a more thorough discussion. It is found as a general result that the concentration of free electrons usually increases and that of free holes decreases with increasing metal vapour pressure; the same conclusion holds for crystals with Schottky disorder. In both cases an increase of the vapour pressure results pre-

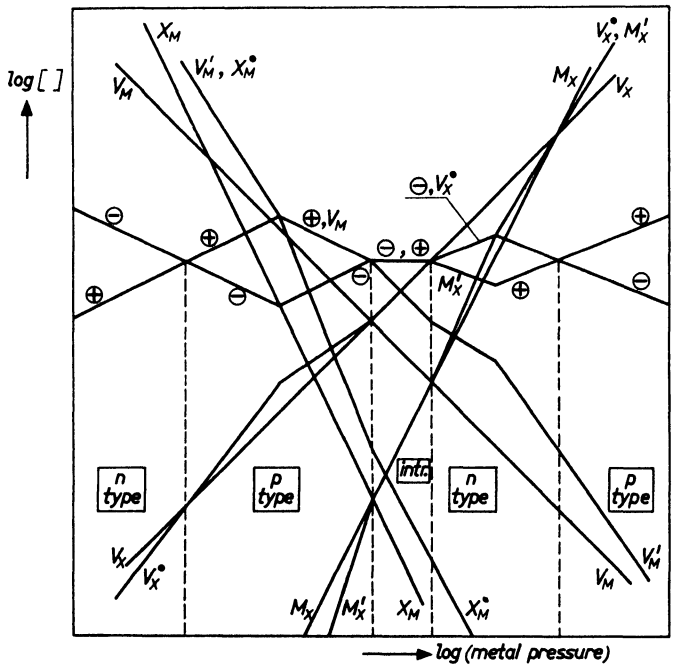


Fig. 5. Equilibrium concentrations of imperfections in a binary compound with simultaneous Schottky and anti-structure disorder, at high temperature in equilibrium with the metal vapour (after KRÖGER and VINK [1955]). In this example the concentrations of free electrons and holes are approximately *independent* of the pressure.

dominantly in more occupied M_I and V_X levels, close to the conduction band.

In crystals with antistructure disorder, however, the opposite is true: there the concentration of free holes increases and that of free electrons decreases with increasing pressure, as then more empty $M_X^{\cdot}$ levels are formed close to the valence band. More complex cases, where several kinds of disorder are present simultaneously, can in some instances show an approximate *in*dependence of the carrier concentrations on the vapour pressure (Fig. 5 gives an example).

3.4. IMPURE CRYSTALS

A still more complicated situation arises when *foreign* atoms are incorporated in the crystal.

For example, impurity ions of deviating charge enhance the ionic conductivity in polar compounds, because, to compensate for the excess charge, point defects must be present simultaneously (section 2). Other impurities can act as traps for electrons or holes, and give rise to colour centres (Chapter XXVI). A third example is constituted by the donor and acceptor action of atoms of deviating valency in homopolar semiconductors (Chapter XXVIII).

It is possible to extend the method of description set forth in the previous section to the impurity case. Substitutionally placed foreign atoms in the compound MX are logically denoted by F_M or F_X according to the site they occupy. When they carry a positive charge they are denoted by $F_M^{\cdot}$, etc. Interstitial foreign atoms are designated by F_I, once ionized they become $F_I^{\cdot}$, etc. A substitutional foreign atom with more valence electrons than the atom which it replaces, usually gives off easily the excess electron(s) to the conduction band. In this it resembles an interstitial M atom, and acts as a donor. A substitutional foreign atom with less electrons accordingly resembles an interstitial X ion, and usually shows acceptor action. In actual cases this distinction cannot often be clearly made. It is difficult to predict more or less generally the position of the levels associated with an *interstitial* foreign atom. Electropositive elements (metals) will usually act as donors since they tend to give off one or two of their valence electrons to the conduction band; electronegative interstitial atoms usually behave as acceptors.

The equilibrium conditions between the concentrations of the

various defects in the crystals remain unchanged in the presence of foreign atoms. However, the neutrality condition now also contains terms in the concentration of the latter. Further, the set of rate equations is extended by (for instance) equations describing the possible association equilibria between impurities and defects, and between impurities and electrons (ionization equilibria). Finally, the sum of the concentrations of the foreign atoms F in all possible states must be equal to the total introduced concentration of F. Again a complete and sufficient set of equations is obtained and Brouwer's method can be applied in principle. A complete survey of the large number of various possibilities is not even contained in the Kröger-Vink paper, only a few particular examples are discussed there.

The next step following logically in this work is the theoretical derivation of the influence of the impurities on the various *physical* properties mentioned earlier. Once "concentration charts" have been obtained, this is easy. The ionic conductivity and the diffusive properties of the compound are usually determined preponderantly by one kind of lattice defect (e.g. M-vacancies in the alkalihalides, see next section). From the chart, the change of concentration of *this* defect in the presence of impurities can be immediately derived. Moreover, it can be checked whether other defects, which ordinarily play no role of importance, may possibly have increased so much in concentration with respect to the former, that they cannot any longer be neglected. In this way various remarkable "irregularities" in the conductivity-temperature relation of ionic compounds have been quantitatively explained. Generally, only impurity atoms of deviating valency, which can exist as ions differing in charge from the ions they replace, influence the ionic conductivity in a notable way; ions of the same valency have no great effect, except on the mobility of charge carriers.

The regulation of the *electronic conductivity* by the incorporation of foreign atoms of deviating valency, thereby changing the number of charge carriers, was called *controlled valency* by VERWEY and others [1950]. A typical example is formed by the compound NiO in which some Li is incorporated. It is assumed that Ni is present as Ni^{2+}-ions, oxygen as O^{2-}-ions, and that Li^{+}-ions substitute for Ni^{2+}-ions. To compensate, Ni^{3+}-ions are then formed in equal numbers. Each Ni^{3+}-ion represents a free hole, unless it is within the attraction sphere of a Li^{+}-ion. Thus, hole conductivity is produced in the originally insulating compound.

In semiconducting compounds, the introduction of foreign atoms with deviating valency results inter alia in a change of the external (e.g. metal) vapour pressure at which the transition from n to p type occurs.

Both examples can be regarded as two aspects of the same process, namely the change in position of the *Fermi level* with the concentration of impurities. Atoms with more valence electrons than the normal number tend to rise the Fermi level, those with less, to lower it.

The method of controlled valency has been applied with success to a large number of compounds. Most of these are more or less heteropolar oxydes; another large group is formed by the intermetallic III–V and II–VI compounds.

Interesting effects arise from the influence of the presence of a second type of impurity on the solubility of the first. This is particularly pronounced in ionic crystals, where the introduction of a differently charged ion is greatly facilitated by the simultaneous insertion of an other ion with the opposite charge difference. For instance, the solubility of Ag in CdS is greatly enhanced by the simultaneous introduction of Cl. Ag enters as Ag^{+}-ions substituting Cd^{2+}-ions. Charge can be compensated by S^{2-}-vacancies, but only to a limited extent. When Cl^{-}-ions are present which can replace S^{2-}-ions, much more silver can be introduced. Again quantitative relations can be obtained from the solution of a set of simultaneous reaction equations. It is, however, no longer possible to construct a concentration chart in two dimensions only.

4. Diffusion and Ionic Conductivity

The diffusion of vacancies and interstitial atoms is described by a diffusion coefficient:

$$D = D_0 \exp^{-Q/kT} ; \tag{12}$$

as before D_0 contains, except for the usual term $\beta \nu_0 a^2$ where β is a geometrical factor (e.g. $^1/_{12}$ in the NaCl structure), an *entropy* factor. This may well differ from unity; in most polar materials it varies between 10 and 10^4, and is smaller when interstitial diffusion is concerned than when vacancy diffusion occurs (the difference may amount to a factor 100 or more; it arises from the contrary influence of interstitials and vacancies on the frequency of lattice vibrations). Diffusion

by the four types of simple point defects is described formally by the same formula, but the actual diffusivities are greatly different. To determine these diffusivities experimentally is usually easier than in metals, owing to the charged nature of the defects. Self-diffusion in a polar crystal gives rise to electrolytic conduction if the mobilities (velocities attained in a given electric field) of the differently charged moving ions are unequal. Since the electrical conductivity of polar crystals is predominantly ionic, its dependence on temperature, the phenomena occurring at cathode and anode, etc., thus yield indirect information on the diffusivity of the ions and thus of the defects (i.e., on the self-diffusion in the solid and on the nature of the defects responsible for this). As a diffusion process, the ionic conductivity increases very rapidly with increasing temperature, in strong contrast to (electronic) conduction in metals. The absolute magnitude of the conductivity is much lower in the former case, however.

Usually the temperature dependence of the conductivity is not characterized by a single energy of activation, and depends sensitively on purity. We shall discuss this in Chapter XXV.

It cannot be unambiguously decided from conductivity experiments alone whether the diffusion of the ions takes place by an interstitial or by a vacancy mechanism (i.e. by the motion of Frenkel or Schottky defects respectively). However, purely geometrical considerations (size of the ions, etc.) may give additional insight into the expected magnitudes of the activation energies for migration. Such considerations lead for instance to the expectation that the very large interstitial anions in the alkali- and silverhalides will not take part in the ionic conduction process. Actually, in AgCl the current seems to be carried predominantly by interstitial Ag^+-ions. In how far the Ag^+-vacancies also contribute to the conduction is not easy to determine, but Cl^--ions certainly do not take part in the process: in thermal equilibrium, as in all other silverhalides, only *Frenkel defects* are present. This stands in contrast to the situation in the alkalihalides, where *Schottky defects* play the dominating role. There, the positive-ion vacancy is much more mobile than the negative-ion vacancy, and *both* kinds of interstitial ions are normally absent. This difference between silverhalides and alkalihalides is due presumably to a difference in the homopolar component of the binding force between ions which is much stronger in the silverhalides than in the alkalihalides. The Ag^+-ion seeks to occupy a place where it is tetragonally surrounded by

halogen ions. Such places are just the interstitial tetrahedral holes in the AgCl lattice. A corresponding tendency to occupy interstitial sites is absent for Na^+-ions in NaCl.

A determination of the so-called *correlation factor* of the diffusivity can give new information on the nature of the diffusing defects. This factor arises in the following way. The diffusivity D of *tracer ions* and the ionic conductivity σ, due to a given defect mechanism, are interrelated by the so-called *Einstein relation* (see section XVIII–1):

$$\frac{D}{\sigma} = \frac{kT}{Ne^2}; \tag{13}$$

this can be derived easily by considering the balance that must exist in equilibrium between the forces exerted on the ions by the applied electric field and by the chemical gradient respectively; N denotes the number of diffusing ions per unit volume. In geometrically simple cases, for instance for direct interstitial migration, this equation holds exactly. However, in many other cases, an ion which has just left a lattice site has a greater chance to *return* to this, now open, site during

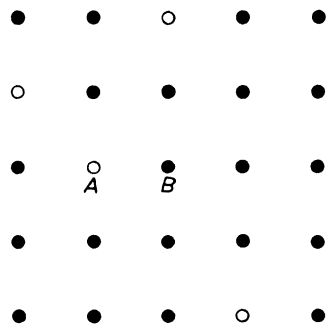


Fig. 6. An ion having arrived at B by jumping from the site A has a larger probability in its next jump to return to A than to one of the other seven places surrounding B, because A is *certainly* empty and whether the other sites are empty is only a matter of chance. Filled circles: occupied sites, empty circles: vacant sites.

its next jump than corresponds to a random walk probability (Fig. 6) To take account of this, a *correlation factor* f is introduced on the right-

hand side of (13), which is smaller than unity: a certain fraction of the jumps does *not* contribute to diffusion, but it does of course contribute to the electrolytic conductivity. The numerical value of the correlation factor depends on the lattice geometry and on the mechanism of diffusion. Therefore, by measuring accurately ionic conductivity and tracer diffusion together, one can gain valuable insight into the details of the diffusion process. We return to this in section XXV–3. We shall call equation (13) with the factor f included the "extended Einstein relation".

Some measured values of activation energies for migration of the various defects are also contained in Table 1.

The problem of ionic diffusion in polar crystals is exceedingly complex by reason of the simultaneous existence of the various possible mechanisms. (In metals, at least in the common metals, only vacancies play a role in diffusion.) In this book (next chapter) we can do no more than to point out the essentialities of the problem. We refer for a more thorough discussion of the fundamentals to the books of Mott and Littleton, and of STASIW [1959].

5. Dislocations in Polar Crystals

We now turn to the characteristics of dislocations in polar crystals. Here again, the condition of electrical neutrality affects in a peculiar manner dislocation geometry and properties. The structure of an edge dislocation (or a dislocation with at least some edge character) is no longer only determined by the criterion of least elastic energy, but rather by the much more stringent condition that the "additional half-plane" shall not disturb the charge neutrality in the crystal. In the *sodiumchloride* structure, which is the structure we shall discuss here most extensively, the shortest lattice vector connecting ions of the same sign is along a ⟨110⟩-direction; its length is $a\sqrt{2}$, when a is the cube edge. As will be discussed in Chapter XXVI, plastic glide in the alkalihalides (at not too high temperatures) takes place on the {110}-planes. A simple edge dislocation having such a glide plane, namely (011), lying along [100] and with a Burgers vector along $[0\bar{1}1]$ is depicted in Fig. 7. Its "extra half-plane" can be regarded as consisting of *two* $(0\bar{1}1)$-planes perpendicular to the (011)-glide plane. These two planes intersect in Fig. 7 with the (100)-plane of drawing along two rows, of negative and positive ions respectively. In this way charge

neutrality is ensured. Of course, this choice of the extra half-planes is not unambiguous, one of the (010)- or (001)-planes indicated by c and c' can also be regarded as such. These planes consist of rows of

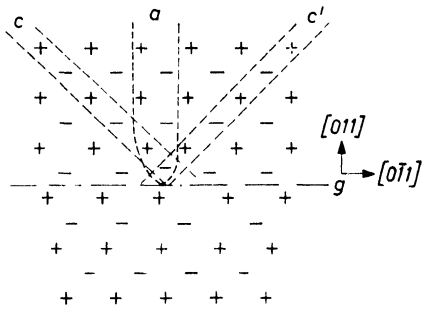


Fig. 7. Ionic structure in (100)-plane perpendicular to an edge dislocation along [100], with Burgers vector along [0$\bar{1}$1], in the sodium chloride lattice. The glide plane g is (011). As "extra half-plane" can be considered the *two* (0$\bar{1}$1)-planes indicated by a, or either one of the (010) and (001)-planes c and c'.

ions of alternate sign parallel to the plane of drawing, and need thus not be doubled.

The two (0$\bar{1}$1)-extra half-planes do not end at equal heights above the glide plane. Therefore, it is possible to distinguish a row of "lower

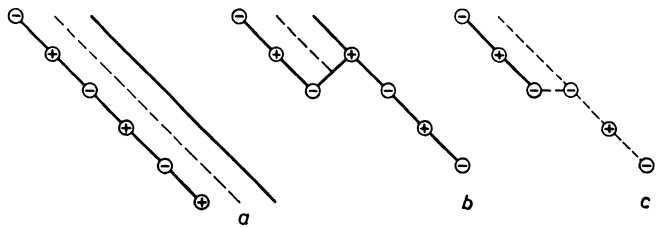


Fig. 8. *a*) Row of lower ions along the axis of the dislocation of Fig. 7. *b*) Neutral jog in the dislocation. The alternation of + and — charges is not disturbed. *c*) Charged jog or incipient vacancy.

ions" along an edge dislocation line (cube edge direction), that consists of ions of alternate sign. Fig. 8*a* depicts this situation (in the system of Fig. 7). A jog in such a dislocation means a step in the row. When this jog is caused by the intersection with another edge dislocation

of the same type, with a Burgers vector along [011], the step changes the position of the first dislocation line by two consecutive (011)-planes (Fig. 8*b*) and no break in the alternation of + and — ions along the dislocation occurs: one speaks of a *neutral* jog. However, a jog consisting of a step of only one (011)-plane in height, produces the configuration of Fig. 8*c*, in which two "lower ions" of the *same* sign (negative in our example) appear next to each other. This configuration (charged jog) was called an *incipient vacancy* by SEITZ [1954], because of its ability to absorb an ion of the opposite charge. To affect this absorption, the jog has only to move sideways by one ionic distance along the dislocation line. It then becomes an incipient vacancy of the opposite sign. This process is favoured by the circumstance that the jog bears an effective charge, which is half the ionic charge of the bordering ions (thus half the charge of the negative ions in our example). This follows immediately, since the charge of the jog changes in sign but not in magnitude by the absorption of one appropriate ionic charge. Thus in the example of Fig. 8*c*, the charge of the incipient vacancy changes its sign from negative to positive (but not its magnitude) by the addition of a full positive-ion charge; it must thus originally have carried half a negative-ion charge. This charged nature of the jogs determines to a large part the properties of dislocations in polar crystals. It is improbable (but not impossible) that charged jogs are generated by dislocation intersection, at least if only dislocations with the described Burgers vectors exist. Such jogs will be certainly present in thermal equilibrium, however, being formed either by a diffusion process (SEEGER [1956]), that is, by "evaporation" of atoms from the dislocations, or by the dissociation of neutral jogs, which might reduce the elastic energy of the system. Charged jogs cannot exist in pure screw dislocations.

Charged jogs interact with the (equally charged) vacancies and interstitial ions in the polar crystals, much more so than the neutral jogs or unjogged dislocations. Moreover, charged jogs can interact with electrons and holes. We shall return to this in the following chapters, but note here already that dislocations will influence in this way the local thermal equilibrium of the point defects, by giving rise to space charge effects (section XXV–2).

The interaction between edge dislocations and point defects must be expected to be exceptionally strong, because apart from the strong elastic interaction due to the large elastic deformations around

vacancies, and similarly large dislocation strains (large Burgers vectors) respectively, the above mentioned electrostatic interaction must be considered.

The occurrence of partial dislocations is improbable in the sodium chloride structure, if not impossible. Burgers vectors which are not simple lattice vectors generally connect ions of opposite sign, and a system of two partials of this kind necessitates a stacking fault with an abnormally high stacking fault energy, owing to electrostatic forces. In this respect it is interesting that mechanical twinning in alkali halides has not been observed under normal circumstances.

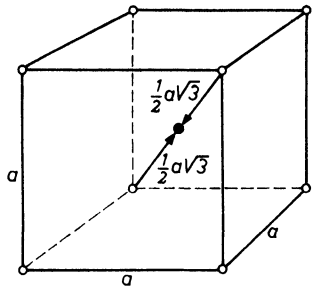


Fig. 9. A dislocation with Burgers vector $a\sqrt{3}$ along the space diagonal of the cubic unit cell of the caesium chloride lattice can easily be split into two half dislocations with Burgers vectors $\frac{1}{2}a\sqrt{3}$.

Of course, pure screw dislocations have none of the specific electrostatic properties mentioned above. However, the motion of screws can result in the production of (charged) defects, which again brings up electrostatic influences.

In the *caesium chloride* structure the electrostatical influences are less conspicuous than in the sodium chloride structure. The simplest dislocations imaginable in the former structure have a Burgers vector along ⟨100⟩, of length a (the cube edge). These much resemble the ⟨100⟩-dislocations in the NaCl structure. For example, edge dislocations (along another ⟨100⟩) are again characterized by a pair of extra half-planes, each of which now ends, however, in a row of ions of the same sign (but different for the two planes). Charged jogs as well as neutral jogs can exist in these dislocations. However, other dislocations, with Burgers vector along a ⟨111⟩-direction are also conceivable. Although this vector has a larger magnitude, ($a\sqrt{3}$), the

dislocation may well split into two partials with Burgers vectors $\frac{1}{2}a$ ⟨111⟩ each (Fig. 9). The stacking fault formed of course again represents a large electrostatic energy, but less than in the case of NaCl, because of the smaller difference in distance between ions of opposite and of equal signs respectively.

We shall not discuss here any other crystal structures, since the variation of structures among polar crystals is exceedingly large and completeness unattainable. A beautiful survey has recently been made by AMELINCKX [1958]. We refer the reader to this and to the original papers that now and again appear on specific crystals. The extensive paper by KRONBERG [1957] on dislocations in sapphire (hexagonal Al_2O_3) deserves special mention here, as it gives some interesting views on the various possibilities of more than two-fold partial splitting of dislocations in these structures.

6. Electronic Conduction and Colour Centres

Owing to their charged nature, vacancies and interstitial ions interact with the outer electrons of the constituting ions of the crystal. The electrons are in many instances rather mobile, which results in a finite electronic conductivity. In the ionic crystals we are concerned with (alkali- and silverhalides), the positive metal ions (alkalimetals or silver) are not easily ionized twice and therefore have no electrons available to give up to the crystal as a whole. On the other hand, the negative halide ions acquired an extra electron from the metal ions, and it is these electrons that are mobile. Such an electron can, for instance under the influence of thermal vibrations, exposure to optical radiation, etc., be transferred back to a positive ion, converting this to a neutral atom. The electron can jump fairly easily from this atom to a neighbouring positive ion, leaving the former again positively charged and making the latter neutral, etc. It will interact but slowly with the negatively charged, and therefore repulsive, halide ions, although it may ultimately be trapped by the neutral halide atom it came from, or by another neutral halide atom. These moving electrons constitute a means for transport of electricity, and can thus carry a current. The neutral halide atoms from which the electrons just considered originated, can in their turn accept an outer shell electron from a neighbouring halide ion, and thus become negatively charged again. The latter ion has then become neutral, of course. This process

can be repeated, and can then be considered as the migration of a *defect electron* or *hole* through the sublattice formed by the negative ions. Such a hole is formally equivalent to an electron except for its charge, which must be taken as positive. As electron migration, hole migration also constitutes transport of charge. Obviously, one can distinguish between electron (n-type) and hole (p-type) conductivity in these crystals. Usually one of the mechanisms dominates over the other.

Of course, both conductivities are associated with an *energy of activation*, viz. the energy required to remove an electron from a negative ion and place it on a positive ion. This usually amounts to at least an electron volt in the crystals under consideration.

In terms of the band theory of solids, the lower band is, at low temperatures and in the absence of radiation, completely filled by the outer electrons of the negative ions. A hole in this lower band corresponds to a "neutral negative ion", an electron in the conduction band corresponds to a "neutral positive ion". The energy distance between the band edges is to be considered as the energy of formation of an electron-hole pair.

The foregoing schematic description applies with minor alterations also to polar crystals consisting of multivalued ions, or having more than two constituents. We refer for a full description to the wellknown book by MOTT and GURNEY [1946]. The large mobility of electrons and holes (once excited), arises from the *periodicity* of the electrostatic potential in the lattice. In a purely periodic potential electrons are perfectly mobile, when their energy is confined to certain energy bands. If no thermal vibrations existed the electronic mobility would, as in a perfect metal, be infinite. The difference between metals and ionic crystals lies, of course, primarily in the *number of carriers* available, or, otherwise stated, in the magnitude of the band gap. The latter is absent in metals, and may easily amount to a few eV or more in a polar crystal, corresponding to the high values of the ionization energies of the ions under consideration.

Thermal vibration results in the existence of a mean free path for the current carriers in an ionic crystal, which decreases very rapidly when the temperature is raised. It is of the order of one ionic distance at the Debye temperature (near roomtemperature or higher) and increases to many tens of ionic distances at the temperature of liquid air. Not only thermal vibrations, but also impurity ions and lattice

defects reduce the mean free path of the carriers. This normal scattering (*free-free* transition) must be distinguished from the influence exerted by defects and impurities, which consists in *trapping* of the carriers (*free-bound* transition). The carriers then occupy a bound state in the field of the defect and are lost to the conduction process, until they are released again (e.g. by thermal activation). Trapping is possible only when the perturbation potential due to the defect varies sufficiently rapidly compared to the periodicity of the crystal field. Usually, the slightly screened electrostatic field at the site of a point defect (or impurity; we shall in the following not discuss specific impurity trapping) obeys this condition. We then speak of a *centre*. An electron trapped in a centre describes a closed orbit around it, the radius of which is determined by the attractive force and by the crystal properties as a whole (dielectric constant) respectively; it may well extend over many ionic spacings. In this way, the centre can also be considered as consisting of a lattice defect and an electron which is *shared* by the neighbouring ions of opposite sign. Conversely, the defect may also favour the presence of a defect electron on one of its nearest neighbours. In that case one speaks of a trapped hole. In terms of the band theory, a centre possesses one or more localized energy levels lying below the conduction band. In the purely ionic crystals all (charged) point defects

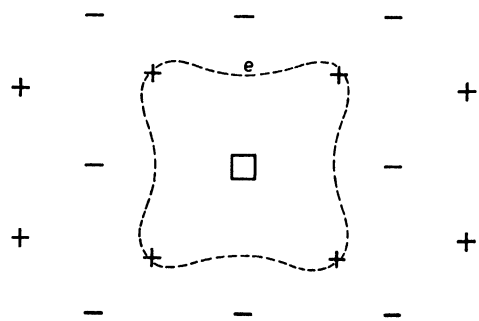


Fig. 10. Schematic picture of a centre (F-centre) consisting of a negative-ion vacancy that has trapped an electron *e*. The latter describes a fairly wide orbit around the former and can more or less be considered as being shared by all surrounding neighbouring positive ions (12 in the actual NaCl-structure).

can become centres; in metals, on the contrary, owing to the strong screening effect of the conduction electrons, bound states do not often occur.

Electrons will be trapped primarily at a negative-ion vacancy. Then a very common type of centre is formed, which is usually denoted as the *F-centre*. An F-centre can be regarded as a negative-ion vacancy and an electron that is equally shared by the positive ions surrounding this vacancy (Fig. 10). The electron can also interact with a positive interstitial ion. Conversely, a hole may be trapped at a positive-ion vacancy or at a negative (interstitial) ion, in which case one speaks of a V- and H-centres. Sometimes electrons and holes mutually interact so strongly in ionic crystals that they form themselves already a stable configuration, called an *exciton*. Electrons and holes may, of course, also be bound to more complex lattice defects such as vacancy pairs, and often a defect can trap more than one electron or hole. Finally, centres occur in which the trapping agency is not a physical defect, but for instance a divalent impurity in a monovalent crystal. It is therefore not surprising that a great many types of centres exist and have been observed in polar crystals. We shall return to this in Chapter XXVI.

REFERENCES

AMELINCKX, S. (1958); Suppl. al Nuovo Cimento No. **2**, p. 569.
ASCHNER, J. F. (1954); Thesis Univ. Illinois.
BROUWER, G. (1954); Philips Res. Rep. **9**, 366.
COMPAAN, K. and Y. HAVEN (1956); Trans. Faraday Soc. **52**, 786.
COMPTON, W. D. and R. J. MAURER (1956); Phys. Chem. Solids **1**, 191.
ETZEL, H. W. and R. J. MAURER (1950); J. chem. Phys. **18**, 1003.
FUMI, F. G. and M. P. TOSI (1957); Trans. Faraday Soc. **53**, 191.
FRIAUF, R. J. (1957); Phys. Rev. **105**, 843.
GRÜNDIG, H. (1960); Z. Physik **158**, 577.
HARRISON, L. (1958); Trans. Faraday Soc. **54**, 106.
HAVEN, Y. (1950); Rec. Trav. Chim. **69**, 1259.
HAVEN, Y. (1958); Philips techn. Rev. **20**, 69.
KELTING, H. and H. WITT (1949); Z. Physik **126**, 697.
KRÖGER, F. A. and H. J. VINK (1955); Solid State Physics **3**, 307.
KRÖGER, F. A. and H. J. VINK (1958); Halbleiter und Phosphore, p. 17 (Vieweg, Braunschweig).
KRONBERG, M. L. (1957); Acta Metall. **5**, 507.
KURUSOWA, T. (1958); J. Phys. Soc. Japan **13**, 153.
MAPOTHER, D., H. N. CROOKS and R. J. MAURER (1950); J. Chem. Phys. **18**, 1231.
MOTT, N. F. and R. W. GURNEY (1946); Electronic processes in ionic crystals, 2nd edition (Oxford, Clarendon Press), Chapters I and II.
MOTT, N. F. and M. J. LITTLETON (1938); Trans. Faraday Soc. **34**, 485.
PRATT, P. L. (1954); Report Conference on Defects in Solids, Bristol, p. 402 (London, The Physical Society).
PRATT, P. L. (1958); Vacancies and other Point Defects (London, Inst. of Metals).
SCHAMP, H. W. and E. KATZ (1954); Phys. Rev. **94**, 828.
SEEGER, A. (1956); Encyclopedia of Physics, Part VII–1 (Springer, Berlin).
SEITZ, F. (1946); Rev. Mod. Physics **18**, 384.
SEITZ, F. (1954); Rev. Mod. Physics **26**, 7.
STASIW, O. (1959); Elektronen und Ionenprozesse in Ionenkristallen (Springer, Berlin).
TELTOW, J. (1949); Ann. Physik **5**, 63 and 71.
VERWEY, E. J. W., P. W. HAAYMAN, F. C. ROMEYN and G. W. VAN OOSTERHOUT (1950); Philips Res. Rep. **5**, 173.
WAGNER, J. and P. HANTELMANN (1950); J. chem. Phys. **18**, 72.

CHAPTER XXV

LATTICE DEFECTS AND THE ELECTRICAL PROPERTIES OF IONIC CRYSTALS

1. Intrinsic Ionic Conductivity

The most important property of polar crystals that is determined principally by lattice defects is the ionic conductivity. Owing to the small number of conducting electrons or holes (due to the large energy of formation of a pair) in polar crystals the electricity is carried practically exclusively by ions. The conductivity is therefore smaller by many powers of ten in these crystals than in metals, that is, ionic crystals are essentially insulators.

The jump frequency of a defect can be formally expressed as

$$\nu = \nu_0 \, e^{-Q/kT} , \tag{1}$$

where Q is a free energy of activation, that may be temperature dependent. When an electric field is present, the probability for jumping of an ion varies according to the direction of the jump. For the simplest case, that of a positive interstitial ion with charge e moving in an isotropic lattice, energy is *gained* in a jump *in* the direction of the applied field F (forward jump); such jumps occur therefore at the increased rate:

$$\nu' = \nu_0 \, e^{-\frac{Q - 1/2 \, e \, a \, F}{kT}} , \tag{2a}$$

whereas jumps in the reverse direction occur at the *reduced* rate:

$$\nu'' = \nu_0 \, e^{-\frac{Q + 1/2 \, e \, a \, F}{kT}} \tag{2b}$$

Here a denotes the full lattice periodicity. From this it follows directly that in not too strong fields ($eaF \ll kT$) the current density

(net number of ions crossing unit area in the direction of the applied field), is given by:

$$j = nea(\nu' - \nu'') = \frac{ne^2a^2F}{kT}\nu_0 e^{-Q/kT} = \frac{ne^2a^2\nu F}{kT}, \tag{3}$$

where n is the number of migrating (interstitial) ions per unit volume. Often a quantity μ, named the *mobility*, is introduced to describe the electrical properties:

$$\sigma = ne\mu, \tag{4}$$

where σ is the conductivity. In our case where evidently $\sigma = na^2e^2\nu/kT$, one has

$$\mu = \frac{a^2e\nu}{kT}. \tag{5}$$

It is easily verified that, apart from a possible numerical factor between one and ten, similar formulae hold for other defect-conduction mechanisms. This numerical factor is entirely of a geometrical nature; it takes account of the various possible directions for jumping of a defect in an actual lattice.

Any circumstance that influences the concentration of defects or their mobility or both affects the ionic conductivity.

The most conspicuous variation has already been mentioned: increasing the *temperature* usually (but not always) results in an increase of n as well as of μ. The conductivity then increases rapidly with temperature; this increase is characterized by an energy of activation which is the sum of the formation energy U of a defect and its migration energy Q. Usually the formation, not of a single defect but of a *pair* of defects (Schottky or Frenkel pair) is given and then according to XXIV–(2), one has

$$\sigma \sim e^{-(\frac{1}{2}U+Q)/kT}, \tag{6}$$

if U denotes this pair energy.

Often it is found experimentally that the variation of conductivity with temperature is not characterized by a single activation energy, but is different at high temperatures and at low temperatures.

Among the many possible internal factors that influence the conductivity, those that interest us here are the presence of other lattice

imperfections and especially of the impurities just mentioned. We shall briefly discuss these influences now.

2. Dislocations and Ionic Conductivity

The presence of dislocations influences the ionic *mobility* only to a very small amount: because of their large mass, the moving ions are but slightly scattered when encountering a dislocation. On the other hand, dislocations can influence the *concentration* of lattice defects to a large extent, either by the dynamic moving-jog mechanism described in Chapter III, or by their activity as *static* generators of point defects. The first process can be fully compared to the generation of vacancies and interstitials by moving dislocations in metals: the formation of defects during plastic deformation follows entirely by geometrical necessity. In polar crystals, to preserve charge neutrality, jogs can only move in such a way that each produces positively and negatively charged defects at equal rates. That is, moving dislocations leave *Schottky defects* or "anti-Schottky defects" (pairs of oppositely charged interstitials) in their wake; Frenkel defects can not easily be formed in this way. When the neutrality condition is not fulfilled, the jog is practically immobile under stress. This imposes some restriction on the mobility of dislocations as a whole, as will be discussed in section XXVII–3.

A stationary dislocation influences by its mere presence the local concentration of point defects, because it can act as a source or sink of individual vacancies or interstitial ions, which are formed *independently* as has been explained in the previous chapter. Schottky or Frenkel defects can be formed at jogs (or, for that matter, also at the surface or at grain boundaries) and, since usually the energies of formation of the two oppositely charged components of such a defect pair are unequal, the *thermal equilibrium* concentrations according to XXIV–(1) are unequal too. Far from the dislocation (or from the surface) the difference in free energy of formation is exactly compensated by the electrostatic energy due to uncompensated charges, and, as we have seen before, the equilibrium concentrations of both components become both governed by half the energy of formation of the complete Frenkel or Schottky defect. However, near to the sources of the *individual* defects this is no longer so, the concentration of the (negative) component A being characterized by its own energy

U_A and that of the (positive) component B by $U_B \neq U_A$. Thus a space charge is formed and an electrostatic potential φ is set up, which influences in turn the free energy of formation of the two components. One obtains for the local equilibrium concentrations:

$$\begin{aligned} c_A &= \exp\{-(U_A - e\varphi)/kT\}, \\ c_B &= \exp\{-(U_B - e\varphi)/kT\}; \end{aligned} \tag{7}$$

where the potential φ depends on the space charge density $e(c_A - c_B)$ according to Poisson's equation:

$$\Delta\varphi = -\frac{4\pi e}{\varepsilon}(c_A - c_B), \tag{8}$$

ε being the dielectric constant. Inserting (7) into (8), one obtains a differential equation for φ which is identical with the *Debye-Hückel equation* governing the potential in an electrolyte. Its solution far away from the dislocation (or the surface) is, as has been mentioned earlier:

$$c_A = c_B = \exp\{-(U_A + U_B)/2kT\}, \tag{9}$$

but within a distance

$$\delta = \sqrt{\frac{\varepsilon kT}{8\pi e^2 n}}, \tag{10}$$

where n is the volume density of defects, the concentrations of A and B are different. Owing to the independent formation of individual point defects a dislocation in an ionic crystal thus becomes surrounded by a space charge region of width δ. The dislocation itself acts as a line charge with potential

$$\psi = \frac{U_A - U_B}{2e}, \tag{11}$$

which according to the known data amounts to -0.28 V in NaCl. The width δ of the space charge cylinder (and of the analogous surface layer, which is of course present as well) varies from about 10^{-6} cm at 900 °K to a few times 10^{-5} cm at 600 °K in this material.

PRATT [1958] discussed the possible theoretical consequences of a charge on the dislocations and the surface on conductivity and mechan-

ical properties. AMELINCKX *et al.* [1959] have confirmed the charged nature of dislocations by the observation of an electric alternating voltage produced by cyclic bending of a NaCl crystal.

BASSANI and THOMSON [1956] have drawn attention to another possibility for a dislocation to acquire a charge, viz. by a difference in *interaction* with oppositely charged point defects. Positive-ion vacancies may for instance be present in large concentrations as the dissociation products of pairs of divalent impurities and vacancies (see below). When these are attracted by a dislocation, this acquires an excess negative charge, which may not be compensated by a similar interaction with positively charged defects. This effect will be of special importance in impure crystals, where the concentration of defects is enhanced over the equilibrium value.

The generation of point defects by moving dislocations leads to very large changes in the electrolytic conductivity. This is not surprising, since after a given amount of plastic strain, the number of Schottky defects or anti-Schottky defects formed can be estimated to be of the same order as in metals, that is about 10^{18} to 10^{19} defects per cm^3 after 10% strain. These numbers are larger by several orders of magnitude than the usual thermal equilibrium concentrations of the common defects at room temperature. As the new defects can usually contribute to the transport of electricity as well as the old ones (if their mobility is large enough) it is not surprising that plastic deformation of ionic crystals usually results in an *enhancement* of the ionic conductivity. This effect of deformation, which contrasts strikingly with the influence of plastic strain on metallic (electronic) conductivity, was first experimentally observed by GYULAI and HARTLY as early as 1928 and has since been confirmed by many others. The conductivity of NaCl crystals at room temperature can be increased several hundred fold after moderate compressive deformation. The point defects that are responsible for this increase apparently anneal out at room temperature in about 20 minutes, since the increase of conductivity is only transient, with a decay time of this order of magnitude. More recent experiments (CAFFYN and GOODFELLOW [1955]) indicated that the effects of plastic deformation were extremely irreproducible.

SEITZ [1954] has interpreted the effect as being due to the formation of simple Schottky defects by moving dislocations, the positive-ion vacancies of which are sufficiently mobile in alkali halides at room

temperature to contribute markedly to the conductivity. The annealing process would then be the association of pairs of vacancies of opposite sign (resulting in an uncharged defect), again brought about by the migration of the positive-ion vacancies. The negative-ion vacancies are practically immobile at room temperature in NaCl.

In contrast to these and other observations, FISHBACH and NOWICK [1956, 1958] observed the *low*-temperature conductivity of NaCl, measured at −60 °C (thus in the impurity range, see next section) to be *decreased* by plastic deformation and increased again by subsequent annealing. The temperature dependence of the conductivity was not affected (the activation energy, see section 1, remained 1.07 eV, irregardless of the treatment). The effect is too large to be explained by the scattering of moving vacancies by dislocations. The explanation proposed by the authors was based on the large influence of accidentally present divalent impurity ions (Ca^{++}-ions) in the crystal on this low-temperature electric conductivity (see next section). These ions can be considered as sources of (mobile) vacancies. It was assumed that plastic deformation resulted in the precipitation of the Ca^{++}-ions, which then lost their charge. The vacancy concentration, only present to neutralize this charge, should drop accordingly, and the conductivity with it.

The same authors also observed [1958] a flow of charge under stress in deformed NaCl without an electric field being present. They interpreted this in terms of the motion of charged dislocations.

3. Impurity Conduction

Impurities of deviating charge affect to a very great extent the ionic conductivity; as has already been mentioned this is due to their influence on the equilibrium defect concentration (at not too high temperatures, where the intrinsic concentration is still small). To preserve charge neutrality, a given concentration of divalent impurity ions in a monovalent crystal must necessarily be accompanied by an equal concentration of oppositely charged defects. The defect that most easily performs the role of neutralizer in the alkali halides is the (positive-ion) vacancy. Studied most extensively is the effect of divalent metal ions (Ca^{++}, Sr^{++}, Cd^{++}, Mg^{++}) on the ionic conductivity of NaCl and KCl. Fig. 1 represents some of the results. The ionic conductivity at low temperatures is enhanced above the "pure" value

by an amount proportional to the concentration of divalent ions. Simultaneously, the temperature dependence has become less steep.

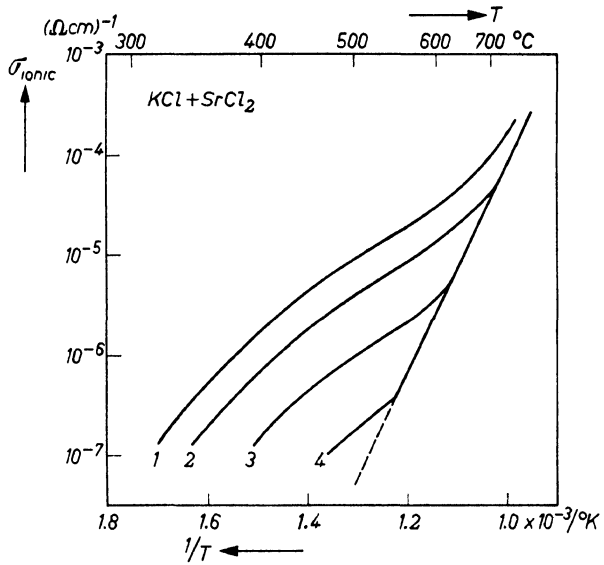


Fig. 1. Ionic conductivity σ of KCl + certain amounts of $SrCl_2$ as impurity, as a function of temperature (after KELTING and WITT [1949]). Curve *1*: 19 × 10^{-3} mol% $SrCl_2$; *2* : 6.1 × 10^{-3} mol% $SrCl_2$; *3* : 1.9 × 10^{-3} mol% $SrCl_2$; *4* : ≪ 10^{-3} mol% $SrCl_2$. The larger the impurity content, the higher the transition temperature between the intrinsic and the extrinsic (impurity) region of conductivity.

At high temperatures the effect of the impurities becomes negligible. (Indications—but no more—exist as to another "knee" at very low temperatures, the origin of which is not clear.) The influence of impurities on the ionic conductivity has been carefully reviewed by LIDIARD [1957] and we shall limit ourselves here to the most important aspects only. It is assumed that to preserve charge equilibrium the incorporation of each divalent impurity ion is accompanied by the introduction of one positive ion vacancy, which is free to move through the crystal, not being properly *bound* to the impurity ion. At low temperatures, the equilibrium concentration of vacancies in a pure crystal is very much smaller than the impurity concentration. The latter then determines the vacancy concentration completely (in an impure crystal), and consequently the conductivity is proportional to the impurity content. The temperature dependence of the conductivity

is governed only by the activation energy for migration of the vacancies, thus only by the activation energy Q of the mobility. When the temperature increases, the thermal equilibrium concentration of vacancies becomes comparable to the impurity concentration and ultimately surpasses it. Then the conductivity is determined primarily by this thermal equilibrium concentration; this *intrinsic* conductivity depends on temperature according to formula (6). The temperature dependence is much steeper, as shown in the figure, since the activation energy is now the sum of the energy of migration Q and of (half) the energy of formation of Schottky defects U.

By careful measurement of the slopes of the two parts of the conductivity—temperature curve of an impure crystal, separate information can thus be gained on the energies of migration and of formation of defects.

A certain energy of *association* between the impurity and its accompanying vacancy exists, which results inter alia in a slight deviation from the proportionality between *low-temperature* (*extrinsic*) conductivity and impurity content (ETZEL and MAURER [1950]). From the temperature variation of these deviations the association energy can be estimated: 0.4 eV for Cd^{++}-ions in NaCl but less than 0.08 eV for Ca^{++}-ions (Bean, unpublished). Theoretical calculations (BASSANI and FUMI [1954]) yielded 0.44 and 0.38 eV respectively in NaCl. The associated positive-ion vacancy, if situated on one of the 12 next-nearest neighbouring sites of the impurity (the nearest sites are occupied by negatively charged ions) forms a so-called *complex* with the impurity ion (SEITZ [1954]).

In rocksalt containing a few per cent $CaCl_2$ as impurity, the complex of a Na^+-ion vacancy together with a Ca^{2+}-ion is characterized by an *electric dipole* moment. In an *alternating* electric field, such dipoles contribute to the electric displacement current by an effectively complex conductivity: the current due to the vibration of the dipoles lags behind the applied alternating voltage. This leads to energy dissipation, thus to dielectric losses (for a description of the theory of this phenomenon, see the textbooks and for instance FRÖHLICH [1949]). The relation between phase lag and energy dissipation has been discussed in Chapter XVII. In the terminology of that chapter:

$$\tan \delta_a = \frac{16\pi n_a p^2}{3\varepsilon kT} \frac{\omega\tau}{1 + \omega^2\tau^2} \tag{12}$$

where n_s is the number of dipoles (associated Na^+-vacancy complexes) per cm^3 and p is their dipole moment, which is equal to the charge e times the distance d between vacancy and Ca^{++}-ion. This distance is a face diagonal of the elementary cubic lattice cell (Fig. 2). The quantity ε measures the real part of the dielectric constant and ω is the frequency. τ is the *relaxation* time of the complex: $1/\tau$ is 2 times the mean frequency with which a neighbouring Na^+-ion jumps into the vacancy and back, in such a way that the latter *remains* a neighbour of the Ca^{2+}-ion (Fig. 2)*. Since there are two equivalent ways in which

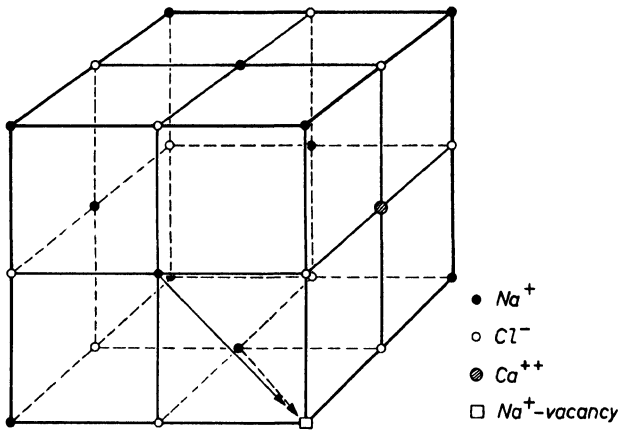


Fig. 2. Two possibilities by which sodium ions can fill a sodium ion vacancy, neighbour to a Ca^{++}-ion, thereby leaving another vacancy with the same property.

a vacancy can jump, the factor 2 has to be introduced. The losses reach a maximum for $\omega = 1/\tau$; this frequency-dependent dielectric loss is depicted in Fig. 3.

Free sodium ions contribute to the *direct* current conductivity, as has been discussed before. Since in each jump a Na^+-ion moves the distance d, the d.c. conductivity is given by:

$$\sigma_f = \frac{4n_f e^2 d^2}{kT\tau_d}. \tag{13}$$

In this formula n_f is the number of free vacancies per unit volume

* Also the Ca^{2+}-ion itself can jump into the vacancy; however, owing to its double charge, this ion is generally considered to have a much lower jumping frequency, and its motion is neglected.

and τ_d the jumping frequency of such a free vacancy; d is, as in the former expression, the length of the face diagonal of the elementary cell. This frequency-independent d.c. conductivity of course also gives rise to dielectric losses in an alternating field, but in contrast to the former case, the "conductivity" losses are simply proportional to the number of cycles per unit time, that is to the frequency.

Both losses together give rise to the solid curve of Fig. 3. From the

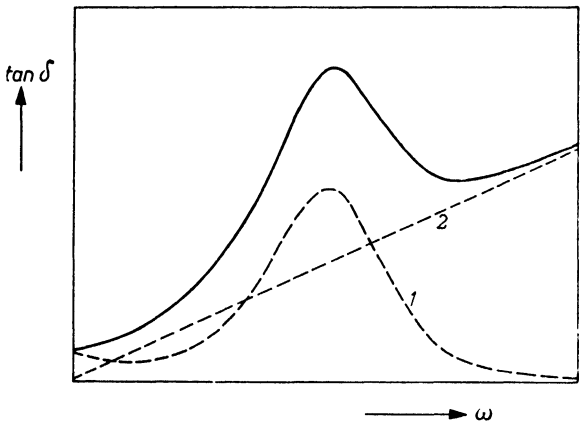


Fig. 3. Curve *1*: Frequency dependence of dielectric loss angle due to the relaxation of associated vacancies in impure sodium chloride in an alternating electric field. The damping is maximum for $\omega = 1/\tau$, where $1/\tau$ is the mean jumping frequency of such a vacancy. Curve *2*: The direct current losses due to the migration of single vacancies are proportional to the frequency. Solid curve: The total dielectric loss angle of an impure sample of sodium chloride (schematic).

frequency where the peak occurs and from the low frequency conductivity (left branch) respectively, n_a and n_d can be determined. By measuring the loss angle as a function of frequency at different temperatures, information on the energy for association of a complex can now be obtained. Of course, what determines the loss angle is always the *product* of concentration and jump frequency, and it is only *this* quantity that can actually be derived, for the free vacancy and the complex respectively. The relaxation time also varies with temperature according to an exponential law

$$\tau = \tau_0 \, e^{+Q/kT} , \tag{14}$$

where Q is again the activation energy of the mobility of the defects.

This can be determined by measuring the conductivity in the low-temperature impurity region, where the degree of association is 100%, or in the high temperature region, where it is essentially zero.

From an analysis, along the lines indicated, of the observed conductivity-temperature-frequency relations in NaCl HAVEN [1954] found the jump frequency of an associated vacancy to be about three times larger than that of a free vacancy at room temperature. That is, a Na^+-ion jumps more easily into an associated vacancy than into a free vacancy. This difference should be due to the electrostatic effect of the neighbouring doubly charged Ca^{2+}-ion. Since the *activation* energies of both jump frequencies were equal within the errors of measurement, it is in the vibration frequency of the Na^+-ion near the vacancy that the difference must be sought.

The degree of association between vacancies and Ca^{++}-ions could not, according to this same author, be very well described by a single activation energy, nor was there close agreement between the fractional concentrations of associated vacancies and free vacancies and the total Ca^{2+}-ion concentration. At 100 °C and 0.3 mol% Ca, only 0.003 mol% free vacancies and 0.07 mol% associated vacancies were present. This, according to Haven, points to the occurrence of the majority of the remaining defects, either in the form of higher complexes or as a second phase. At high temperatures the association equilibrium is further perturbed by the influence of a second equilibrium reaction:

$$Na^+\text{-vacancy} + Cl^-\text{-vacancy} \rightleftarrows \text{complex}$$

which only takes place when the thermal equilibrium concentration of Cl^--vacancies has reached a sufficiently high level.

Combined conductivity and after-effect measurements on impure polar crystals have been carried out by a great number of investigators (see BRECKENRIDGE [1952], MAPOTHER *et al.* [1950]) and we refer for a further discussion to LIDIARD's review paper [1957], in which also extensive theoretical considerations are contained (see also next section).

4. Diffusion

A simultaneous measurement of the ionic conductivity σ and the diffusivity D of the ion (or rather the defect) responsible for it, makes

possible an investigation to the validity of the "extended Einstein equation":

$$D = f \frac{kT}{Ne^2} \sigma , \tag{15}$$

where f is the correlation factor (see previous chapter). The diffusivity of the ions must be measured under conditions most resembling those of the stationary state, that is, in the absence of chemical activity gradients. This can be realized by making use of *radioactive isotopes* (*tracers*). In the absence of complicating factors, and when only one kind of ion takes part in the diffusion process, equation (15) should hold exactly for tracer diffusion, and by comparing measured and computed values of the diffusivity the correlation factor can be numerically determined. Since this factor can also be evaluated theoretically for the various defect diffusion mechanisms, information about the actual mechanism can be obtained in this way.

However, complications do usually occur. When more than one constituent takes part in the conductivity process, only that fraction of σ that corresponds to the contribution by the tracer in question should be used in the comparison, and an additional factor $\alpha \leqslant 1$, called the *transport number* of the ion, must be inserted on the left hand side of (15). LIDIARD [1957] gives a list of observed transport numbers in a variety of crystals. Even when this correction is duly taken into account, deviations between the measured tracer diffusivity and that computed from the conductivity are usually observed in practice. These are in many cases due to the fact that also uncharged defects such as vacancy pairs or associated complexes take part in the diffusion process, which of course do not contribute to the conductivity, since their transport numbers are necessarily zero.

Most investigations have again been carried out on the alkali-halides and the silver halides. Of the latter group, the work of COMPTON and MAURER [1956] should be mentioned here. The conductivity and diffusivity data of silver ions differ at high temperature although the activation energies are practically equal. In the low-temperature (impurity) region, where the vacancy concentration is determined solely by the divalent impurities present, and the activation energy determining the conductivity-temperature relationship is correspondingly small, the differences are small and again temperature independent; they indicate a correlation factor of 0.8, characteristic for

a vacancy mechanism in the silverchloride lattice. Since the transport number of silver ions is unity, and no neutral pairs are formed (these should *enhance* the measured diffusivity over the computed value instead of *lowering* it, as observed), the discrepancy at high temperature must be found in a much smaller correlation factor, the magnitude of which (about 0.3) can only be explained by assuming that an *interstitialcy* mechanism (section XVIII–2) describes the migration of Ag^+-ions (at least partly) at these temperatures. Theoretically, $f = \frac{1}{3}$ for this mechanism in the silver chloride lattice. For a detailed discussion the reader is referred to the original paper. Evidence for interstitialcy migration in AgCl has also been advanced on theoretical grounds by HOVE [1956], whereas new experimental evidence by FRIAUF [1957] is also in accord with this assumption. According to Hove's calculations, ordinary interstitial migration can not occur, as the energy of activation for this is very high. The activation energy for the interstitialcy mechanism should lie theoretically between 0 and 0.8 eV, in fair agreement with the measured value of the activation energy of intrinsic conduction in AgCl (0.87 eV).

In the alkali halides the comparison between the measured tracer-diffusion coefficient of Na^+-ions and the diffusion coefficient computed from conductivity measurements shows discrepancies that are most conspicuous in the structure-sensitive, low-temperature region. There, the measured diffusivity is too high, which was interpreted by MAPOTHER *et al.* [1950] in terms of the simultaneous migration of uncharged associated complexes between Na^+-vacancies and the ever present divalent Ca^{2+}-ions. The latter have nearly the same ionic radius as the Na^+-ions, and are therefore easily soluble in NaCl. Dielectric loss measurements (see previous section) also point to the presence of such complexes, although the numerical agreement between the two kinds of experiments, as far as the number of complexes is concerned, is not good.

Similar observations were made by WITT [1953] on the diffusion of K^+-ions in KCl containing strontium as an impurity. Here again K^+-vacancy/Sr^{++}-ion complexes are responsible for a large part of the observed diffusivity.

LIDIARD [1957] has studied theoretically the association equilibrium between divalent impurities and vacancies in sodium chloride. Since there are 12 positions to be occupied by cation vacancy next to a divalent positive ion (compare Fig. 2), the equilibrium concentration of

associated pairs is described by the equation

$$\frac{pc}{(1-p)^2} = 12 \exp\left\{\frac{\Delta G}{kT}\right\}. \tag{16}$$

Here ΔG is the free energy of association of an associated pair, c the concentration of impurity ions and p the degree of association of these ions with Na^+-vacancies. When a concentration gradient is present the pairs give rise to a diffusion current. If it is assumed that the vacancies do not impede the motion of the complexes, that is, if the jumping frequency of the vacancy, ν_v, is higher than that of the associated impurity ion ν_i, the diffusion coefficient of complexes is given by

$$D = \frac{1}{3} a^2 \nu_i \frac{d}{dc}(pc) . \tag{17}$$

The factor $\frac{1}{3}$ must be included for geometrical reasons; a is the lattice parameter. If the vacancy itself moves as fast as, or more slowly than the ion, then in this expression ν_i must be replaced by the reduced frequency

$$\frac{\nu_i \nu_v}{\nu_i + \nu_v} .$$

Hence, D depends on the impurity concentration. Inserting (16) into (17), it is found that

$$D(c) = \frac{a^2}{3} \frac{\nu_i \nu_v}{\nu_i + \nu_v} \left\{1 - \left[1 + 48c \exp \frac{\Delta G}{kT}\right]^{-\frac{1}{2}}\right\}. \tag{18}$$

A measurement of D thus yields information on the degree of association p.

The derivation of (18) is based on the assumption that the lifetime of complexes is long compared to their mean jumping time. If this is not so, less simple formulae are obtained which are of little practical use. In NaCl containing calcium the above approximation is, however, indeed applicable.

Also in metals the diffusion of complexes of this kind can be of importance (Johnson effect), but not in the noble metals, since there the energy of association of a divalent impurity and a vacancy can be estimated to be small; in *copper* it was found less than 0.1 eV, in remarkable disagreement with Seeger's value (quoted in Chapter V)

of 0.3 eV for the association energy between two vacancies in this metal. Anyhow, at the elevated temperatures where diffusion takes place in the noble metals the dissociation is always complete.

Haven has compared his low-frequency conductivity data discussed in the previous section also with experimental results on the diffusion of radioactive Na-tracer ions. The correlation factor for vacancy diffusion in this type of lattice is 0.78, the transport number of the Na^+-ion 0.86. Associated Na^+-Cl^- vacancies also take part in the tracer diffusion process. Again these do not influence the conductivity, but the diffusion coefficient of the tracer ions must now be written as

$$D = D_{ass} + \frac{0.78kT}{Ne^2}\sigma . \tag{19}$$

In practice, the various observational results can be very well described mathematically by disregarding D_{ass} altogether, at least at moderate temperatures. The degree of association between Na^+- and Cl^--ion-vacancies is thus obviously very small, or the mobility of associated pairs is very low. In contrast, at low temperatures the contribution to the tracer diffusion of associated Na^+-vacancy- Ca^{2+}-ion pairs is of course *not* found to be negligible.

We have discussed Haven's work in some detail in this chapter as an example of the large amount of information that can be drawn from simple direct and alternating current conductivity measurements.

Irradiation with ionizing radiation (X-rays or γ-rays) decreases the ionic conductivity of impure NaCl containing divalent metal ions (SEITZ [1954]). This is because the radiation produces free electrons, and these are captured by the halogen-ion vacancies to form *electrically neutral* F-centres (see Chapter XXVI). The vacancies concerned can therefore no longer take part in the conduction process. Also the self-diffusion rate is observed to decrease under these circumstances.

The ionic conductivity of *pure* NaCl is but slightly affected by irradiation. Thus presumably the capturing process is enhanced by the simultaneous presence of associated divalent ions. Possibly only the associated vacancies become occupied, and only the association equilibrium is disturbed. The number of free unoccupied vacancies will then also decrease, leading to the observed decrease of the conductivity.

A particular form of conduction in insulating crystals is *dielectric breakdown*, which takes place under a very high applied field strength (10^6 V/cm). The atomic theory of breakdown is fairly complicated in

detail and will not be discussed here. The presence of lattice imperfections generally *increases* the breakdown strength considerably, especially when the voltage is not applied suddenly, but increased gradually to its ultimate breakdown value. In particular plastic deformation (e.g. compression) can increase the breakdown strength appreciably (ITOH and SUITA [1957]). The dislocations and point defects formed presumably provide nucleation centres for the precipitation of metal ions, resulting in well-conducting patches of metal, which lower the internal field strength.

The phenomenon of photoconductivity will be discussed in the next chapter, as it is of an electronic rather than of an ionic nature.

REFERENCES

AMELINCKX, S., J. VENNIK and G. REMAUT (1959); J. Phys. Chem. Solids **11,** 170 and id. **16,** 158 (1960).
BASSANI, F. and F. G. FUMI (1954); Nuovo Cimento **11,** 274.
BASSANI, F. and R. THOMSON (1956); Phys. Rev. **102,** 1264.
BRECKENRIDGE, R. G. (1952); Imperfections in nearly perfect crystals, p. 219 (Wiley, New York).
CAFFYN, J. E. and T. L. GOODFELLOW (1955); Nature **176,** 878.
COMPAAN, K. and Y. HAVEN (1956); Trans. Faraday Soc. **52,** 786.
COMPTON, D. W. and R. J. MAURER (1956); Phys. Chem. Solids **1,** 191.
ETZEL, H. W. and R. J. MAURER (1950); J. chem. Phys. **18,** 1003.
FISCHBACH, D. E. and A. S. NOWICK (1958); Phys. Chem. Solids **5,** 302.
FRIAUF, R. J. (1957); Phys. Rev. **105,** 843.
FRÖHLICH, H. (1949); Theory of dielectrics (Clarendon Press, Oxford).
GORDON, R. B. and A. S. NOWICK (1956); Acta Metall. **4,** 514.
GYULAI, Z. and D. HARTLY (1928); Z. Physik **51,** 378.
HAVEN, Y. (1954); Report Conference Defects in Solids, Bristol, p. 261 (London, The Physical Society).
HOVE, J. (1956); Phys. Rev. **102,** 915.
ITOH, N. and T. SUITA (1957); J. Phys. Soc. Japan **12,** 214.
KELTING, H. and H. WITT (1949); Z. Physik **126,** 697.
LIDIARD, A. B. (1957); Encyclopedia of Physics, Part XX (Springer, Berlin).
MAPOTHER, D. E., H. N. CROOKES and R. J. MAURER (1950); J. Chem. Physics **18,** 1231.
PRATT, P. L. (1958); Vacancies and other Point Defects in Metals (London, The Institute of Metals).
SEITZ, F. (1954); Rev. Mod. Phys. **26,** 7.
WITT, H. (1953); Zs. Physik **134,** 1170.

CHAPTER XXVI

OPTICAL PROPERTIES OF DEFECTS IN IONIC CRYSTALS

Dislocations and point defects strongly influence the optical properties (such as the position and shape of optical, infrared and ultraviolet bands) of polar crystals. Conversely, irradiation with light affects the electrical conductance and other electronic properties of these crystals. Both influences are primarily caused by the effect of *centres*, of the kind that have already been mentioned in Chapter XXV. In this chapter we shall discuss very briefly the properties of *colour centres* in the alkali halides and silver halides, two groups of polar materials in which centres play an important role. The situation in this field of research is in a rapid state of development, and many aspects are still unclear. It is therefore that we content ourselves by giving only a superficial discussion.

1. Colour Centres in the Alkali Halides

Preparation in excess metal vapour or irradiation with X-rays at ordinary temperatures produces in most alkali halides strong visual

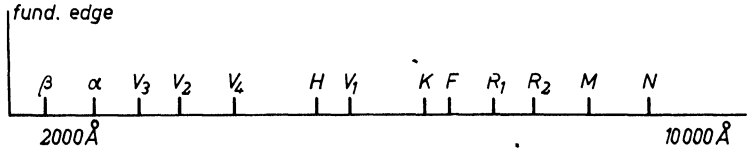


Fig. 1. Approximate positions of the various defect absorption bands in alkali halides (after VARLEY [1954]).

coloration (darkening). For instance, X-rayed KCl crystals become blue, NaCl crystals yellow. Closer examination of the absorption

spectrum reveals primarily the presence of a wide absorption band, the so-called *F-band* (F from Farbzentrum, German for colour center). Many characteristics of this absorption were described by Pohl and his collaborators in the 1930's. Table 1 gives the approximate wavelengths of the F-band centres in several alkali halides (see also Fig. 1).

TABLE 1

Position of F-band in the alkali halides

Crystal	*F-band*
NaCl	4650 Å
NaBr	5400 Å
KCl	5630 Å
KBr	6300 Å
KI	6850 Å
RbCl	6240 Å

DE BOER [1937] has proposed the following model for a hypothetical *centre* that should be responsible for the F-absorption. This "F-centre" should consist of an (effectively positively charged) halogen-ion vacancy which has captured one electron and thus become neutral. The electron must then actually be considered as shared by all the neighbouring metal-ions, as has been described in Chapter XXV. The quantum mechanical description of this system closely resembles that of a hydrogen atom; the stationary orbits in the crystal have radii which are comparable to the corresponding radii of the stationary electron orbits in the hydrogen atom except for a factor ε, the dielectric capacity of the crystal. Above the groundstate a converging series of discrete excited states exist, which terminates in an ionization continuum. The F-absorption proper corresponds to transitions from the ground state to the first excited state. This model has been shown to explain very well the actual state of affairs in many if not all polar crystals.

A detailed investigation of the form of the absorption curve mostly reveals the presence of a hump on the short wave-length wing of the F-band. This has been termed the K-band and was believed to be associated with transitions of the electron in the centre to higher excited states. Recent research on the bleaching behaviour (see below) has brought to light that K- and F-band differ in their properties, and are therefore presumably associated with different centres.

The short wavelength side of the F-band extends rather far to the ultraviolet. This corresponds to the finite ionization probability of the centre.

Crystals coloured by preparing them in excess metal vapour are called *additively coloured* crystals. The coloration is associated with the presence of a non-stoichiometric number of ions of one kind in the crystal. In order to preserve charge equilibrium, vacancies are then simultaneously introduced at the sites ordinarily reserved for the other kind of ions, or interstitial positions become occupied. When these vacancies, or other defects as the case may be, capture electrons (or holes), centres are produced.

As has already been mentioned, similar centres can be formed (although usually fewer in number) in alkali halides of normal composition by irradiation with X-rays or, for that matter, with γ-rays or fast electrons (*photochemical coloration*, see also next chapter). It is assumed that by absorption of the high-energy radiation lattice defects (vacancies, not interstitial ions; these are not normally encountered in pure alkali halides), as well as *pairs* of *electrons* and *holes* (*excitons*) are formed (see below). Of the various recombination processes that can be imagined, the formation of F-centres is again a very likely one. Positive-ion vacancies, which can trap a *hole*, will also be produced. Centres which have trapped a hole are called V-centres; they give rise to absorption bands of their own (Fig. 1).

A third method of colouring an ionic crystal is by *electrolysis* at an elevated temperature. The metal ions collect at the cathode and, when certain precautions are taken, they diffuse back into the crystal, thus locally producing a stoichiometric excess of metal, the charge of which is neutralized by the simultaneous presence of halogen vacancies. F-centres can thus be formed, but these are partly dissociated at the temperature of the procedure into positively charged vacancies and electrons. The vacancies in turn take part in the electrolysis process. At the anode similar phenomena occur, and the various defects formed can interact together in the bulk. One can describe the complete process in alkali halides most easily by saying that F-centres enter the crystal at the cathode, V-centres at the anode.

Room temperature irradiation of a coloured crystal with light in the wavelength region of the F-band (thus visible or even white light) causes the F-absorption gradually to decrease (*bleaching*). Instead of it, other absorption bands, situated more to the long wavelength side,

appear. These have been called R_1, R_2, M and N-bands respectively in the order of increasing wavelength (Fig. 1). They have been interpreted by SEITZ [1954] as caused by the excitation of larger centres, consisting of more than one vacancy, which are formed during the bleaching by coagulation of the F-centres. That such coagulation indeed will take place becomes clear when it is realized that, when an F-centre ionizes under influence of the radiation, the corresponding vacancy becomes positively charged again. This can then associate (if the temperature is not too low) with a moving positive-ion vacancy that is negatively charged, to form a neutral pair. The latter vacancy is very mobile at room temperature in alkali halides as shown in the previous chapter. The neutral pair is also very mobile and can in its turn be trapped at other defects. Even colloidal alkali-metal particles can be formed in this way, especially so at more elevated temperatures.

It should be noted here already that the assignment of specific centres or complexes of centres to certain bands is still largely a matter of speculation. Often even the nomination of centres and of the

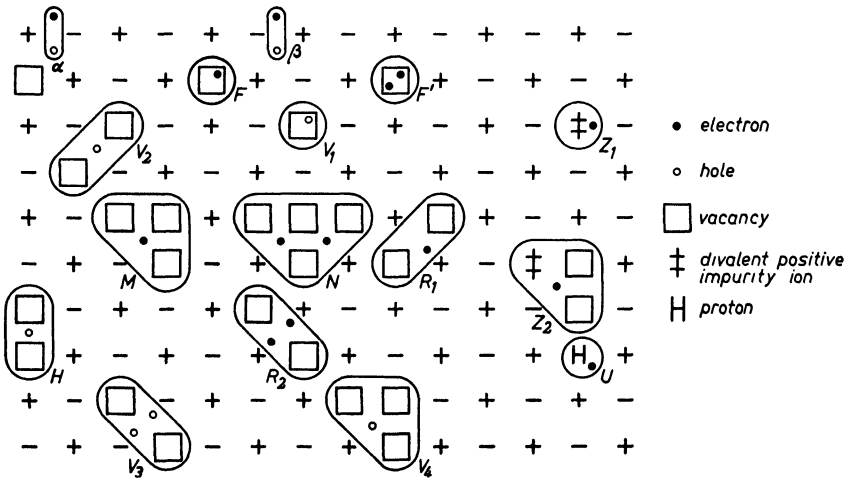


Fig. 2. Interpretation of the various colour centres in the alkali halides according to Seitz. Especially the interpretation of the higher F-centres, the V-centres and the impurity centres is very doubtful.

absorption bands they are associated with are different, because the interpretation of the latter has already been changed several times.

Another way of producing M-bands etc. is by additively colouring a

crystal at a high temperature, thus producing a very large number of potential F-centres, and afterwards quenching the crystal. Subsequent annealing at room temperature then leads first to the appearance of the F-band and successively to that of M, R and N-bands. These are again due obviously to clustering of the vacancies, present in large supersaturation.

The following series of identifications was proposed by Seitz (Fig. 2) but is by now to be regarded as obsolete.

M-centre = F-centre + neutral vacancy pair = 2 halogen vac. +1 metal vac. +1 electron

R_1-centre = F-centre + neg.-ion vacancy = 2 halogen vac. +1 electron

R_2-centre = F-centre + F-centre = 2 halogen vac. +2 electrons

N-centre = three neg.-ion vacancies +1 or 2 metal vac. + electron (very dubious).

Doubt has been cast on these interpretations by various experimental results, including those of UETA [1952] and VAN DOORN and HAVEN [1955, 1956]. The latter authors irradiated a KCl crystal with "F-light" *polarized* in a [011]-direction; the beam followed an [100]-direction. They then measured (with another polarized beam of light) the absorption in the F, M, R and N-bands (which were already there before the irradiation, the crystal having been quenched), and found that a few minutes after the first irradiation, the F-absorption for [011]-polarized light (polarized, that is, in the same plane as the irradiation beam) *decreased*, whereas the M-absorption in this polarization direction *increased*. Conversely, perpendicularly $[0\bar{1}1]$ polarized light became gradually more absorbed in the F-band, and less in the M-band (Fig. 3). This *dichroism* of the M-bands and F-bands can only be observed at temperatures below 200 °K. Irradiation in the M-band had no effect whatsoever. Nor did the F-band become dichroitic after F-band irradiation if no M-centres were already present before this.

The explanation proposed for all this was the following: it was assumed that the M-centres are *anisotropic*, and have a satellite absorption band (F(M)-band) nearly coinciding in wavelength with the F-absorption but further independent of it; the axes of the M-centres are always directed along ⟨110⟩ directions. The dichroism in the F-band then arises, not from the (isotropic) F-centres, but from the F(M)-band absorption. Irradiation in the F-band (or to be precise, in the F(M)-band) causes a *reorientation* of the axes of the M-centres,

that is, it directs most of them along the electric vector. Once this has happened, the absorption for light polarized in this direction in the F(M)-band has decreased, that in the proper M-band has increased.

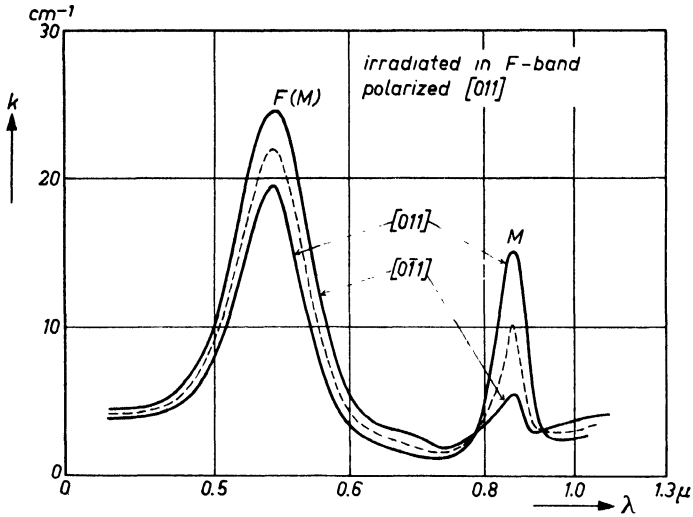


Fig. 3. F- and M-absorption in KCl after irradiation with F-light polarized in the [011]-direction. For light polarized in the same direction the F-absorption has decreased, whereas for light polarized perpendicularly it has increased over the pre-irradiation absorption (dotted curve). After VAN DOORN and HAVEN [1955].

To understand this, it must be assumed that the transition responsible for F(M)-absorption is only sensitive for light polarized perpendicularly to the length direction of the centre, whereas the real M-absorption is sensitive only to light polarized *in* the direction of the centre.

The model proposed by VAN DOORN [1957] for the anisotropic M-centre is that of two neighbouring F-centres (two halogen-ion vacancies and two electrons, Seitz's R_2-centres). The R-centres would, according to this author, consist of combinations of at least three centres. Observations with polarized light indicated for the R-centres a preferred direction along [111] (UETA [1952]). If true, they could not be the centres proposed by Seitz.

Several investigators have recently confirmed details of this work and of related studies and have extended it to other cases, a description of which must be omitted. For example, KONITZER and MARKHAM

[1957] found that the bandwidth of the F-band increased on bleaching. They deduced from this the presence of an F(M)-band component situated close to the F-band which of course becomes increasingly important on bleaching of the latter.

Valuable new information on the anisotropy of centres has been obtained by paramagnetic resonance (KÄNZIG [1955], CASTNER [1957]). The very detailed interpretation of this work also points to the existence of anisotropic colour centres in the alkali halides. The method is based on the determination of the interaction between the electron belonging to a centre with the neighbouring ions as it moves along them. The orbit can be considered as a linear combination of atomic orbits on these ions. Many such combinations are possible, which leads to many spin resonance absorption lines lying very close to each other. Analysis of such band systems with a high-resolving power spectrometer gives information on the symmetry of the centre.

Colour centres of the types described appear to be luminescence centres as well: irradiation in their own absorption band causes a characteristic luminescence in the near infrared region of the spectrum. In KCl, for instance, irradiation in the F-band gives rise to luminescence around 10010 Å, in the M-band around 10610 Å, and in the R-band around 12340 Å (VAN DOORN and HAVEN [1956]).

Narrow absorption bands which have been called *exciton bands* and are usually denominated by greek letters α and β, have been observed very close to the fundamental absorption limit of the crystal (which corresponds to the transition of an electron from a halogen ion to a metal ion, that is, from the valence band to the conduction band). These bands are due presumably to the excitation of electrons situated on halogen ions surrounding a halogen vacancy (α-band) or an F-centre (β-band). The excited electron becomes nearly ionized, but not quite, its "hole" remains weakly bound to it and to the vacancy. The lifetime of such an "*exciton*" is very short (smaller than 10^{-8} sec). The α and β bands thus are associated with *trapped excitons*. By irradiation in the β band the F-absorption is increased; evidently the halogen ions now become fully ionized.

Free excitons give also rise to an absorption band, still closer to the fundamental absorption edge, and extended to the long wave length side. This is the so-called "long-wavelength tail" of the fundamental absorption limit. We shall see below that it is strongly structure sensitive.

On X-ray irradiation at a sufficiently low temperature (liquid air) or on preparing the salt in excess halogen vapour (MOLLWO [1935]), optical absorption in the short-wavelength *V-bands* is observed next to that in the visual F-band. At least six separate absorption bands have been distinguished in various crystals, called V_1, V_2, ... , V_7, V_k and H-bands. The centres responsible for this absorption, the V-centres, have already been mentioned; they were until recently regarded as the counterparts of the F-centres: *holes* trapped by vacancies. SEITZ [1954] has tentatively proposed the following very hypothetical identification (see Fig. 2).

V_1-centre = positive-ion vacancy + hole = counterpart of F-centre

V_2-centre = 2 positive-ion vacancies + hole = counterpart of R_1-centre (Seitz's interpretation)

V_3-centre = $2\,V_1$ = 2 positive-ion vac. + 2 holes = counterpart of R_2-centre (Seitz's interpretation)

V_4-centre = neutral vacancy pair + V_1 = 2 pos.-ion vac. + 1 neg.-ion vac. + 1 hole = counterpart of M-centre (Seitz's interpretation).

The V-centres are only stable at low temperatures. The binding energy of a hole to a positive-ion vacancy should evidently be much smaller than of an electron to a halogen vacancy since already on slightly warming the crystal the centres become "ionized", to judge from the disappearance of the absorption. This is one of the reasons why the study of V-centres in alkali halides is in a much less satisfactory state than that of F-centres. Doubt has been cast again on Seitz's interpretations by spin resonance experiments (KCl, CASTNER and KÄNZIG [1957]; LiF, WOODRUFF and KÄNZIG [1958]), which showed quite convincingly that some V-centres are not spherically symmetric in these crystals. The hyperfine structure of the resonance lines could be best interpreted as due to the existence of a *Cl_2^--molecule ion* type of configuration in some cases. In other words, one hole is shared by two neighbouring halogen ions and may be regarded as "self-trapped" (Fig. 4). The axis of symmetry of this system is ⟨110⟩, in accordance with the experimental evidence. The description of a centre, especially when consisting of more than one defect as a *molecule* formed by the "remaining" atoms, is attractive. The stability of such a centre may well be related to that of the molecule outside the lattice (Kröger, private communication).

A distinction can be made between several possible configurations

showing the characteristics of a molecule-ion and the required symmetry. The *self-trapped hole* proper is shared by two neighbouring halide-ions, which are then drawn somewhat together, forming the molecule situated in the (large) space originally destined for two single halogen-ions. This configuration is assigned to the V_k-centre, which is responsible for the V_k-band. Another centre, observed in LiF and closely related to the V_k-centre, has an axis of symmetry deviating slightly from the [110]-direction. It has been interpreted by KÄNZIG [1958; 1960] as an F_2^--molecule-ion situated next to a Li-vacancy. This may also be thought of as a hole trapped on the latter vacancy, Seitz's V_1-centre again.

Irradiation at a still lower temperature (liquid hydrogen) does not produce the V_1-band, but another band, slightly further shifted to the short-wavelength side (together with the F-band). This has been called the H-band, and is generally interpreted again as due to a centre containing a trapped hole, or a halogen ion on which an electron is missing. On warming-up, these H-centres lose their weakly bound hole (it is even less stably bound here than in the V_1-centre). The hole can either be trapped at a halide-ion (formation of a V_k-centre) or eliminated at an F-centre. The H-band indeed bleaches easily on warming up to liquid air temperature, and the F-band, simultaneously formed at 20 °K, is then also found to decrease in strength.

Previously, the H-centre was thought of as a hole trapped on a metal-ion vacancy–halide-ion vacancy pair. However, spin resonance experiments have again shown this interpretation to be untenable. An explanation originally advanced by VARLEY [1954] seems more probable. According to the mechanism explained in section XIV–2.2.4 irradiation of ionic crystals should produce negative-ion interstitials. The H-centre is now considered as such an interstitial which has trapped a hole, thus as a neutral interstitial halogen ion. The close relation to the V-centres is apparent.

Very-low-temperature irradiation leaves the F-band practically unchanged, at least initially, only the rate at which it is formed is somewhat retarded. However, an F-band formed by irradiation at liquid helium temperature is found to be inherently unstable; it bleaches spontaneously, even in the absence of light and when the material is stored at this temperature. To explain this, it has been assumed that the various centres of F- and V-type are formed very *close together* under these circumstances, and that by quantum mecha-

nical *tunnelling* the electrons can be transferred from one centre to another. This *clustering* of the defects is an indication that they are formed (under X-irradiation) near *preferred sites*. We shall see below that these are presumably dislocations. At normal temperatures the clusters dissolve so rapidly by diffusion that tunnelling becomes impossible.

Another weak band, intimately associated with the F-band is present at low temperature. It has been called the F'-band and is considered as due to absorption by F-centres that have trapped an *additional* electron. These F'-centres (Fig. 2) are, of course, very unstable, and dissociate into F-centres and electrons on warming. Irradiation in the F'-band causes the F'-centres to be ionized of their second electron. The electrons are trapped by empty halogen vacancies to form new F-centres, and in doing this give rise to a typical luminescence in the infrared, which can be directly compared to the F, M, etc., luminescence mentioned above. The reverse process, formation of F'-centres from F-centres, indeed occurs on low-temperature irradiation in the F-band of crystals containing a fair amount of the latter.

Impurities in alkali halides give rise to still other absorption bands. Adding divalent metal ions results in a number of absorption bands that are commonly designated by the letter Z. The centres responsible for these bear some analogy to F-centres etc., the (positively charged) halogen vacancy having been replaced by an (equally positively charged) M^{2+}-ion. The most conspicuous bands in KCl containing Ca are the Z_1- and Z_2-bands on the long wavelength side of the F-band. The Z_1-centre was regarded by Seitz as a Ca^{2+}-ion + one electron, inter alia because F-band illumination produces Z_1-absorption. New evidence obtained by PICK [1958] on the bleaching and photoconductivity behaviour of these centres rather indicated the presence in the centre of a positive-ion vacancy as well. The Z_2-centre was thought by Seitz to consist of a Ca^{2+}-ion + a neutral vacancy pair + one electron, as it is formed by heating a crystal containing Z_1-centres. Pick, however, considered this Z_2-centre as the analogue of the F-centre. The discrepancy, like so many others in this field, has not yet been resolved.

A final type of centres occurs when alkali *hydrides* are mixed with the halides. H-ions are then substituted for some of the halogen ions, and these substitutions act as absorption centres called U-centres. They give rise to absorption bands in the near ultraviolet (2280 Å in

KBr). Illumination in its own light at room temperature bleaches the U-band and develops the F-band. This indicates that neutral H-atoms escape from the crystal, leaving vacancies behind. When the irradiation takes place at very low temperatures, however, the α-band appears instead of the F-band. Evidently the protons leave the crystal in this case alone, unaccompanied by their electrons.

Impurities have other effects on the optical properties, e.g. simply by their influence on the defect concentration, and also by their influence on the presence of dislocations (see below).

2. Photoconductivity

A valuable tool for the study of the nature and properties of the different centres in heteropolar crystals such as the alkali halides is the measurement of *photoconductivity*. Irradiation with light or ionizing radiation of coloured crystals nearly always gives rise to some electronic conduction. The magnitude and dependence of this on temperature, wavelength, etc., gives valuable information on the centres responsible for the supply of conduction-electrons. Many of the conclusions just mentioned are based primarily upon the outcome of photoconductivity measurements (e.g. those of PICK [1958]). The theory of this phenomenon cannot be discussed in this book, however, we refer the reader to the discussion by SEITZ [1951], to a well-known paper by APKER and TAFT [1952] and to the review article by Pick.

3. Silver Halides, Photolytic Darkening

We have, till now, confined our attention to the alkali halides only. In the *silver halides* the situation with regard to optical centres is in many aspects quite different. These substances are very important because of the optical properties which lead to their use as photographic materials, and we shall devote a few lines to them here. It was already pointed out that in the silver halides, owing to the more homopolar character of the interionic binding forces (AgI, the last of the series, even crystallizes in the zincblende crystal structure, a typical homopolar structure, see Chapter I), Frenkel disorder is present rather than the Schottky disorder such as prevails in the alkali halides. At not too elevated temperatures Ag^+-ion vacancies and Ag^+ interstitial ions are the dominating equilibrium defects in AgCl and AgBr.

The interstitial ions are very mobile, their jump frequency is of the order of 10^{10} sec^{-1} at room temperature and is still 10^4 sec^{-1} at 100 °K; the activation energy of the mobility is only about 0.1 eV. (That of an Ag^+-vacancy is much larger, viz. 0.36 eV; COMPTON and MAURER [1956].) Above a few hundreds degrees centigrade some Schottky defects are presumably also present in the silver halides, but not much is known about them quantitatively. An indirect indication for these Schottky defects is the high-temperature distortion of the various absorption band contours (SEITZ [1954]), which distortion could be caused by the condensation of vacancies on dislocations.

Well-defined absorption bands, such as those in the alkali halides, do not occur in pure silver halides. Neither is the fundamental absorption limit sharp; it has an extended long-wavelength tail. Irradiation of AgBr and AgCl in this tail, or with ionizing radiation, gives rise to absorption over a wide frequency band in the visible region, that is, to visual *darkening*. This is the well-known *photolytic darkening* by visible and ultraviolet light. Presumably the irradiation produces dissociation of excitons into electrons and holes; these migrate to suitable defect sites and form absorption centres. In the first stage of the darkening, the centres responsible for the absorption, like the colour centres in the alkali halides, only contain one or a few atoms. In later stages of *development* coagulation occurs if the temperature is not too low, and colloidal silver particles consisting of many silver atoms collect at the centres (*silver specks*). The extinction is now primarily caused by *scattering*. The darkenability under irradiation is very much enhanced by previous cold work; some previous cold work is even required to produce any darkening at all at low temperatures. Presumably the dislocations formed act as preferred sites for the formation of photolysis centres and silver specks. This was confirmed beautifully some years ago by the work of Mitchell and collaborators (*decoration* of dislocations), which has been mentioned in Chapter V.

The primary photolytic centres are probably not simple Ag^+-ions + electrons; Seitz mentions a number of objections to this idea. There is, generally, rather too little interstitial silver available to produce the observed darkening. Further, one would not expect the electron to be very tightly bound to the Ag^+-ion, and the combination should be extremely mobile, thus unstable against precipitation on dislocations. Seitz proposed that *incipient halogen-ion vacancies* (see Chapter XXV) or actual halogen ion vacancies, which have trapped an electron, in

other words, incipient or actual F-centres, were the centres responsible for darkening. An incipient halogen vacancy (situated, of course, on a dislocation) which has trapped an electron bears an effective charge $e/2$ and will therefore attract interstitial Ag^+-ions, which ions will then become trapped at the dislocation. The end result is the same as if the interstitial ion had first captured an electron and then migrated to a dislocation. The advantage of Seitz's model over that which assumes that Ag^+-ions are the darkening centres is, that after the described sequence has been completed the incipient halogen vacancy is reset and can attract another Ag^+-ion. The role of the dislocations is only that of an intermediary, namely to provide an infinite source of incipient F-centres. As various kinds of dislocations are possible, various types of photolytic centres are imaginable and will occur, resulting in a broad absorption band. The preparation of AgCl in excess halogen vapour produces yellow coloured crystals. This is due to the formation of V-centres, as in the alkali halides. Coloration by impurities has also been observed. Excess silver gives darkening similar to photolytic darkening, but addition of divalent metal ions usually decreases the photolytic darkenability. The substitution of Cl^--ions by S^{2-}-ions (heating in Ag_2S vapour) very much enhances the sensitivity to light however: sulphur is a well-known *sensitizer*. Presumably the presence of S^{2-}-ions enhances the concentration of (incipient) F-centres. It is not certain whether the sulphur ions actually occur doubly charged in the lattice.

We have discussed the silver halides because of their practical importance, and because of the important role played by the dislocations in these materials. There are many other groups of polar substances that show interesting optical properties due to lattice defects, but to treat these here would take us too far. We only mention yet the thallium halides, because they strikingly resemble the silver halides in most of their properties, and the very interesting case of quartz, a substance on which not only optical but also dielectric and mechanical studies have been performed, which have resulted in a wealth of detail regarding the nature of the various defects and combinations of defects in this material (VOLGER [1960]).

REFERENCES

APKER, L. and E. TAFT (1952); Imperfections in nearly perfect crystals, p. 247 (Wiley, New York).
BOER, J. H. DE (1937); Rec. Trav. Chim. Pays Bas **36,** 301.
CASTNER, T. G. and W. KÄNZIG (1957); Phys. Chem. Solids **3,** 178.
COMPTON, W. D. and R. J. MAURER (1956); Phys. Chem. Solids **1,** 191.
DOORN, C. Z. VAN (1957); Philips Res. Reports **12,** 309.
DOORN, C. Z. VAN and Y. HAVEN (1955); Phys. Rev. **100,** 753.
DOORN, C. Z. VAN and Y. HAVEN (1956); Philips Res. Reports **11,** 479.
KÄNZIG, W. (1955); Phys. Rev. **99,** 1890.
KÄNZIG, W. (1960); Phys. Rev. Letters **4,** 117.
KÄNZIG, W. and T. O. WOODRUFF (1958); Phys. Rev. **109,** 220.
KONITZER, W. and J. MARKHAM (1957); Phys. Rev. **107,** 685.
MOLLWO, E. (1935); Nachr. Akad. Göttingen, p. 215.
PICK, H. (1939); Z. Physik **114,** 127 and Ann. der Physik **35,** 73.
PICK, H. (1958); Suppl. al Nuovo Cimento VII, no. 2, 498.
SEITZ, F. (1951); Rev. Mod. Phys. **23,** 328.
SEITZ, F. (1954); Rev. Mod. Phys. **26,** 7.
UETA, M. (1952); J. phys. Soc. Japan **7,** 107.
VARLEY, J. H. O. (1954); J. Nucl. Energy **1,** 130.
VOLGER, J. (1960); Progress in Semiconductors **4.**
WOODRUFF, T. O. and W. KÄNZIG (1958); Phys. Chem. Solids **5,** 268.

CHAPTER XXVII

PLASTIC DEFORMATION AND IRRADIATION OF IONIC CRYSTALS

1. General Features

In contrast to the extensive discussion of the plastic behaviour of metals in Part II we shall be brief about that of ionic crystals. Much less detailed knowledge exists on this subject, since the mechanical strength of these materials is usually not a property of much interest. Most ionic materials are under normal circumstances hard and brittle. Only by taking special measures such as wetting the crystal with a suitable solute, they can be made to deform (slowly) in glide and incidentally by twinning. The latter mode is of special importance in the hard non-cubic compounds such as corund (Al_2O_3) and calcite ($CaCO_3$). Cubic polar crystals (except those with strong homopolar character that crystallize in the ZnS-structure) usually glide along ⟨110⟩-directions, parallel to rows of equally charged ions. The glide plane is normally {110}, but in some cases the cube planes {100} also act as such, for instance in wet rocksalt and in fluorspar CaF_2. (This crystal has a remarkable structure, the fluorine ions being situated on a face-centred cubic lattice, in the tetrahedral interstices of which the Ca-ions are placed; a few of these sometimes occupy octahedral interstices instead.) In rhombohedral and hexagonal lattices the {110}-planes or their equivalents have lost their favourable position, and the only slightly less symmetrical {111}-planes and {0001}-planes respectively take over their function in plastic deformation. Octahedral glide is occasionally observed in cubic crystals, namely at elevated temperatures.

The normal $[110](1\bar{1}0)$-glide systems in cubic crystals do not favour the occurrence of single glide, since the shear stress is always equal in magnitude on two systems simultaneously (namely on those two that transform into each other by interchanging the indices of plane and direction). Silver chloride has been very extensively investigated,

because its properties resemble very much those of metals, also in so far as alloying and recrystallization are concerned (MOELLER *et al.* [1951]). It was found by NYE [1949] to deform in *pencil glide*, similarly to α-iron. The glide direction is again invariably ⟨110⟩, but the glide plane is not uniquely determined, although the slip lines are rather straight. The investigations by NYE and SPRACKLING [1957] form an interesting sequel to these researches: they investigated the birefringence of deformed silver halides, which revealed the existence of appreciable inhomogeneities in the stress distribution, after unloading, of twisted or bent single crystals. Such inhomogeneities are obviously connected with the mechanism of generation and motion of dislocations, and the role of the surface in this. Mixed crystals of various thallium halides (which in many respects resemble the silver halides) have been studied by SMAKULA and KLEIN [1951]. These have {110}-glide planes as usual, but [001]-glide directions. Many other observations exist, but will not be discussed here (see also PRATT [1954]).

Many hard polar substances, and especially non-cubic crystals such as the hexagonal crystals sapphire and beryl, resemble diamond: they *fracture* by *cleavage* instead of deforming. Cleavage planes are the hexagonal base planes or prism faces, or the cube planes and octahedral planes (diamond). The "ductile" materials discussed earlier also generally have only a limited range of ductility; they usually fracture already after a few percent of strain. In this respect the polar substances clearly differ from (cubic) metals, the single crystals of which can usually withstand strains up to a few hundreds of percents. Magnesium oxyde crystals have been found to be brittle in tension but ductile to 20% strain in bending, before cleavage occurs (WASHBURN *et al.* [1959]). This is due to the inhomogeneity of the stress in the latter case which causes dislocation concentrations to form only in the stress-free neutral zone, where they cannot develop into cracks.

The slope of the stress-strain curve of rocksalt crystals decreases slightly with increasing strain. The rate of work-hardening and the flow stress are both very temperature dependent. Occasionally in impure crystals an indication of delayed yielding (Chapter IX) has been observed (photoelastically). A strong Bauschinger effect (Chapter XIII) has been found in bent rocksalt crystals (SCHMID and BOAS [1935]).

Some aspects of plastic flow which are only of minor importance in the plastic deformation of metals, are of foremost importance in

the alkali-halides. Among these are *size effects* and the *Joffé-effect.*

The flow stress and work-hardening coefficient of cylindrical NaCl-crystals of thickness less than 1 cm have been found to depend sensitively on the diameter. Very remarkable is also the large increase in fracture strength, leading to a considerable increase of the ultimate strain before fracture of NaCl-crystals, when they are immersed in water. (Normally brittle rocksalt crystals can be easily bent and twisted in warm water.) This effect, which was first discovered by JOFFÉ [1924], occurs in many heteropolar crystals when tested in a number of inorganic and organic liquids which can act as solvents. Also the rate of work-hardening is less in the wet than in the dry state. The enhanced ductility remains present during a certain period after drying.

The Joffé-effect is either due to the dissolution of a surface layer that prevents plastic glide, or it is connected with the charged nature of dislocations (see below).

An apparently similar influence of the surface is present in metals, as has already been discussed in Chapter XIII; viz. the decrease of plasticity associated with the presence of oxyde, hydroxyde or other layers on the substance and the Rehbinder effect in active liquids respectively. These effects occur in polar crystals also, but there they are usually of even less importance than in the surface-active metals.

Mechanical recovery and polygonization are common processes in ionic crystals. Recrystallization after cold-work and annealing is usually less conspicuous than in metals, but it has been shown to occur in deformed AgCl and NaCl. Often, however, polygonization takes place to such an extent that recrystallization proper can no longer take place (see Chapter XX). Irradiation with X-rays strongly slows down any recovery or recrystallization process, if present, presumably because the dislocations become pinned (see below).

Of the various studies of the *physical* (and mechanical) recovery of deformed ionic crystals we mention here only as an example the work of KANZAKI [1955] on AgBr. This author studied the decrease of the additional ionic conductivity (section XXV–2) induced by plastic deformation, as a function of the annealing temperature. This decrease takes place in three recovery steps, two of which are characterized by an activation energy of about 0.4 eV, and one by only 0.2 eV. By comparison with measurements of KURNICK [1952] on the temperature dependence of the conductivity itself, it was concluded that the smaller

activation energy was connected with the migration of interstitial Ag^+-ions whereas the higher energy should correspond to Ag^+-ion vacancy migration. The recovery is due to recombination of these defects, respectively with Br^--interstitials and Br^--vacancies, to form neutral pairs. This recombination occurs preferentially at *jogs in dislocations* (compare previous chapter); the energy liberated is rather large and has been measured also: 30% of the strain energy can be recovered, which is about 10 times as much as in metals.

Evidently, the dislocation motion has generated Frenkel, Schottky, as well as anti-Schottky defects, the latter two by geometrical means, the first by thermal "evaporation" from jogs, however. From the annealing behaviour, Kanzaki draws the conclusion that Frenkel defects have been formed in a concentration of 10^{-7} and Schottky defects in a concentration of 10^{-4} after 35% plastic strain. These defects recombine in the way suggested above; recombination of Frenkel pairs as such seems not to take place, to judge from the large amounts of external energy liberated during each recovery process.

We conclude that, apart from a few differences, the plasticity of ionic crystals follows the same lines as that of metals and does therefore not need special discussion here. We shall in the following discuss only these typical differences in more detail.

2. Joffé Effect and Surface Sources

From the observations discussed in the previous section it is evident that the condition of the surface exerts a large influence on the plastic properties of (not too thick) alkali halide crystals. This has been explained by assuming that a large fraction of the active dislocation sources in the crystal lies close to the surface. This could be understood in principle by noting that dislocation segments ending on the surface can, owing to the decrease in strain energy arising from the elastic attraction of a dislocation to the surface, act as sources of new dislocations at *half* the critical shear stress that would be needed to activate a source situated wholly in the interior (Chapter III). A recent study by T. SUZUKI [1957] of the plasticity of KCl crystals throws some more light on this point. Annealed KCl crystals were compressed at room temperature along known crystallographic directions. Loaded along a cube direction, single glide was observed during the first one or two per cents of strain. Although not less than 4 slip systems $[01\bar{1}]\,(011)$

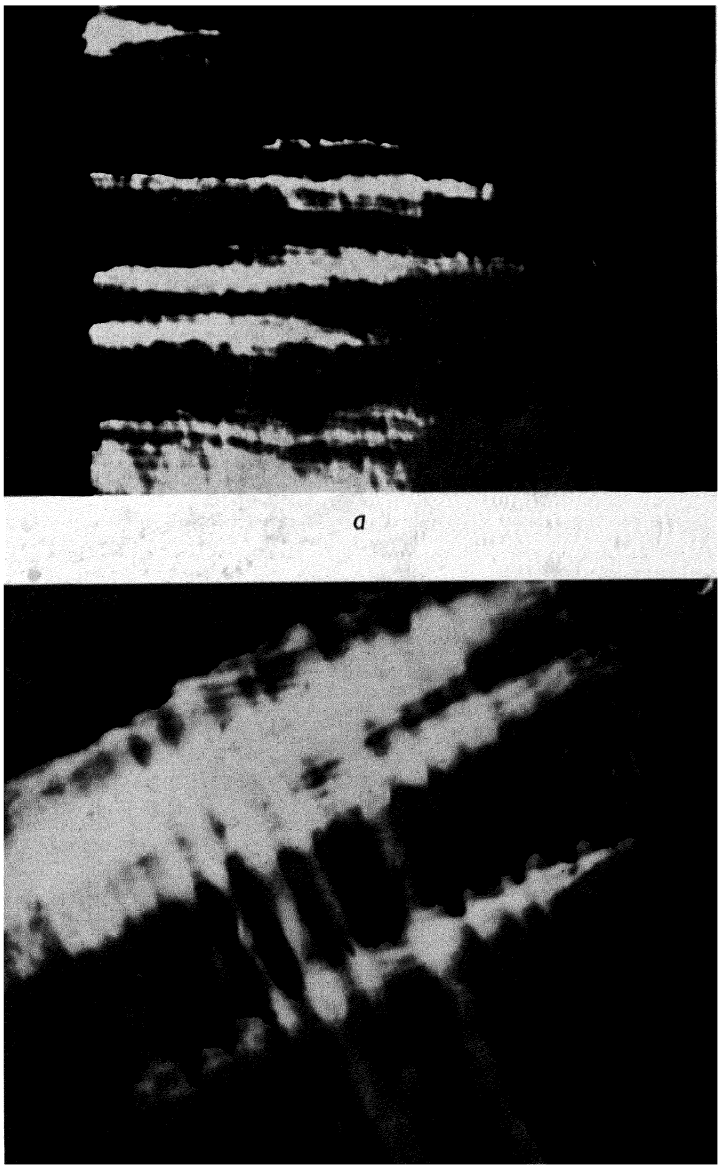


Fig. 1. Birefringence bands in compressed KCl crystal bar. Direction of compression [100]. *a*) After 0.4% compression bands belonging to one single slip system only are present. *b*) At 2% strain bands belonging to a second slip system appear in the surface region. (After SUZUKI [1957].)

are equivalent, only one of them, randomly chosen, is visibly activated; the stress-strain curve was comparable to that of a face-centred cubic metal crystal in easy glide. Photoelastic investigation with light polarized in the glide direction revealed the existence, during the very beginning of the deformation, of strong birefringence bands belonging to the active system *near the free surfaces* (cube faces) *only*. After a strain of about 0.3–0.4% had elapsed bands began to appear in the central region too (Fig. 1*a*). Still later ($\varepsilon > 2\%$) bands belonging to a second system appear (Fig. 1*b*). This is perpendicular to the first one and its bands can again be seen only when the microscope is focused just on or below the surface. It is clear that in these crystals slip is indeed initiated in the regions near the free surfaces, on the first as well as on the second system. Further investigation showed that only such sources were activated of which the Burgers vectors are parallel to the surface.

Polishing with a wet cloth during an intermission of the compression test produced a decrease of the flow stress (which had been raised by work-hardening) by a constant factor, independent of the thickness of the removed layer, as long as this remained less than about 100μ. Further polishing eliminated the work-hardening almost completely, and led in some cases to a flow stress equal to the original critical flow stress. The work-hardening, therefore, seems to take place wholly in the surface layer of about 50 to 100μ thickness. Removing this layer frees new surface sources that are able to operate at a lower shear stress than the already hardened original sources. It seems a fair conclusion to take the thickness d of the active surface layer as two times the effective mean length of the sources, l_0. The critical shear stress is then

$$\tau_0 = \frac{Gb}{2l_0} = \frac{Gb}{d}, \tag{1}$$

which was found to be in very good agreement with the observations.

This work leads to an interpretation of the Joffé effect in terms of the dissolution of the surface layers in which the generation of dislocations as well as (source-)hardening takes place. As soon as unwetted regions of crystal are present, much more rapid hardening takes place in these regions, leading ultimately to local stress concentrations high enough to cause fracture. In the original, undeformed and unwetted crystals, a very wide distribution of source lengths is

present, leading to multiple slip and rapid hardening from the start.

The occurrence of birefringence points to pile-ups of dislocations as the sources of hardening (because these have long range strains around them). They presumably block the sources by their stress field.

Annealing of a deformed crystal has, as usual, the effect of decreasing the flow stress. This might be due to the removal, possibly by migration under the influence of the surface (image) forces, of part of the dislocations present in the surface layer, which is then left relatively empty of dislocations; those which *are* left are presumably longer than the average length before annealing.

PRATT [1957] has advanced a fully different explanation of the Joffé effect, which may well apply to other crystals. It is based on the concept, discussed in Chapter XXV, of *space charge* cylinders surrounding the dislocations and space charge layers bordering the surface, respectively, in ionic crystals. Both charges have the same sign and repel each other; therefore dislocations will be repelled by the surface, unless this is sufficiently rapidly etched away. The surface will thus exert a hardening influence.

3. Mutual Interaction between Dislocations and Point Defects

3.1. COLORABILITY BY IONIZING RADIATION

A very interesting aspect of the plastic deformation of polar crystals is the mutual influence of plastic flow and optical irradiation.

It has already been remarked in Chapter XXIV that dislocations play an important role as incipient sources of defects in alkalihalogenides. It is therefore not surprising that the colorability of deformed alkalihalides under X-ray irradiation is increased by plastic deformation. The exact mechanism of X-ray coloration is not yet well understood at present: as the end products are primarily F-centres, electrons as well as point defects must certainly have been created by the ionizing radiation in some way or another. SEITZ [1954] has proposed that the X-rays produce in first instance excitons, that is, traveling excited states of negative ions (hole-electron pairs, section XXVI–1). When one of these reaches a pair of charged incipient vacancies or *double jog* in a dislocation (Fig. 2), one component (either the hole or the electron) of the exciton is trapped, and a vacancy carrying the same charge is released from the jog under the influence of the liberated excitation energy. For instance, the electron of an

exciton is trapped at an incipient negative-ion vacancy, whose initial charge $+e/2$ then becomes $-e/2$. A nearby positive-ion-incipient vacancy (charge $-e/2$) is now repelled and diffuses away as a real

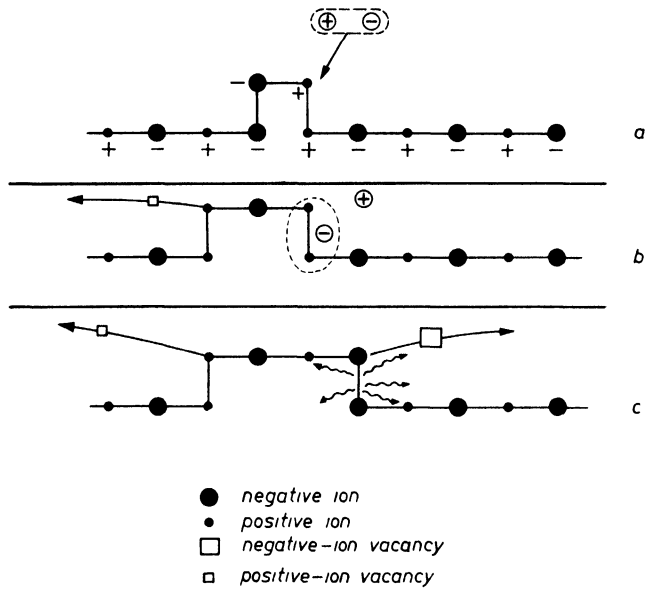


Fig. 2. Seitz's mechanism of formation of point defects by absorption of excitons at dislocations. *a*) An exciton arrives at a negative-ion incipient vacancy, part of a double jog. *b*) The electron is absorbed and a positive-ion vacancy (bearing a negative charge) is emitted from the second jog. An incipient negative-ion vacancy is left in its place. *c*) The hole is subsequently also absorbed, and the energy released in this recombination leads to the evaporation of the negative-ion vacancy.

positive-ion vacancy. A new incipient negative-ion vacancy is then automatically formed together with the already present "incipient F-centre". The latter subsequently traps the remaining hole and becomes a negative-ion vacancy again, which also diffuses away (by the heat generated). Two vacancies of opposite sign have thus been formed in the process, which can elsewhere trap electrons or holes and become centres. The heat released by a recombining exciton is so large (about 6 eV in the alkalihalides NaCl and KCl), that even from unjogged edge dislocation sites vacancies can evaporate. This means that the dislocation must *climb* over a considerable distance in order to produce a reasonable vacancy concentration. CRAWFORD and YOUNG

[1960] did not find any traces of such excessive climb under the microscope. Another difficulty in Seitz's exciton mechanism is, that the vacancies formed in this way cannot diffuse very far at low temperatures. Near to the dislocations the density of F-centres should thus become very high, leading to large internal stresses, and DEXTER [1954] has pointed out that this should influence the shape of the F-absorption band in a way that has not been observed. At liquid nitrogen temperatures some evidence of vacancy clustering after irradiation was indeed found by UETA and KÄNZIG [1954]. The clusters resolve again on heating to room temperature.

GORDON and NOWICK [1956a] have studied the structure sensitivity of the X-ray colorability of NaCl in more detail. They distinguished between two mechanisms of colouring, after plastic deformation. The first, *rapid* colouring, saturates already at an F-centre density of about $10^{17}/cm^{-3}$ and should be due to vacancies already present before the irradiation, for instance as the association partners of divalent positive impurity ions. In accordance herewith, the rapid colouring is not dependent on the amount of previous plastic deformation.

Another, slower type of coloration saturates at much higher F-centre concentrations ($10^{18}\,cm^{-3}$), and depends strongly on the amount of previous plastic deformation. To this type only, Seitz's mechanism or an analogous one should apply. In the next section another aspect of this problem is reviewed. A third way in which vacancies and centres could be formed is by the *Varley mechanism*, already discussed in section XIV–2.

3.2. INTERNAL FRICTION AND DIELECTRIC LOSS

GORDON and NOWICK [1956b] also studied an effect of ionizing irradiation which can, in some aspects at least, be considered as the mechanical counterpart of the optical effects discussed in the previous section. They observed a lowering of the dislocation mobility after X-radiation of NaCl, as measured by means of the internal friction (or rather the elastic modulus). The damping of deformed crystals at 85 kc decreased practically to zero and the elastic modulus increased appreciably. Analogous results had already been obtained before by FRANKL [1953], by a different method. The internal friction concerned is of the strain-amplitude dependent type discussed in section XVII–2.2.4. It depends markedly on the amount of previous plastic defor-

mation, and is therefore presumably also in NaCl associated with the irreversible oscillation of dislocation segments. Subsequent bleaching with visual light has no effect whatsoever, but annealing at 200 °C results in a return of about $\frac{2}{3}$ of the original damping.

Irradiation at liquid-air temperature does *not* reduce the internal friction. Only after subsequent bleaching at room temperature the damping decreased markedly.

An explanation of these phenomena was put forward in terms of the *pinning* of dislocations by point defects. Mild annealing directly after plastic deformation has the same influence as irradiation: both treatments result in a lowering of the dislocation mobility, owing to the diffusion of point defects to the dislocations, to which they become bound by electrostatic effects. However, if pinning of dislocations by vacancies is indeed the cause of a decrease of the damping, it can no longer be assumed that the same dislocations act as *sources* of these vacancies during irradiation; they behave rather as *traps*. The vacancies must then come from other sources, e.g. from neutral vacancy pairs dissociated by the X-ray quanta. The negative-ion vacancy should then become an F-centre, the positive-ion vacancy should migrate to the dislocation because this is positively charged (Chapter XXV). Slight warming of a crystal and *simultaneous* illumination with white light removes the pinning and thus causes the return of the damping. Presumably charged negative-ion vacancies are then produced by the ionization of F-centres, and these are attracted electrostatically by the positive-ion vacancies to form neutral pairs again, which can pin the dislocations no longer. PRATT [1954] has severely criticised this interpretation, inter alia because V-centres are also formed (next section).

From their experiments Gordon and Nowick concluded that the dislocation density in a NaCl crystal compressed by 4% lies between 10^8 and 10^9 cm^{-2}.

The low-frequency (1000 c/sec) damping in alkali halides has also been investigated, among others by ISHIGURO *et al.* [1954]. Deformation as well as quenching from an elevated temperature results in a type of frequency independent damping (hysteresis ?), which increases rapidly with increasing temperature. Superposed on it is a relaxation type of damping (see Chapter XVII), with a low damping maximum near 120 °C. Subsequent irradiation with X-rays does not affect the temperature-dependent part of the internal friction but removes the relaxation

peak. This returns after room temperature bleaching, or after mild annealing. Strong annealing removes it again, together with all other effects of the deformation or the quenching. It was suggested that dislocation segments, having become jogged in an irregular manner by the interaction with vacancies, are the cause of the relaxation damping. It is possible, however, that in this case the cause must be sought in the effect of vacancies or vacancy pairs alone, rather than in the interaction with dislocations, since quenching and deformation produce identical results.

Apart from on the mechanical losses, dislocations and vacancies have some influence on the *dielectric* losses of insulating solids. This subject has in first instance been studied in the materials quartz and rocksalt. We shall refrain from a discussion of this very interesting field, because as far as individual lattice imperfections are considered, not much more evidence is obtained about their properties by this method than by measuring the mechanical damping or the d.c. conductivity. The structural imperfections affecting the dielectric loss generally have dimensions exceeding the lattice parameter. For an authoritative discussion we refer to a paper by LIDIARD [1954].

3.3. MECHANICAL PROPERTIES OF IRRADIATED AND QUENCHED CRYSTALS

It has already been mentioned that NaCl crystals often show an abrupt yieldpoint under (compressive) stress. Unlike in metals, this is not necessarily an impurity effect, since in this case also *physical* point defects interact strongly with dislocations. The experimental data on the mechanical properties of quenched and irradiated ionic crystals are only scarce; we refer for a discussion to PRATT [1954, 1957]. Quenching results in an enhancement of the colorability, an increase of the low-frequency damping and a pinning of the dislocations. Irradiation as well as quenching increases the fracture strength and yield stress, whereas the rate of work-hardening in X-ray coloured crystals is much greater than in clear crystals (where it is usually but very small). Bleaching with white light at room temperature does not remove this extra hardening. Addition of divalent impurity ions results in very considerable increases of the fracture strength, much more so than addition of monovalent ions. This may again be due to point defect pinning.

The exact nature of the interaction between vacancies and dislo-

cations responsible for all these effects is not yet clear at the present time, electrostatic and elastic interactions presumably both are important. PRATT [1954] has shown that the lattice relaxation around a neutral vacancy pair is even larger than that around a single charged vacancy (see also section XXIV–2). This should result in a rather strong interaction of *pairs* with dislocations, even if they are neutral, the more so as the lattice deformation is anisotropic. Once a single (charged) vacancy has trapped an electron or hole and thus has become a neutral *centre*, the strains around it are reduced, since the neighbouring ions are no longer repulsed. Even a change of sign of the dilatation may occur, since the electron in the centre does not obstruct by its presence the surrounding ions from moving inwards. Electrostatic effects are also much smaller in centres than in charged defects, thus in general the interaction between centres and dislocations will only be slight.

Apart from volume hardening, irradiation results in an exceptionally strongly hardened surface layer, which was attributed by Pratt to jogging of the climbing dislocations which must produce the vacancies responsible for darkening (section 3.1).

It should be remarked finally that to move a dislocation with its space charge cylinder, defects must be carried along with it by diffusion. This situation is comparable to that occurring in microcreep of metals (Chapter IX). The most conspicuous result would be a characteristic strong dependence of the flow stress on temperature. This has indeed been observed in rocksalt (CAFFYN and GOODFELLOW [1955]).

4. Optical Absorption and Plastic Deformation (not Colour Centres)

4.1. FUNDAMENTAL ABSORPTION

The effect of plastic deformation on the fundamental absorption limit of alkali halides has been studied by DEXTER [1954]: the absorption tail at the long-wavelength side is displaced to even longer wave lengths. This was explained by noting that electronic transitions which are ordinarily forbidden may become allowed in the strongly strained regions near imperfections. Moreover, a dilatation, if any, of the lattice would in itself result in a local shift of the absorption limit. Both effects would be expected to occur near dislocations. The experimental investigation of the optical effects is made difficult by the

simultaneous occurrence after plastic strain of new exciton bands (α and β bands, see section XXVI–1) in the same region of the absorption spectrum, which are associated with excitons created near negative-ion vacancies. Plastic deformation at liquid-air temperature suppresses the creation of excitons in the alkali halides, however. (Apparently very few vacancies are formed by moving dislocations at low temperature.)

Luminescence induced by plastic deformation has been observed by LEIDER [1958] in X-ray coloured KBr. The motion of dislocations evidently triggers some recombination process.

4.2. MICROSCOPIC EVIDENCE FOR DISLOCATIONS IN POLAR CRYSTALS

We have already mentioned in Chapter V the so-called *decoration* of dislocations. By this is meant the microscopic indication of dislocation lines in a crystal through the presence of impurities, which have precipitated on them. The precipitates scatter the transmitted light and the course of the dislocation lines is made optically distinguishable. (In contrast to the photoelastic method, here the dislocation *strains* are not responsible.) Decoration of dislocations was first observed by HEDGES and MITCHELL [1953] in silver bromide which had been "physically developed". Irradiation with light in the fundamental absorption band of slightly impure AgBr layers that had been accidentally deformed during their preparation, showed on subsequent microscopic investigation the presence in the interior of a network of dark lines along which silver specks had been formed. A beautiful hexagonal structure, corresponding to a polygonized dislocation array, was observed in these cases.

Following this work, DEKEYSER, AMELINCKX *et al.* [1954] have developed other still more powerful techniques for decoration. We refer to Chapter V for some discussion. One of their most succesful decoration techniques consists in the heating of an alkali chloride crystal in chlorine gas in the presence of gold vapour, a method which has also been used by the english workers. The crystal thus treated remains colourless when hot, an indication that the gold is dissolved as chloride in the crystal. On cooling, however, the gold chloride decomposes, and colloidal gold particles precipitate along the dislocations.

Decoration has revealed the peculiar dislocation forms like helicoidal dislocations discussed already in Chapter III (Fig. III-3). Presum-

ably these are formed by the action of the stresses set up during the precipitation.

As to the exact atomic mechanism of precipitation, not much is known at present. Presumably jogs in the dislocations play an important role as nucleation centres (see also Chapter XXX).

H. KANZAKI [1956] has studied the correspondence between decoration patterns (obtained by physical development) in the interior and chemical etchpit patterns (see Chapter V) on the surface of annealed and deformed silver chloride and bromide crystals. No unambiguous correspondence was observed: on some crystallographic faces ({100}) the decorated dislocations did not produce etchpits, whereas on other crystal faces ({110}-surfaces) the etch pits did not correspond to decorated dislocations. This discrepancy was interpreted in terms of the differences that are present in the relative orientation of the {110}-extra-half planes of edge dislocations and these surfaces.

4.3. MICROSCOPIC EVIDENCE FOR THE GENERATION AND MOTION OF DISLOCATIONS

The microscopic method has met with special success when applied to silver halides (MITCHELL [1957]) and to alkali halides: NaCl (AMELINCKX [1957]), lithiumfluoride (GILMAN and JOHNSTON [1957]). The application to metals and semiconductors is discussed elsewhere. For a detailed discussion of the phenomena observed we refer to the extensive and beautifully illustrated papers cited. Mitchell was able to study the peculiarities of the three-dimensional network of dislocations in AgCl and AgBr as revealed by his decoration technique, and to follow the influence of annealing and deformation on this network. The formation of sub-boundaries, the object of much speculation in connection with the recovery phenomenon (Chapter XI) could be clearly followed. By applying a special etching technique to LiF, Gilman and Johnston observed the motion of individual dislocation lines emitted by sources. Dislocation nucleation took place at unexpectedly small stresses, of the order of 1 kg/mm^2. By applying transient loads the same authors (GILMAN and JOHNSTON [1959]) were able to show that dislocations move with a velocity which depends exponentially on the stress (see also Chapter XXX). They related this to the observed brittleness of lithiumfluoride under rapidly varying stress. The apparent nucleation of dislocations out of nothing observed by these investigators at low

applied stresses, is a matter still under extensive investigation. A similar discrepancy is found when observing the beginning of plastic flow in dislocation-free germanium crystals. There, one must assume the presence of surface sources that cannot be detected by present means (VAN BUEREN [1959]). Presumably, similar sources are also present in the seemingly perfect lithium fluoride crystals. Also interior sources, maybe consisting of platelike *condensations* of point defects formed during growth can be present. AMELINCKX and BONTINCK [1957] have pointed out that the collapse of such platelets does not lead to a single dislocation loop, but that by the process of helicoidal climb (Chapter III), followed by mutual intersection of the various spiral branches, separate closed rings can form instead. A discussion of the generation of dislocations from vacancy condensates has already been given elsewhere in this book (Chapter XIV).

5. Electron and Neutron Irradiation Effects in Polar Crystals

Irradiation of alkali halides with moderately fast electrons results equally well in the production of lattice defects and colour centres as X-ray irradiation, or even better. As in the latter case, it is assumed that primarily excitons are produced, which by recombination at incipient vacancies then produce the defects. The latter cannot usually be formed directly, since the energy needed for this is too high. The generation of defects can be followed by photoelastic methods. By the same methods, it was found that in high melting point polar crystals such as quartz relatively mild deuteron bombardment did *not* produce any defects. At present the effects of electron and neutron bombardment on various ionic materials have not yet been studied systematically, although new experimental evidence is continuously forthcoming.

Very large irradiation effects have been observed in neutron-bombarded lithium fluoride. They are due to the action of *fission products*. The lithium nucleus is split by thermal neutrons in a He^3 and a He^4 nucleus, and the He gas collects in small voids, of platelike shape. The lattice parameter as well as the density of LiF crystals change appreciably (and equally) during pile irradiation. This points to the formation of Frenkel defects also (see also section XV–4). For a discussion of the bearing of this work on the expected influence of

radiation on other polar crystals, we refer to the review paper by SEITZ and KOEHLER [1956].

Optical effects of neutron irradiation of quartz have been discussed by DITCHBURN *et al.* [1954]. In this material the optical properties are governed to a large extent by the presence of chemical impurities, especially of aluminium and excess oxygen. We shall not go into this work here, but refer the reader to the original paper.

Several investigations have been carried out on the annealing of radiation damage in the alkali halides, but since these are still very incomplete a full discussion is not warranted here. Very interesting are RÜCHARDT's observations [1956] on the presence of very low (< 10 °K) temperature annealing steps in X-ray irradiated KBr, which point to an inherent instability of "vacancies" produced at such a low temperature. They presumably occur in close clusters, and recombine after making but a few jumps. This recombination can be observed as a change of the optical absorption (compare section XXVI-1).

KOBAYASHI [1956] observed three thermal recovery steps in NaCl irradiated by 300 MeV protons at room temperature. The electrical conductivity, the release of stored energy and the optical absorption were studied together. The first stage occurs between room temperature and 150 °C, in it the resistivity goes through a maximum, presumably because positive-ion vacancies are trapped by other defects to form pairs or clusters. In the second stage (near 200 °C) the resistivity goes through a minimum and then rises again, much energy is released, and the number of colour centres decreases fastly. This process can be compared with the bleaching process (Chapter XXVI); the centres dissociate and become charged defects, which afterwards recombine. In the third stage (near 300 °C) the resistance and the absorption gradually approach the normal values pertaining to unirradiated crystals. Apparently several structural rearrangements occur together, which lead to elimination of the radiation effects. This work is still in its first stage, and the interpretation must be considered with much care.

REFERENCES

AMELINCKX, S. (1957); Dislocations and the mechanical properties of crystals, p. 3 (Wiley, New York).

AMELINCKX, S. and W. BONTINCK (1957); Acta Metall. **5**, 435.

BUEREN, H. G. VAN (1959); Physica **25**, 775.

CAFFYN, J. E. and T. E. GOODFELLOW (1955); Nature (London) **176**, 878.

CRAWFORD, J. H. and F. W. YOUNG (1960); J. applied Physics **31**, 1688

DEKEYSER, W., S. AMELINCKX, W. BONTINCK and F. SEITZ (1954); Phil. Mag. **2**, 355.

DEXTER, D. L. (1954); Rep. Conf. Defects in Solids Bristol (The Physical Society, London).

DITCHBURN (1954); Rep. Conf. Defects in Solids Bristol (The Physical Society, London).

ESHELBY, J. D., C. W. A. NEWEY, P. L. PRATT and A. B. LIDIARD (1958); Phil. Mag. **3**, 75.

FRANKL, D. R. (1953); Phys. Rev. **92**, 57.

GILMAN, J. J. and W. G. JOHNSTON (1957); Dislocations and the mechanical properties of crystals, p. 116 (Wiley, New York).

GILMAN, J. J. and W. G. JOHNSTON (1959); J. applied Physics **30**, 150.

GORDON, R. B. and A. S. NOWICK (1956a); Phys. Rev. **101**, 977.

GORDON, R. B. and A. S. NOWICK (1956b); Acta Metall. **4**, 514.

HEDGES, A. and D. MITCHELL (1953); Phil. Mag. **44**, 223 and 357.

ISHIGURO, M., F. ODA and T. FUJINO (1954); Phys. Rev. **95**, 1347.

JOFFÉ, A., M. W. KIRPITSCHEWA and M. A. LEWITSKY (1924); Zs. Physik **22**, 286.

KANZAKI, H. (1955); J. phys. Soc. Japan **10**, 663.

KANZAKI, H. (1956); J. phys. Soc. Japan **11**, 120.

KOBAYASHI, H. (1956); Phys. Rev. **102**, 348.

KURNICK, S. W. (1952); J. chem. Phys. **20**, 218.

LEIDER, H. R. (1958); Phys. Rev. **110**, 990.

LIDIARD, A. B. (1954); Rep. Conf. Defects in Solids Bristol (The Physical Society, London).

MITCHELL, J. W. (1957); Dislocations and the mechanical properties of crystals, p. 69 (Wiley, New York).

MOELLER, R. D., F. W. SCHONFELD, C. R. TIPTON and J. T. WABER (1951); Trans. ASM **43**, 39.

NYE, J. F. (1949); Proc. Roy. Soc. **A198**, 190.

NYE, J. F. and SPRACKLING (1957); Phil. Mag **2**, 772.

PRATT, P. L. (1954); Rep. Conf. Defects in Solids Bristol (The Physical Society, London).

PRATT, P. L. (1957); Vacancies and other Point Defects in Metals (The Institute of Metals, London).

RÜCHARDT (1956); Phys. Rev. **103**, 873.

SCHMID, E. and W. BOAS (1935); Metallplastizität (Springer, Berlin).

SEITZ, F. (1954); Rev. Mod. Physics **26**, 7.

SEITZ, F. and J. S. KOEHLER (1956); Solid State Physics **2**, 307.

SMAKULA and KLEIN (1951); Phys. Rev. **84**, 1043.

SUZUKI, T. (1957); Dislocations and the mechanical properties of crystals, p. 215 (Wiley, New York).

UETA, M. and W. KÄNZIG (1954); Phys. Rev. **94**, 1390.

WASHBURN, J., A. E. GORUM and E. R. PARKER (1959); Trans. Met. Soc. AIME **215**, 230.

CHAPTER XXVIII

THE ELECTRICAL PROPERTIES OF ELEMENTAL HOMOPOLAR SEMICONDUCTORS (GERMANIUM AND SILICON)

1. Electrical Conductivity

As in metals and ionic crystals, the quantummechanical stationary states of the valence electrons in homopolar semiconducting crystals fall into separate *energy bands*. In contrast to metals, however, in pure semiconductors (as well as in pure ionic crystals, see Chapter XXV) at the absolute zero of temperature the upper filled band is exactly full (valence band), and the band above it, the conduction band, is completely empty. This corresponds to the localization of all valence electrons near their particular atoms. Each atom in germanium or silicon carries 4 valence electrons, which it shares with its neighbours so as to form 4 covalent bonds each consisting of two electrons. There is no question of possible transport of electricity through the crystal by the migration of bound electrons from atom to atom: for each electron that jumps in the direction of an applied electric field, another one must necessarily jump in the opposite direction. Therefore, the substance is an *insulator*.

When the temperature is raised, however, thermal vibrations may cause the breaking of one or more bonds, and the promotion of a few electrons into the conduction band, where they are "free" to migrate through the lattice. For each liberated electron, an unpaired bond electron or *hole* is left. This can also move through the lattice, since the unpaired bond can interchange its place with a neighbouring paired one. In many aspects a migrating hole can be considered as the positively charged analogue of a migrating electron. Both the mobile hole in the valence band, and the mobile electron in the conduction band, can transport electricity: the substance has become an *intrinsic*, strongly temperature dependent (semi)conductor. Fig. 1 illustrates the conduction mechanism schematically.

In pure, intrinsic semiconductors the number of free holes is always exactly equal to the number of mobile electrons. In the band picture, the formation of a free electron with its hole corresponds to the

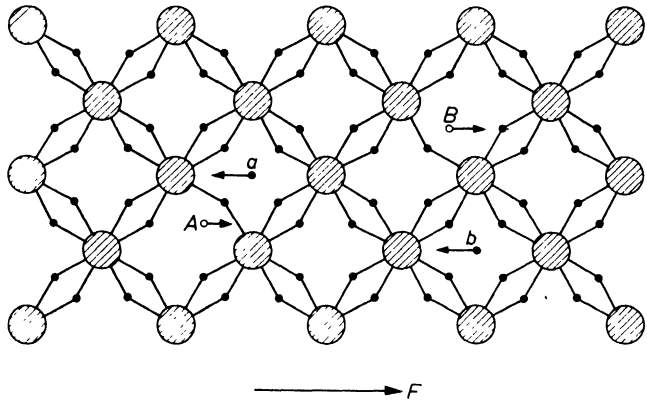


Fig. 1. Schematic picture of electron-hole conductivity in a homopolar semiconductor. The atoms are mutually connected by valence bonds consisting of two electrons (black dots) each. Valence electrons *a* and *b* orginally belonging at *A* and *B* have been liberated and thus become mobile. Holes (denoted by circles) have been left in their place, and these are also mobile. In an electric field *F*, electrons and holes will drift as indicated, both contributing to the current. The tetragonal symmetry of the three dimensional lattice has been replaced in this diagram by a simple quadratic symmetry.

excitation of an electron out of the valence band into the conduction band (Fig. 2). An unoccupied level is then left in the former, and this corresponds to the hole (a "hole" in the valence band). The excitation energy, that is the width of the *band gap* E_1, corresponds to the minimum energy needed to remove an electron from an atomic bond, as described above. Thermal excitation and thus intrinsic conduction will only give rise to a measurable conductivity if the band gap is not too wide compared to kT. At room temperature intrinsic semiconducting properties are notable only if the band gap does not exceed about 1 eV; in germanium the gap is 0.785 eV wide, in silicon 1.11 eV, and in diamond about 3 eV. The latter substance is therefore, at room temperature and even far above it, still a perfect insulator.

Not only thermal vibrations, but also foreign atoms with a number of valence electrons differing from 4 can give rise to a finite electronic conductivity, often even at very low temperatures. Consider the insertion of a trivalent atom such as gallium into the germanium lattice. Gallium has only three valence electrons to form covalent bonds with

the neighbouring germanium atoms. In terms of the above discussion, a hole must thus necessarily be present on one of the Ga-Ge bonds. This hole is not confined to the four sites next to the Ga-atom but it can, as before, interchange with another valence electron and thus migrate through the lattice. Ga-atoms hence act as generators of holes,

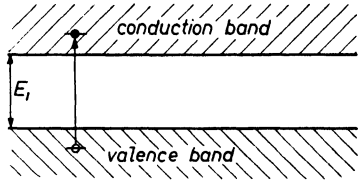


Fig. 2. Energy band scheme (very much simplified), of an intrinsic homopolar semiconductor. Thermal vibrations raise occasionally an electron from the valence band (bound state) into the conduction band (free state), leaving a (mobile) hole in the former.

but *not* of electrons. In other words Ga can accept an electron from the valence band and produce a hole: such impurities are called *acceptors*. Their presence makes the substance semiconducting; however, now holes contribute predominantly to the charge transport. These are positive carriers, and the type of conductivity is correspondingly called *p-conductivity*.

The opposite situation is obtained when a pentavalent atom such as antimony is introduced into the lattice. This atom has one of its five valence electrons to spare. This can migrate through the lattice like the hole in the previous example, giving rise to electronic or *n-conductivity*. The presence of these atoms thus leads to the introduction of an electron into the conduction band without an accompanying hole. Such impurities are called *donors*.

By careful control of the number of donors and acceptors ("doping"), germanium and silicon crystals can be prepared showing any conductivity (above the intrinsic conductivity), of either p- or n-type, that may be wanted.

In the band scheme, donors and acceptors of the kind described give rise to *localized* energy levels within the forbidden energy gap (Fig. 3). These can be occupied or unoccupied, according to an electron being present or not in the centre. At the absolute zero of

temperature a donor centre is filled, an acceptor centre is empty, according to the definition, and the impurity centres are electrically neutral. To excite the electron from a donor level into the conduction

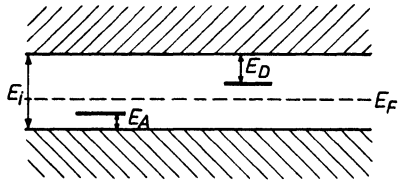


Fig. 3. Energy band scheme of an impurity semiconductor. The acceptor level is characterized by an activation energy E_A, the donor level by an activation energy E_D. Usually (but not always), the Fermi level lies somewhere about the middle of the energy gap.

band, that is, to render the extra electron mobile, a certain amount of energy is needed; also some energy must be provided to render the hole of an acceptor centre mobile, that is, to excite an electron from the valence band into the acceptor level. The distances of the donor and acceptor levels from the conduction and valence band edges respectively thus determine the temperature dependence of the *extrinsic* or *impurity* semiconductivity. These distances are usually considerably smaller than the gap width, and lead to measurable low-temperature conduction. Some acceptors can trap more than one electron, or donors can be twice or more ionized; we shall not go into these details here, nor shall we dilate on the possibility that charge transport can occur from impurity direct to impurity (impurity-band conduction).

It is a problem of quantum statistics to obtain, once the positions of the various bands and levels are known, the probability for a given energy state to be occupied by an electron or not. The statistics applicable here are the Fermi-Dirac statistics, in which the distribution function of energy is given by

$$f = \frac{1}{1 + e^{(E-E_f)/kT}} . \tag{1}$$

Here E_f denotes the Fermi energy, the position of which is determined by the condition that the material as a whole must be electrically

neutral. Of course the function f must be normalized over all possible energy states. The Fermi energy takes the place of the thermodynamic potential in the solid, and must be the same everywhere when the solid is in thermodynamic equilibrium.

Usually (but not always) the Fermi level lies somewhere midway in the energy gap (Fig. 3). Since kT is 0.1 eV or less in practical cases, that is, always much less than the width of the energy gap in Ge or Si, the distribution function of the fraction of the electrons which are in the conduction band, respectively of the fraction of holes which are in the valence band, reduces to the classical Boltzmann distribution. The total concentration of electrons in the conduction band of such a *nondegenerate*, not necessarily intrinsic semiconductor is then given by:

$$n = N_c \, e^{-\frac{(E_c - E_f)}{kT}}, \tag{2}$$

and that of holes by

$$p = N_v \, e^{+\frac{(E_v - E_f)}{kT}}. \tag{3}$$

N_c and N_v denote the densities of states available in the conduction and valence band respectively, E_c the energy of the lower edge of the conduction band, E_v that of the upper edge of the valence band. The product of the concentrations of electrons and holes is constant, namely

$$np = N_c N_v \, e^{-E_i/kT} = n_i^2, \tag{4}$$

where E_i denotes again the energy gap and n_i is the concentration of either holes and electrons in the *intrinsic* semiconductor with the same band width. If the material is extrinsic, with N_A acceptors and N_D donors per unit volume, at *moderate* temperatures where *all* centres are ionized, the number of electrons in the conduction band of an n-type semiconductor is given by

$$n \approx N_D - N_A, \tag{5}$$

and there are no holes in the valence band. At very *high* temperatures we of course again obtain the intrinsic relation $n = p \approx n_i$. At very *low* temperatures, where the distance between donor levels and conduction band becomes large compared to kT, one can derive the

following expression valid for an n-type semiconductor:

$$n = \sqrt{\frac{N_D N_c}{2}}\, e^{\frac{-(E_c - E_D)}{2kT}} \tag{6}$$

where N_c denotes again the number of levels available in the conduction band (2 times the number of atoms per cm^3 in the monatomic diamond-structure crystals) and E_D is the energy distance from donor level to that band. In a p-type semiconductor an analogous expression holds for the number of holes in the valence band.

The dependence of the energy E of an electron or hole on the "crystal momentum" or wave vector $\boldsymbol{k}$ is rather intricate in germanium and silicon, owing to the complexity of the band structure. For practical purposes, the energies close to the bottom of the conduction band or to the top of the valence band are nevertheless often described by the effective-mass approximation, in which a simple scalar parabolic E–k relationship resembling that for a free electron is adopted. For an electron in the conduction band one writes

$$E = E_c + \frac{\hbar^2}{2m_e} k^2 , \tag{7}$$

and for a hole in the valence band

$$E = E_v - \frac{\hbar^2}{2m_h} k^2 , \tag{8}$$

Here m_e and m_h must be determined from experiment. These quantities are the so called *effective masses* of the electrons and holes respectively, they can differ very appreciably from the mass of a free electron and depend on the crystallographic orientation of the electric field. The actual numerical values to be used to describe different physical phenomena also vary considerably, indicating that a tensorial energy-momentum relation should really be preferred.

Transitions from the valence band to the conduction band can be induced by optical excitation, as was the case for polar crystals. Below a certain wavelength, corresponding to the width of the energy gap, germanium and silicon are opaque, at longer wavelengths they are highly transmittant (no absorption). The transition lies in the near infrared, it is not sharp but has a characteristic structure, partly owing to the absence of a well-defined bandwidth. The transparency

for infrared light is of much interest for the study of defects, as it offers the possibility for the photoelastic investigation of internal strains and for the study of decorated dislocations.

The wavelength of the optical absorption edge and the width of the energy gap depend on the external pressure. A change of the lattice constant is accompanied by unequal changes of the edges of both energy bands; hence the energy gap varies in width. This results inter alia in the scattering of charge carriers by lattice strains. A theory for this "deformation potential" has been developed by BARDEEN and SHOCKLEY [1950], and extended by HERRING and VOGT [1956], who also take the variation of the band edges with shear strains into account.

The electrical conductivity of semiconductors is determined by two quantities; the number of charge carriers (n or p or both) and the *mobility* of the carriers. The latter quantity is defined in the same way as in Chapter XXV, that is, according to the formula

$$\sigma = e \sum_k n_k \mu_k \tag{9}$$

where the summation extends over the various charge carriers.

We have seen before that the n_k are determined by the concentration of donors and acceptors and by the temperature. μ_k is determined by the scattering properties of the crystal lattice for electrons or holes. If a single isotropic *relaxation time* τ describes the decay in time of deviations from the equilibrium Fermi-distribution (by the influence of the lattice vibrations), thus if:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{f - f_0}{\tau}, \tag{10}$$

one has for the mobility limited by the lattice vibrations:

$$\mu_k = \frac{e}{m_k} \tau \approx \frac{e}{m_k} \frac{l_k}{v_k}. \tag{11}$$

Here l_k is the mean free path between collisions, which is inversely proportional to the temperature at not too low a temperature, and v_k is the mean thermal velocity of the carriers. This, for a nondegenerate semiconductor, is proportional to $T^{1/2}$. Thus, for lattice scattering only, one obtains

$$\mu \sim T^{-3/2}. \tag{12}$$

Experiments confirm this relation fairly well in n-germanium but not in n-silicon, p-germanium and p-silicon. When charged impurities are present, these also scatter the charge carriers and relation (10) is invalidated; when this scattering dominates, the mobility *in*creases with increasing temperature. Scattering by neutral impurities and uncharged dislocations results in a more or less temperature-independent reduction of the mobility.

According to (9), the dependence of σ on the temperature is predominantly governed by the carrier density, and is therefore in most cases approximately exponential. It is difficult to determine n and μ separately from the conductivity alone. But by measuring the *Hall effect* (transverse voltage across a current-carrying conductor in a magnetic field), one can affect their separation. For a nondegenerate semiconductor with one kind of carriers (in the approximation of spherical energy surfaces and lattice scattering only), the Hall constant R is given by

$$R = \frac{3\pi}{8}\,\frac{1}{nec}\,, \tag{13a}$$

or

$$R\sigma c = \frac{3\pi}{8}\,\mu\,. \tag{13b}$$

In other, more intricate cases the latter relation retains approximate validity, apart from slight changes in the value of the numerical factor. Often the quantity

$$\mu_M = \text{“}\left(\frac{3\pi}{8}\right)\text{”}\,\mu$$

is called the Hall mobility.

Really detailed information about the properties of charge carriers (their effective mass and their mobility) can only be obtained by more complete electrical measurement, e.g. by measuring the magnetoresistance, or by high-frequency studies (cyclotron resonance).

Measurement of the Hall effect enables us also to determine the activation energy for the ionization of donors or for the filling of acceptor levels. By plotting the logarithm of the observed Hall constant as a function of the inverse temperature, one obtains effectively a plot

of log n or log p versus $1/T$. From the slope of this curve, if it is straight, the activation energy can be immediately derived.

Fig. 4 gives an example of this for n-germanium containing trivalent

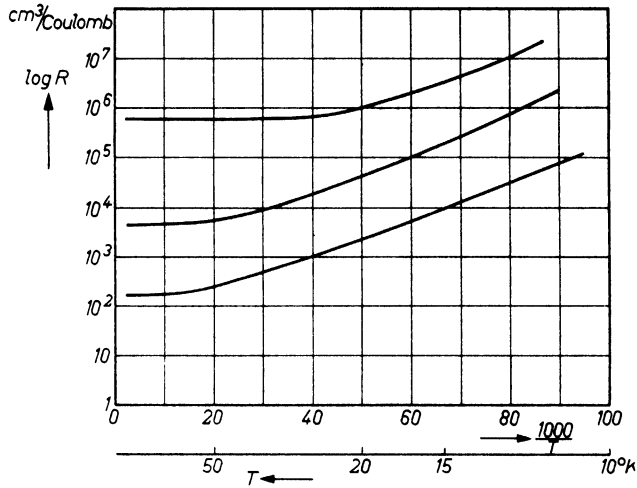


Fig. 4. Hall constant as a function of the inverse temperature for n-Ge doped with different amounts of arsenic (after DEBYE and CONWELL [1954]).

arsenic as an acceptor at three different concentrations. At "high" temperatures, R is essentially constant. This expresses the fact that the donors are completely ionized. Below $T = 50$ °K log R begins to increase and the curve soon attains a fairly constant slope, corresponding to an energy of activation of about 0.01 eV. Not only this energy, but also the absolute concentration of the donors can be determined, namely from the height of the constant part of the curve.

Charge carriers of the type not normally present in equilibrium, for instance electrons in p-germanium, can be *injected* into the crystal by applying a suitable tension to a p–n barrier: a thin region separating n- and p-regions. These *minority carriers* tend to recombine as soon as possible with the equilibrium carriers, that is, they have only a limited *lifetime*. Partly recombination occurs directly: electron-hole recombination, but it takes place far more effectively at specific *recombination centres*. The rate of direct recombination is determined by the possibility that the energy released during this process is taken up by the lattice in a suitable way, in the form of *phonons*, or is radiated in the form of *photons*. The first process is rather difficult, as many

phonons have to be excited simultaneously. Radiative direct recombination is therefore much more probable. Nevertheless, the theoretical lifetime of minority carriers in germanium and silicon is usually much longer than that observed in practice. This is due to the influence of impurities and lattice imperfections that also act as recombination centres. Especially those imperfections that have associated with them localized energy levels in the forbidden band, can serve as traps for electrons and holes and thus appreciably shorten their lifetime. Copper and nickel are typical examples of very effective impurity recombination centres in germanium, and dislocations can also act in this way.

The trapping action of impurities or imperfections has also notable effect on the characteristics of photoconductivity and its decay. It would carry us to far to discuss this here.

2. Effect of Dislocations on the Electrical Properties of Germanium and Silicon

Dislocations in homopolar crystals usually have along their axis a series of unpaired valence bonds. We postpone the full description of this interesting aspect to the next chapter, but consider here already the effect of such a linear array of *dangling bonds* on the electrical properties of semiconductors. In the first place, the dangling bonds along dislocation lines undoubtedly exert some influence on the *carrier concentration* of semiconductors. They can either take up an additional electron, therewith reconstituting the 8-electron-configuration around the Ge or Si atoms on the edge of the extra halfplane, or they can give off the remaining bound electron to the conduction band. In the latter case the dislocation would act as a row of donors, in the former as a row of acceptors. Since the influence of dislocations on the electrical properties of p-germanium has been proved to be very much smaller than on n-germanium (see Chapter XXX), SHOCKLEY [1950] concluded that the donor action of dislocations in Ge is negligible. Under practical circumstances a dislocation than only acts as a row of acceptors.

The absence of donor action means that the energy required for removing the remaining electron from the free bond must be larger than the largest possible distance between the Fermi level and the bottom of the conduction band. The position of the Fermi level is determined by the depth of the chemical acceptor levels (Fig. 3); when these are all emptied, the Fermi level cannot sink lower. It is then

usually quite near to the top of the valence band, and the dislocation donor levels must therefore be expected to lie even closer to that top or perhaps *in* the valence band.

If all the acceptor levels are occupied, the dislocation line is negatively charged, and this has again a strong effect on the *scattering* of the other conduction electrons. READ [1954] has studied the effect of the *spatial correlation* existing between acceptors arranged along a dislocation line. The gain in potential energy obtained by filling the acceptor levels along a straight dislocation increases proportionally to the fraction occupied. However, the repulsive energy between the extra negative charges increases quadratically with this fraction. Thus somewhere at an occupied fraction less than unity, the two influences will balance. Therefore, even at very low temperatures the dislocation acceptor levels are never fully occupied. At higher temperatures the filling fraction of course decreases even further.

Consider a sample of n-germanium with the Fermi level very close to the bottom of the conduction band. The fraction of occupied dislocation sites will be denoted by f_d. At the absolute zero of temperature all electrons are either localized on impurity donor levels, in the valence band or on dangling dislocation bonds; there are none in the conduction band. Assuming that the electrons for the dangling bonds must be provided by the centres lying within a cylinder of sharply defined radius R around the dislocation (*space charge* cylinder), one has from the ever prevalent condition of electrical neutrality:

$$f_d = \pi R^2 a (N_D - N_A), \tag{14}$$

where a is the mean distance between dangling bonds along the dislocation line and N_D and N_A are the volume concentrations of donors and acceptors. The cylinder R contains the equivalent amount of excess positive charge. At the absolute zero its radius is indeed sharply determined as there are no mobile carriers to screen off this space charge. The trapped electrons are then evenly spaced along the dislocation line. When the temperature is raised, the boundary becomes vague since in the outer regions of the cylinder one must now replace $(N_D - N_A)$ by the more general expression $(N_D - N_A - n + p)$. Further, the radius decreases when f_d decreases with increasing temperature.

The electrostatic potential φ in the sharply bounded cylinder is

approximately that of a continuous line charge along the axis of a continuous cylindrical positive charge distribution. We thus have at the distance r ($\gg a$):

$$\varphi = -\frac{f_d e}{\varepsilon a}\left[\ln\frac{R^2}{r^2} + 1 - \frac{r^2}{R^2}\right], \tag{15}$$

where ε is the dielectric constant.

For the total work done in building a space charge cylinder associated with an occupied fraction f_d (equally divided over the number of dislocation bonds), Read derived by the aid of the previous formula:

$$E_s = \frac{f_d e^2}{\varepsilon a}\left[\ln\frac{f_d R}{a} - \text{const}\right]. \tag{16}$$

Let us now assume that each dangling bond can take up one electron, that is, acts as a single acceptor level, and let this level be situated the amount E_2 above the top of the valence band at the dislocation (Fig. 5). The position of this top with regard to the Fermi level will

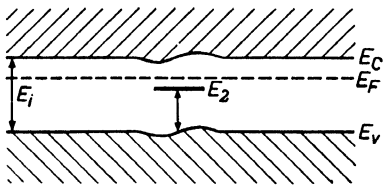


Fig. 5. Acceptor level of dislocation in semiconductor. E_2 is its activation energy, E_f denotes the Fermi level. The latter is straight whereas the band edges E_c and E_v are slightly buckled, owing to the elastic strains around the dislocation.

near the dislocation differ from elsewhere in the crystal, since the elastic strain around the dislocation causes a shift in the band edge (section 1). We shall neglect this here. The decrease of free energy in transferring *one* electron from the free state, characterized in the mean by the Fermi energy E_f, to a dislocation state, characterized by the energy E_2, is:

$$\Delta F = E_2 - E_f + \mathscr{E}(f_d). \tag{17}$$

where $\mathscr{E}(f_d) = \partial/\partial f_d(f_d E_s)$ is the increase of the space charge energy due to the addition of one electron at the occupation level f_d. The

equilibrium configuration is attained when this quantity $\Delta F = 0$, thus when

$$E_f - E_2 = \frac{\partial}{\partial f_d}(f_d E_s) = \mathscr{E}\,. \tag{18}$$

Inserting (16) and (14) one obtains approximately

$$\mathscr{E} = \frac{f_d e^2}{\varepsilon a}\left[3 \ln f_d - \ln\{a^3(N_D - N_A)\} + \text{const.}\right]. \tag{19}$$

The variation of $\mathscr{E}$ with f_d has been depicted in Fig. 6 for the following example: germanium with $N_D - N_A = 10^{15}$ cm^{-3}, $a = 4$ Å, $E_f = 0.77$ eV (0.01 eV below the conduction band) and $E_2 = 0.55$ eV (acceptor level 0.23 eV below the conduction band). For this situation, Read computed $f = 0.11$ (one electron per 36 Å) and $R = 2.8 \times 10^{-5}$ cm at the absolute zero of temperature. Although the Fermi level lies considerably above the dislocation acceptor level, yet but 11% of the

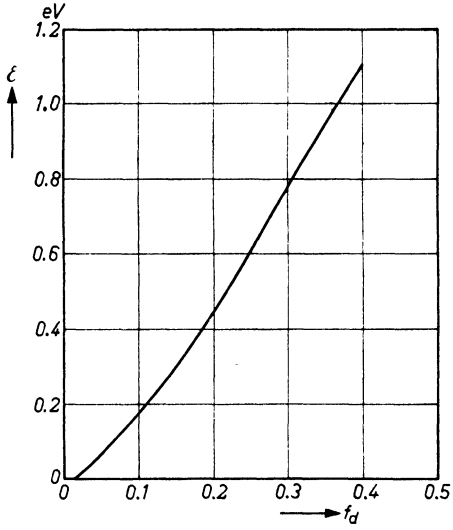


Fig. 6. Variation of $\mathscr{E}$ with f_d for the example mentioned in the text (after READ [1954]).

available dangling bonds have accepted a second electron: the rapid increase in electrostatic repulsive energy prevents their filling to a greater extent. It follows that dislocation acceptors, evidently owing to

their correlated spatial arrangement, cannot be described by simple Fermi statistics, which are so well applicable to randomly spaced zero-dimensional impurities.

The above discussion applies only to very low temperatures. Read has also extended the theory leading to formula (19), to higher temperatures. There the electrons are no longer evenly spaced along the dislocation, since the entropy of the distribution now enters also. Further, the (intricate) variation of E_t with T must be considered. Several different approximative methods were used to estimate the temperature dependence of f_d; the results have been depicted in Fig. 7, which pertains again to the previously mentioned example. It is seen that at room temperature only very few of the dislocation acceptors are still occupied.

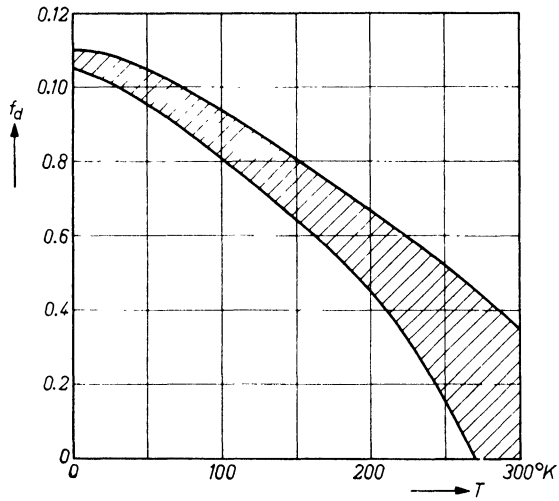


Fig. 7. Extrapolation of dislocation occupation function f_d to higher temperatures. The hatched zone contains the various approximative relations derived by READ [1954].

Apart from their influence on the density of charge carriers, dislocations also affect the mobility, simply by scattering the conduction electrons or holes. This scattering is largest in n-germanium at low temperatures, presumably because of the existence of the space charge regions around the dislocation lines. The mobility of an electron moving parallel to an array of dislocation lines will, of course, only be slightly influenced. The possibility must not even be excluded that electrons

can be transported with exceptional ease *along* the dislocation lines, by jumping from free bond to free bond, in which case the dislocations act in principle as high-conductivity channels. This has not been observed in practice, presumably because of the blocking effect of impurity atoms locked to the dislocations. Electrons moving more or less perpendicularly to the dislocation lines are either scattered by reflection against the space charge cylinder (at low temperature or when the density of dislocations is high) or they are compelled to take winding paths through the dislocation forest (in the case of a low density or a high temperature); their effective mobility is in both cases reduced. Recent experiments have confirmed many of these theoretical predictions.

It will be realized that the treatment of dislocations as rows of acceptors illuminates only one aspect of the problem. The dislocation strains, and their anisotropic distribution, affect the motion of carriers as well. DEXTER and SEITZ [1952] have treated this aspect in first approximation and found the effect to be small under normal circumstances (see also section XXX–7.3).

Dislocations affect the lifetime of minority carriers to a very large extent (see section XXX–7.4). Not much is known of the way this is brought about. A strong anisotropy has been observed by BELL and HOGARTH [1956] in that the lifetime of minority carriers moving parallel to dislocations in polygonized germanium was considerably less affected by the dislocations, than that of carriers moving in other directions.

3. Point Defects

A general theory of the electrical effects of vacancies and interstitial atoms in homopolar semiconductors does not exist. Atoms which can easily be ionized, and are known to be partially present interstitially in germanium and silicon (such as copper, nickel and lithium) are expected to act as donors. Not much can be inferred theoretically about the depth of the donor levels, and thus about the observability of this donor action. Lithium was indeed found to be a donor (SHOCKLEY [1950], REISS *et al.* [1956]) but copper and nickel are, as interstitials, very probably electrically inactive (PENNING [1958]). Substitutionally placed, however, these metal ions act as acceptors. About interstitial atoms of germanium itself not much is yet known (see also

the following chapters). Vacancies are naturally expected to be acceptors. There exists conflicting evidence as to the number of acceptor levels, and to their position, associated with vacancies in the elemental semiconductors. We shall postpone the discussion to Chapters XXX and XXXI.

GENERAL REFERENCES

[A] SHOCKLEY, W. (1950); Electrons and holes in semiconductors (Van Nostrand, New York).

[B] FAN, H. Y. (1955); Valence semiconductors, in Solid State Physics, Vol. 1, 283 (Academic Press, New York).

SPECIFIC REFERENCES

BARDEEN, J. and W. SHOCKLEY (1950); Phys. Rev. **80,** 72.
BELL, C. R. and C. HOGARTH (1956); J. Electronics and Control, **3,** 455.
DEBYE, P. P. and E. M. CONWELL (1954); Phys. Rev. **93,** 693.
DEXTER, J. L. and F. SEITZ (1952); Phys. Rev. **86,** 964.
HERRING, C. and E. VOGT (1956); Phys. Rev. **101,** 944.
PENNING, P. (1958); Philips Research Rep. **13,** 17.
READ, W. T. (1954); Phil. Mag. **45,** 775, and subsequent papers.
REISS, H., C. S. FULLER and F. J. MORIN (1956); Bell Syst. Techn. J. **35,** 535.

CHAPTER XXIX

DISLOCATIONS AND OTHER IMPERFECTIONS IN HOMOPOLAR CRYSTALS WITH THE DIAMOND STRUCTURE

In this chapter we shall discuss some fundamental properties of defects, particularly of dislocations, in the homopolar material germanium. We shall, in particular, mention work of HORNSTRA [1958] on the dislocation structure and other geometrical properties of imperfections in the diamond lattice, the electrical properties having already been dealt with in the previous chapter. To extend the discussion to other materials would lead us too far.

1. Dislocations in the Diamond Structure

1.1. GENERAL

Homopolar substances are characterized by directed valence bonds between the constituent atoms. The structures in which they crystallize are largely determined by this tendency. In the materials we are particularly interested in, the constituent atoms have each four valencies, which are symmetrically directed in space. Accordingly, each atom in the perfect crystal lattice is surrounded by four neighbours situated at the corners of a regular tetrahedron. The cubic unit cell of the so-called diamond structure that results in this way for a monatomic crystal is shown in Fig. 1: the valence bonds form ⟨111⟩-directions. In Fig. 2 the structure is represented in a simplified two-dimensional manner. One set of {111}-planes are horizontal, and the lattice is projected on the $(1\bar{1}0)$-plane through the vertical [111] bonds parallel to the plane of drawing. From this and the foregoing diagram it is seen that the atoms of the lattice can be regarded as belonging to two distinct groups, even if we have to do with a monatomic crystal. Atoms of the first group have a [111]-bond vertically upwards,

atoms of the second group have such a bond vertically downwards. The octahedral lattice planes occur in pairs, each component layer containing only atoms of one kind. The connection between the layers of one pair is formed by one vertical bond per atom; two successive

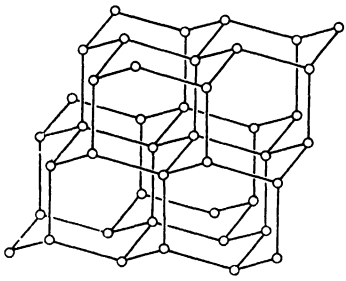


Fig. 1. Diamond lattice with one of the four ⟨111⟩-bond directions vertical.

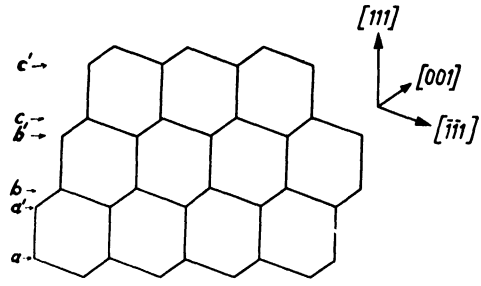


Fig. 2. Two-dimensional diagram of the diamond structure (projection of Fig. 1 on the (1$\bar{1}$0)-plane). *aa'* etc. indicate (111) double layers.

double layers are connected by three bonds per atom. This occurrence of double layers is of great importance for the mechanism of defect formation. We have already encountered a similar situation in ionic crystals. There, however, the occurrence of paired layers was a result of the condition of electrical neutrality, rather than a purely geometrical necessity. In compounds such as InSb, the two kinds of atoms fill separate atom layers of each double layer.

In organic chemistry, the typical staggered hexagon configuration of the atoms evident from Fig. 1 is called the *chair* form. The reverse configuration in which both tips of each hexagon point to the same side is called the *boat* form. It will be encountered when the crystal contains certain dislocations of predominantly screw type.

1.2. SIMPLE DISLOCATIONS

The shortest translation vectors connect nearest neighbouring atoms of the same type along ⟨110⟩-directions. In the notation of Chapter III these vectors are $\frac{1}{2}a\langle 110\rangle$. Most densely packed planes in the lattice are the octahedral planes. It is therefore to be expected that dislocations in the diamond structure will have $\frac{1}{2}a\langle 110\rangle$-Burgers vectors and {111}-glide planes, and it has indeed been observed that plastic glide of substances crystallizing in this structure invariably takes place

on these systems (see also Chapter XXX). Since an arbitrary path in the lattice can always be decomposed into a series of steps in successive ⟨110⟩-directions, the simplest types of dislocation in the diamond lattice are those with axis and Burgers vector both along the same or different ⟨110⟩-directions. There are three such *simple dislocations*: a screw dislocation, the "60°-dislocation" with {111}-glide plane, the Burgers vector of which makes an angle of 60° with the axis; and an edge dislocation with a {100}-glide plane. These three simple dislocations are shown in Figs. 3 to 5 respectively. From an inspection of these diagrams, two characteristic properties of dislocations in the diamond structure become evident. 1) The bond directions of the atoms in the core of the dislocations deviate considerably from their normal tetrahedral orientations. This occurs most conspicuously in the screw dislocation; the hexagons along the axis have assumed the boat form

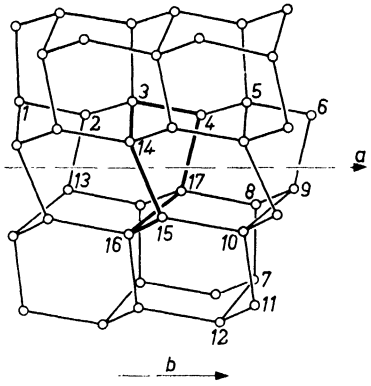


Fig. 3. Simple screw dislocation along [$\bar{1}$10]. This and the following figures have been first published by HORNSTRA [1958]. In them a and b denote axis and Burgers vector of the dislocation respectively.

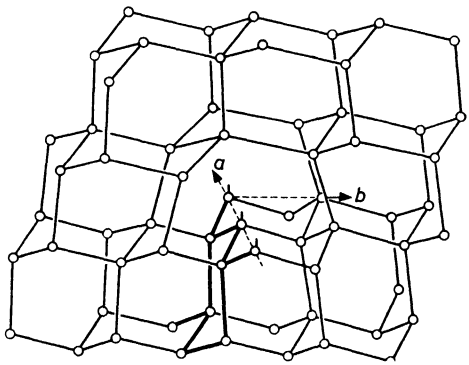


Fig. 4. Simple 60°-dislocation along [$\bar{1}$10] with (111)-glide plane. Heavy lines denote the extra half plane.

instead of the chair form. 2) Along the axis of dislocations of not pure screw character, part of the atoms have unsaturated bonds attached to them, namely those atoms that form the edge of the extra-half plane which is always associated with an edge type dislocation. These *free* or *dangling* bonds, as they were first called by Shockley, can act as acceptors for electrons and determine to a large extent the electrical and also the mechanical properties of dislocations, as we

have seen in Chapter XXVIII, and shall come back to in Chapter XXX. They form the most conspicuous feature of dislocations in the valence semiconductors.

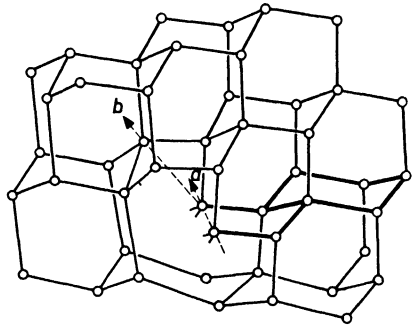


Fig. 5. Simple edge dislocation along [$\bar{1}$10], with (100)-glide plane. Heavy lines denote the extra half plane.

Although it appears at first sight as if only edge or semi-edge dislocations had free bonds, and screw dislocations had only large distortions of the bond structure associated with them, this is not the case. HORNSTRA [1958] has shown that various processes of *"core-rearrangement"* in a dislocation are conceivable, by which the relative importance of the amount of distortion, and the number of dangling bonds respectively, can be varied.

A possible rearrangement of the screw dislocation of Fig. 3 is obtained by breaking the bonds between the atoms marked *2* and *3*, *4* and *5*, etc. In this way, a certain relaxation of the deformed lattice can occur, resulting in less distortion but, of course, at the cost of the presence of broken bonds along the dislocation axis. However, it is theoretically very well possible that the free bonds thus originated subsequently combine in pairs to form *double bonds* between the atoms *1* and *2*, *3* and *4*, etc. This situation has been indicated in Fig. 6. Neither large distortions nor free valencies are now present.

Both rearrangement processes described concern only the *electronic* configuration near the core atoms, no migration of the atoms themselves being necessary to produce them. The edge dislocation of Fig. 5 or any other predominantly edge-type dislocation can, on the other hand, also be rearranged by *removing* the atoms carrying the free bonds. Then, for instance, a configuration as shown in Fig. 7 is obtained, which corresponds to the edge dislocation of Fig. 5, now, however,

without any free bonds. Instead, somewhat more severe distortions of the bond directions occur, giving rise to five- and seven-rings instead of the atomic hexagons. This kind of rearrangement can take place

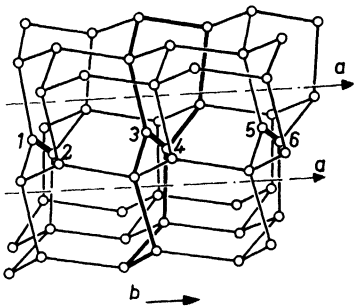


Fig. 6. Electronically rearranged simple screw dislocation with double bonds.

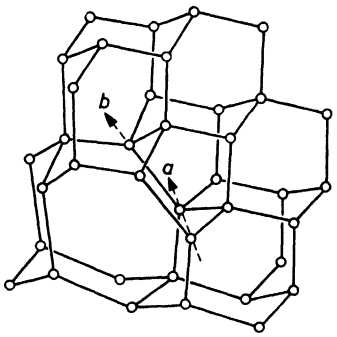


Fig. 7. Diffusionally rearranged simple edge dislocation without broken bonds.

only by *diffusion* of vacancies or interstitial atoms; either it or its reverse process will therefore only take place at elevated temperatures. The stable situation, i.e. the configuration of lowest free energy, might be that without the free bonds. Well-annealed crystals must then be expected to possess edge dislocation in this configuration.

Both examples of rearranged dislocations have lost their easy mobility by the process. The edge dislocation can only glide after a certain diffusion process has taken place, or it can move nonconservatively such as a jogged screw dislocation in metals, leaving point defects in its wake. A screw dislocation with double bonds first needs to transform into the normal configuration of Fig. 3 to retain its mobility. Both processes require an *energy of activation*. This has an important bearing on the process of plastic deformation (Chapter XXX).

Rearrangement of a 60°-dislocation (Fig. 4), the most elementary form of a dislocation in the diamond lattice, is not feasible. However, as we shall see below, a 60°-dislocation can easily be "extended" into partial dislocations that in their turn may become rearranged. Therefore, 60°-dislocations can also exist in relatively immobile forms.

HAASEN [1957] has pointed out that the bonds directly above the extra-halfplane of a 60°-dislocation are very severely stressed and may easily be broken. If that occurs, such a dislocation contains a micro-

crack along its core which, in motion, has to be dragged along by diffusion. The mobility of such a dislocation would therefore also be impeded by a diffusion process. The activation energy of this may of course very well be considerably smaller than that of free-vacancy diffusion, but it should nevertheless be appreciable.

The general question whether it is more favourable from an energetic point of view, for a dislocation to possess free bonds or strongly distorted bonds along its axis has not yet been answered. Except for cases like that of the edge dislocation where it is practically certain that the configuration of Fig. 7 is most stable, it cannot be predicted beforehand whether the rearrangements described or others will actually take place or not. An accurate quantummechanical computation of the bond energy as a function of distance and orientation in materials such as silicon or germanium is required for this.

1.3. COMPOSITE DISLOCATIONS

Hornstra has also discussed the atomic structure of some more complicated types of dislocations, namely those which occur when the axis of a simple dislocation changes alternately from one ⟨110⟩-orientation to another. Then a dislocation not lying along ⟨110⟩ is obtained (of course again with a ⟨110⟩-Burgers vector). There are six distinct possible combinations of two simple dislocations with the same Burgers vector, leading to six different types of composite dislocations, all, of course, with Burgers vector $\frac{1}{2}a[110]$. Table 1 lists these possibilities.

The dislocations I and II have the same glide plane as the simple

TABLE 1

Composite dislocations having a Burgers vector $\frac{1}{2}a[110]$ (after HORNSTRA [1958])

Type	*Axis of the dislocation*	*Angle between axis and Burgers vector*	*Glide plane*
I	⟨211⟩	30°	(111)
II	⟨211⟩	90°	(111)
III	⟨211⟩	73°13′	(311)
IV	⟨211⟩	54°44′	(110)
V	⟨100⟩	90°	(110)
VI	⟨100⟩	45°	(100)

dislocations. They may, together with the latter, form part of curved dislocations or dislocation loops that originate from a Frank-Read source on a {111}-glide plane. Type I is shown in Fig. 8, it has a large screw component (hexagon in boat-form); type II is pure edge and has the same number of free bonds per unit length as the simple edge dislocation of Fig. 5.

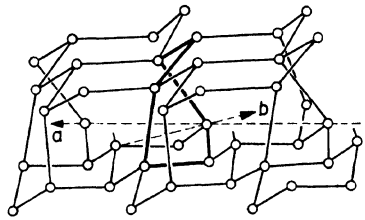


Fig. 8. 30°-dislocation (type I of Table 1) along [$\bar{2}11$], with (111)-glide plane.

The number of free bonds per unit length of different dislocations possessing the same glide plane is proportional to the sine of the angle between their axis and Burgers vector. The constant of proportionality differs for differently oriented glide planes, however.

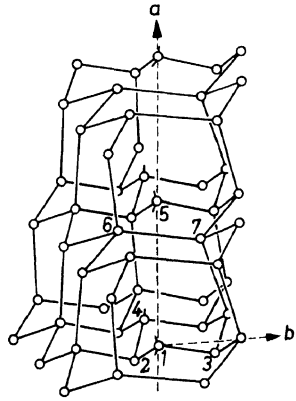


Fig. 9. 73°-dislocation (type III of Table 1) along [211] with (11$\bar{3}$)-glide plane.

A type III-dislocation can exist in two forms, the first of which is depicted in Fig. 9; atoms with one (*4*) and with two broken bonds (*1,5*) alternate along the axis. By diffusional rearrangement, that is

by removing the atoms (*1* and *5*) with two broken bonds, a configuration resembling that of Fig. 7 is obtained, although some broken bonds, e.g. those of *4*, still remain in the dislocation. Since for each atom to be removed only two valence bonds need to be broken, this process is presumably fairly easy. These atoms then become interstitial and a new bond can now be formed between the atoms *2* and *3*, etc. The interstitial atom *1* may subsequently become captured between the atoms *4*, *5*, *6* and *7*, thereby still more reducing the number of broken bonds, and then only atoms with one free bond remain. This configuration is presumably very stable (and very immobile).

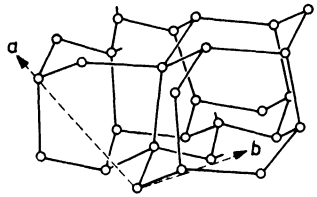
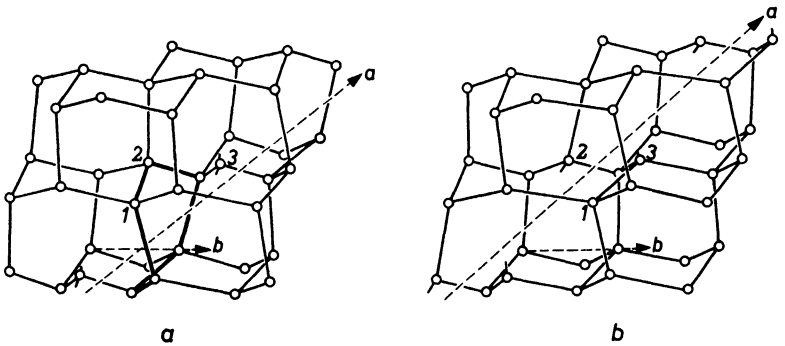


Fig. 10. "Boundary"-edge dislocation (type V of Table 1) along [001] with (110)-glide plane. This has been observed in germanium.

The dislocations IV and V are much alike. The edge dislocation with {110}-glide plane, V, depicted in Fig. 10, has been observed as a constituent of certain small-angle tilt boundaries in germanium



a b

Fig. 11. 45°-dislocation (type VI of Table 1) along [001] with (100)-glide plane. It can exist in two forms, *a* and *b*.

(VOGEL [1955]). It is the only dislocation with a glide plane differing from {111} that has actually been *observed* in germanium. Again diffusional rearrangement can take place, resulting in a configuration

without any broken bonds. Dislocation VI, the 45°-dislocation, can exist in several configurations. Starting from its intermediate position between screw and edge type (Fig. 11*a*), by electronic rearrangement a distorted chair-form structure can be formed (Fig. 11*b*). By diffusional rearrangement a dislocation without any broken bonds, resembling that of Fig. 7, can be created.

1.4. PARTIAL DISLOCATIONS AND STACKING FAULTS

The number of distinguishable partial splittings of dislocations in the diamond lattice is much greater than in the close-packed cubic lattice, because of the existence of double layers parallel to {111} instead of single octahedral planes. The two components of each double layer have the same position in the diamond structure; successive double layers differ by a rotation over $\pi/3$ around their common normal. The regular stacking order can thus be represented by the symbol

$$aa'\ bb'\ cc'\ aa'\ldots$$

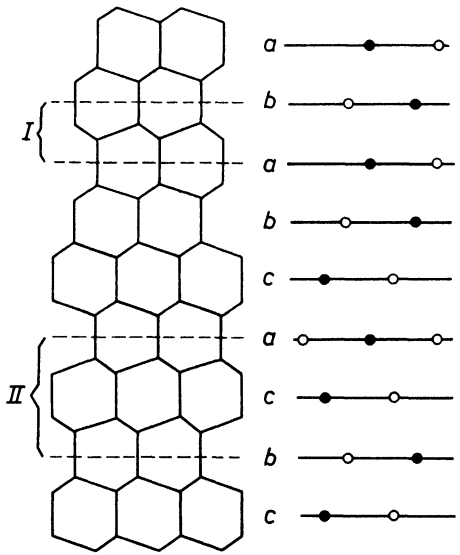


Fig. 12. Intrinsic (I) and extrinsic (II) stacking fault in the diamond lattice.

where *a* and *a'*, etc., represent the two components of the same double layer that have the same position. This notation has also been used

in Fig. 2. In Fig. 12 each double layer has for simplicity been indicated by one letter only.

A stacking fault on a {111}-plane can be constructed in two ways, corresponding to the removal or the insertion respectively of a double layer (Fig. 12), analogous to the formation of intrinsic and extrinsic stacking faults in the face-centred cubic lattice (Chapter III). Each stacking fault corresponds to a thin twin lamella (of width one or two lattice constants respectively).

Where a stacking fault ends, a partial dislocation occurs. This, now, can have several structures, depending on the type of stacking fault it is associated with (intrinsic or extrinsic) and its Burgers vector. The partials associated with a 60°-dislocation (with Burgers vector $\frac{1}{2}a[110]$) are the analogues of the Shockley-partials in f.c.c. lattices; they have Burgers vectors $\frac{1}{6}a[211]$ and $\frac{1}{6}a[12\bar{1}]$ (section III–1). These partials both have their stacking fault in a (111)-glide plane, but in the diamond structure an extended dislocation lying fully in one such plane is not possible. Instead, by *combination* of a perfect dislocation on one component layer and a Shockley-type partial dislocation on the

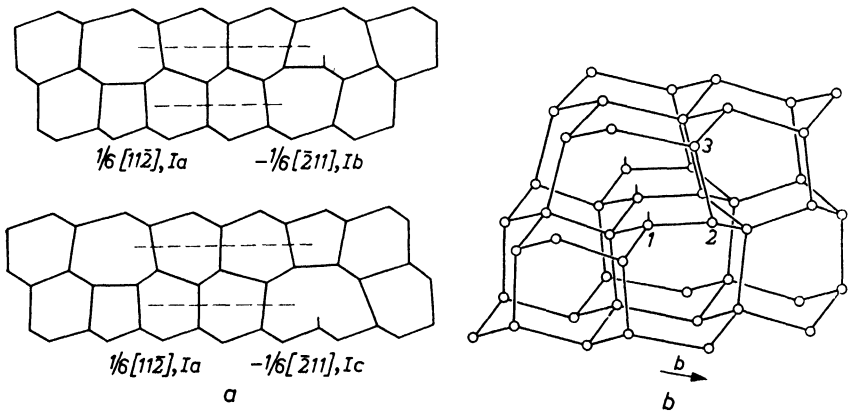


Fig. 13. *a*) Three possible atomic configurations Ia, Ib and Ic of partial dislocations with Burgers vectors along ⟨211⟩, bordering an intrinsic stacking fault on an octahedral plane. *b*) The atomic structure of the partial of type Ib.

other layer, a stable type of "extended" dislocation is formed (HAASEN and SEEGER [1958]). Three possible configurations of these Shockley partials with an intrinsic stacking fault are conceivable; these have been illustrated in Fig. 13. Electronic rearrangement of the partials is very well possible and the actual position of the free bond is very uncertain.

The stacking fault belonging to each partial need not necessarily be in the octahedral plane. If it is not, the partial dislocation is sessile (the analogue of the Frank sessile dislocation of section III–1). Partial dislocations with extrinsic stacking faults cannot easily be combined with a perfect dislocation, since the necessary atomic movements are difficult. This was already more or less the case with the intrinsic partials, since then also whole ribbons of atoms near the plane of the stacking fault must rotate through about 90° (Fig. 14). To form an extrinsic stacking fault, however, *two* ribbons on successive planes must rotate simultaneously.

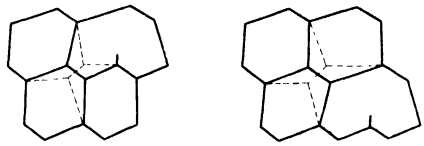


Fig. 14. First step in the combination of a 60°-dislocation with a partial. In both ways shown, ribbons of atoms have to be rotated through about 90° and bonds must be interchanged.

The glissile motion of a "combined" dislocation also needs a continuous rotation of atom ribbons. This constitutes probably an impediment for the motion of the dislocation. In contrast to face-centred cubic metals, in the diamond lattice the mobility of dislocations is thus seriously decreased by partial splitting.

A similar difficulty arises in metals with a complicated lattice, such as β-uranium (KRONBERG [1959]).

Hornstra gives in his paper an exhaustive description of various other forms of partial dislocations. In another paper [1959] he discusses the construction of grain boundaries of arbitrary misfit angle from a discrete number of partial dislocations. The reader is referred to these papers for more details.

It has already been mentioned before (Chapters III and XXIII) that crystals of germanium, silicon and other diamond-structure materials can fairly easily be grown to a very high state of perfection. Crystals containing only a few cm dislocation per cm^3 can be prepared without difficulty by the Czochralski pulling technique. Except for the simultaneous exceptional purity of these crystals, the reason why this should be so in these materials and not in metals is not yet clear

although DASH [1959] has analyzed the circumstances which must be controlled to obtain these "dislocation free" crystals.

A correspondingly degree of perfection has also been obtained in some polar crystals, such as LiF.

2. Point Defects in the Diamond Structure

Vacancies and interstitial atoms are probably as common in a homopolar crystal with the diamond structure as in metal crystals. As the diamond lattice is less dense than most simple metallic lattices, the strains associated with interstitial atoms are presumably relatively small. The cubic unit cell is only half full, and space is available for at least one interstitial atom per cell.

The properties of vacancies and interstitials are so intimately tied up with their influence on the electrical properties, that we postpone their discussion to Chapter XXX. It will be shown in that chapter that the *double* vacancy or vacancy pair also plays a very important role in these substances. This stands in direct connection with the double layer structure of the {111}-planes, described in the previous sections.

Apart from physical defects in the valence semiconductors, the chemical impurities which to a large extent determine the electrical properties, must also be considered. Instead of tetravalent silicon or germanium atoms, trivalent atoms like boron or gallium can be introduced substitutionally into the lattice, or pentavalent arsenic or antimony atoms. The former act as acceptors, the others as donors, and together they determine the type of conductivity of the material. But they can, in principle, also interact with other lattice defects such as dislocations. Not much is known about this interaction at present.

A third group of defects are the electrically inactive impurity atoms such as carbon, tin or oxygen. Even less is known about these. They may occur in relatively very large concentrations (0.1%) in otherwise extremely pure germanium or silicon crystals, containing less than 10^{-8}% electrically active foreign atoms. About their mechanical effects only speculations exist.

Oxygen in silicon can be detected by the infrared absorption it produces, and by its effect on the mechanical damping (Chapter XXXI). Arguments can be advanced that oxygen atoms are incorporated inter-

stitially, at sites between two silicon atoms, as shown in Fig. 15 (KAISER *et al.* [1956]). Stable complexes containing more than one oxygen atom can be produced by annealing, whereas indictions exist for oxygen-vacancy associations as well. The reason for the occurrence of these configurations is chemical rather than physical.

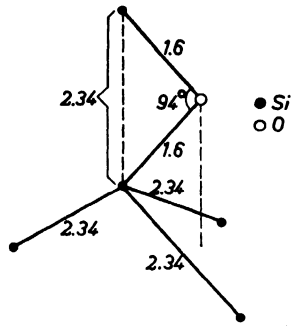


Fig. 15. Possible site for interstitial oxygen in silicon (according to KAISER *et al.* [1956]). Distances in Å.

In general, the chemical forces of interaction between impurities, imperfections, etc., are much more important in homopolar substances than in metals and in most instances dominate over all other interactions in the former. This brings an element of uncertainty in the estimations of the interaction energy.

REFERENCES

DASH, W. C. (1959); Growth and Perfection of Crystals (Wiley, New York).
HAASEN, P. (1957); Acta Metall. **5**, 598.
HAASEN, P. and A. SEEGER (1958); Halbleiterprobleme IV (Vieweg, Braunschweig).
HORNSTRA, J. (1958); J. Phys. Chem. Solids **5**, 129. (See also BUEREN, H. G. VAN, J. HORNSTRA and P. PENNING, Supplement to Nuovo Cimento **2**, 1958.)
HORNSTRA, J. (1959); Physica **25**, 409.
KAISER, W., P. H. KECK and C. F. LANGE (1956); Phys. Rev. **101**, 1264.
KRONBERG, M. L. (1959); J. Nucl. Mat. **1**, 85.
VOGEL, F. L. (1955); Acta Metall. **3**, 245.

CHAPTER XXX

PLASTIC DEFORMATION AND IRRADIATION OF MATERIALS WITH THE DIAMOND STRUCTURE, ESPECIALLY OF GERMANIUM

1. General Description of Plastic Properties of Homopolar Substances

The *plastic properties* of the semiconducting materials have become of increasingly practical importance, owing to the realization that small changes in the density of dislocations and other defects, such as are produced by mechanical handling and heat-treatment, cause large variations in the electrical and other technically important physical properties of these elements. Even their chemical properties, such as the reactivity with alloying elements, are affected. Although a minimum amount of plastic deformation cannot be avoided during the preparation and handling of the crystals, in order to obtain crystals with reproducible properties for use in electrical devices it is of primary importance to control the amount of deformation and to eliminate its influences as far as possible. Moreover, the study of plastic deformation has provided new information on the more general problem of the interaction between dislocations and impurities. Finally, the study of the plastic properties of homopolar materials is interesting in itself, since these properties differ notably from those of metals: the behaviour and specific properties of lattice imperfections in homopolar crystals are to a large extent determined by the valence-type bonding. The study of imperfections and their influences can, in this respect, give new insight into the nature and magnitude of the forces between atoms in crystals. It is for this reason that we shall discuss in this chapter the mechanical properties of valence crystals and especially those of germanium in some detail.

Germanium, as well as silicon, diamond and nearly all homopolar crystalline substances, are very brittle at room temperature. They

fracture by cleavage, usually along {111}-planes, without measurable previous deformation, at a load of 10 to 20 kg/mm². It was discovered by GALLAGHER [1952] that the same substances become progressively more ductile at elevated temperature. The actual temperature region in which the first traces of ductility can be detected varies with the nature of the test. Very carefully performed bending experiments (at a low strain rate) have shown germanium to become slightly deformable at as low a temperature as 375 °C, whereas in silicon the lowest transition temperature ever detected for bending is little above 500 °C (PEARSON, READ and FELDMANN [1957]). In tests carried out at the usual rates of strain the ductility of germanium begins above 450 °C, that of silicon above 650 °C. Under the same circumstances of temperature and rate of strain, the material behaves appreciably less ductile in tension than in bending (see also section 2.1). Diamond, well known as extremely hard and only workable in practice by cleaving ("cutting") becomes slightly ductile above 1800 °C (HAASEN and SEEGER [1958]). Indium-antimonide, another diamond-type compound, deforms above 400 °C, zincblende above 600 °C (CHURCHMAN, GEACH and WINSTON [1956]): the (absolute) transition temperatures always amount to about half the absolute melting temperatures.

As to the geometrical nature of the plastic deformation (unless otherwise stated we shall confine our attention to germanium as the most thoroughly studied representative of the group) all observers (GALLAGHER [1952], TREUTING [1952, 1955], VOGEL [1956], PATEL and ALEXANDER [1956]) agree that it occurs predominantly by glide on the {111}-planes, the glide directions being invariably ⟨110⟩-directions. That is, the dislocations responsible for the deformation are of the type discussed extensively in the previous chapter, their Burgers vectors being always $\frac{1}{2}a\langle 110\rangle$. A shear {111} ⟨110⟩ can be produced on two types of glide planes, namely between two sublayers of atoms belonging to the same double layer, which are connected by one bond per atom and spaced a distance $a\sqrt{\frac{3}{4}}$, or between sublayers of different layers; these are connected by three bonds per atom, and spaced at a three times smaller distance (compare Fig. XXIX–1). All circumstances (energy, geometry of partial dislocations) favour the operation of the first kind of glide plane. In mechanical twinning, however, the second type of shear is encountered.

Deformation by twinning has been observed in many diamond structure materials under certain, rather critical circumstances of stress,

strain rate and temperature. Twinning planes observed were again {111}, but also {123}; the directions of the twinning shear being ⟨112⟩ in the first case and ⟨124⟩ in the second. We shall postpone the discussion of this deformation twinning to section 6.

The degree of ductility, as well as the characteristics of the deformation, vary considerably with the temperature in the ductile region. In this respect the homopolar materials differ essentially from metals and polar crystals.

At "low" temperatures (400–550 °C in germanium), the plastic deformation is slow; the deformation proceeds but gradually (creep), and when the specimen is deformed at a constant strain rate, the form of the stress-strain curve and the coefficient of work-hardening both depend strongly on that rate. The influence of a thermally activated rate-limiting process is clearly present. At higher temperatures (medium region, 550–850 °C in germanium, $0.7 < T/T_m < 0.9$) the time dependence becomes less conspicuous, owing to the rapid decrease of the rate-determining time constants with the temperature. Also the rate of work-hardening decreases and the extent of the plastic strain that can be achieved increases. A few tens of degrees below the melting point, germanium and many other homopolar substances become extremely plastic; nevertheless, at very high rates of strain, the degree of ductility remains limited even then (see section 4).

We shall in the following sections discuss successively the most important aspects of the plastic behaviour, and the imperfection mechanisms underlying it, of germanium in the "low"-, "medium"-, and "high" temperature regions as defined above. Occasionally we shall mention analogous observations and conclusions pertaining to other materials.

2. Low- and Medium-Temperature Deformation of Germanium

2.1. CREEP UNDER CONSTANT STRESS

A creep curve typical for a monocrystal of germanium is shown in Fig. 1; it applies to a crystal loaded in uniaxial stress along the [110] axis, at a temperature of 470 °C. The tensile stress amounted to about 4 kg/mm^2. Three parts can be distinguished in the creep curve (VAN BUEREN, HORNSTRA and PENNING [1958], VAN BUEREN [1958]). During the first few minutes after loading, very little strain at all occurs; this period is sometimes called the *incubation period*, but we prefer

the name *initial period*, for reasons to become clear below. After it has elapsed a remarkable increase of the creep rate is observed: this quantity rises within a few minutes to a value many times larger than the average creep rate during the initial period, and remains *constant* afterwards during a second interval of time, the so-called *steady state period*. Finally, the creep rate decreases again, because of the onset of *work-hardening*. During the initial period sometimes a complete absence of deformation is observed (especially in bending tests) and one speaks of a *delay time*. Recently it was found, however (VAN BUEREN [1958, 1959]), that in carefully prepared pure specimens a real delay

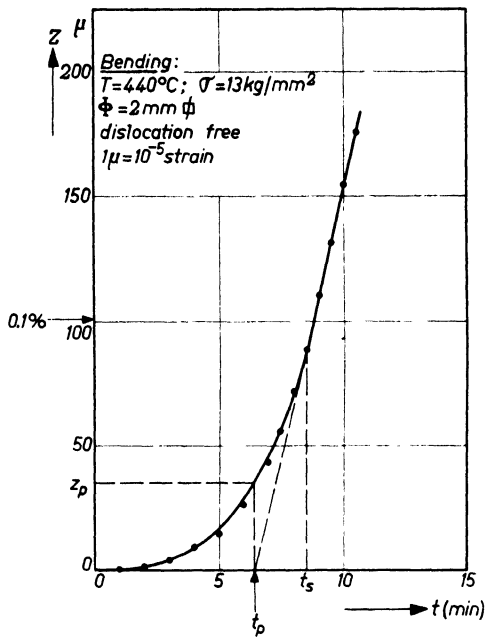


Fig. 1. Creep test of extended germanium single crystal. The tensile axis was along [110]; temperature and load were 470 °C and 4 kg/mm² respectively. The drawn curve is the theoretical curve (see section 5).

time does not exist, and that the crystals always begin to deform immediately, although slowly. The initial part of the creep curve could be well described by a *cubic* time law (see also PENNING and DE WIND [1959]).

On raising the *temperature*, the time scale of the creep phenomenon shrinks fairly uniformly: all three periods decrease in length at about

the same rate, and the constant creep rate during the steady state period increases in about the inverse ratio. That is, the vertical scale of the creep curve remains more or less unaltered, apart from a slight increase in the maximum creep strain obtained during the work-hardening period: the rate of work-hardening decreases also with

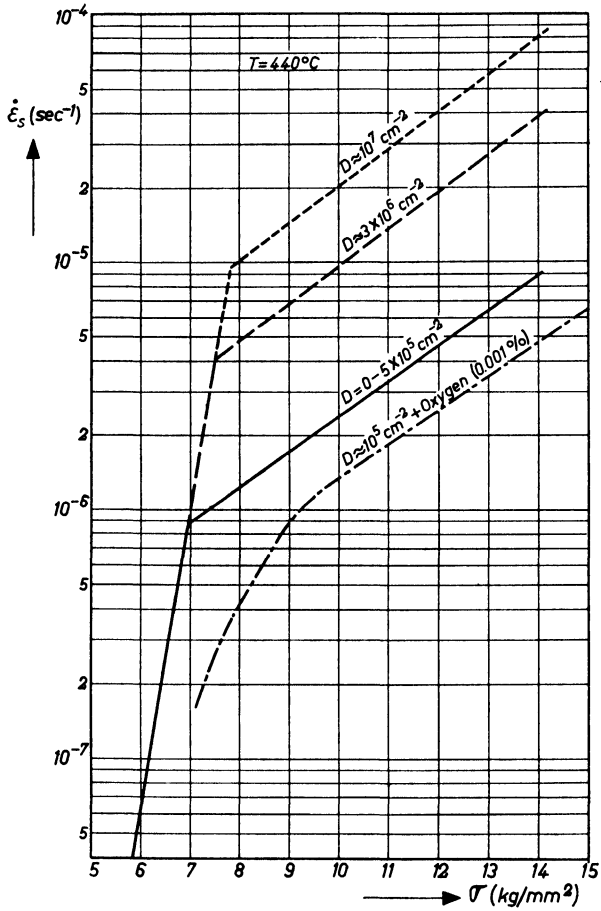


Fig. 2. Stress dependence of steady creep rate $\dot{\varepsilon}_s$ of bent germanium single crystals with original dislocation density D as indicated. The curves are the averages of about 150 observations (after PENNING and DE WIND [1959]).

increasing temperature. The temperature variation of the length of the initial period, t_p, defined as indicated in Fig. 1, and of the inverse constant strain rate during the steady state period, $(\dot{\varepsilon}_s)^{-1}$, can both

be represented in good approximation by the exponential factor:

$$\left.\begin{matrix} t_p \\ (\dot{\varepsilon}_s)^{-1} \end{matrix}\right\} \sim \exp\left[\frac{Q}{kT}\right], \tag{1}$$

where Q lies between 1.5 and 2.0 eV in germanium.

A variation of the *loading stress* also results in a fairly uniform contraction of the time scale of the complete creep curve, whereas the vertical scale is again affected to a much lesser amount. The dependence of the two creep curve parameters on the resolved shear stress σ could be expressed by van Bueren within most of the stress range by one exponential expression of the form (Fig. 2):

$$\left.\begin{matrix} t_p \\ (\dot{\varepsilon}_s)^{-1} \end{matrix}\right\} \sim \exp(-a\sigma), \tag{2}$$

which factor must be used in combination with the temperature factor (1). The constant a amounts to a few times 10^{-7} dyn^{-1} cm^2 in germanium. Penning and de Wind showed that a *transition stress* σ_t exists, below which the stress dependence of $\dot{\varepsilon}_s$ becomes much more rapid, corresponding to $a = 2.10^{-6}$ dyn^{-1} cm^2. The effect of the transition stress on t_p is much less, and maybe not present at all.

2.2. DEFORMATION AT A PRESCRIBED STRAINRATE

Quantitative measurements on the plastic behaviour of germanium stressed or strained at a constant rate have been made by PATEL and ALEXANDER [1956] and by CARREKER [1956]. They compressed and extended respectively germanium single crystals at elevated temperatures and observed a pronounced temperature dependence of the flow stress (Fig. 3) and sometimes even abrupt yielding at testing temperatures between 500 °C and 700 °C (cf. Fig. 10). The magnitude of the flow stress or the yield stress and the work-hardening coefficient decreased rapidly with increasing temperature; accurate quantitative data regarding this temperature dependence are not available, owing to the large scatter in the individual observations, but an exponential dependence corresponding to (1) but with a smaller Q seems indicated for both quantities. Above 700 °C a yield phenomenon has no longer been observed.

The occurrence of a yield point followed by a yield drop very much resembles the situation in several impure metals (Chapter IX). This analogy stretches still further, since it has been found that after

interruption of a deformation test the yield point did not return on restraining, except when during the intermission the material was kept at a sufficiently high temperature for a sufficiently long interval of time. Apparently the phenomenon of *strain-ageing* discussed in section X–1 also takes place in deformed germanium and silicon.

Bending tests with corresponding results have been carried out on thin silicon crystals and *whiskers* down to diameters less than 100μ,

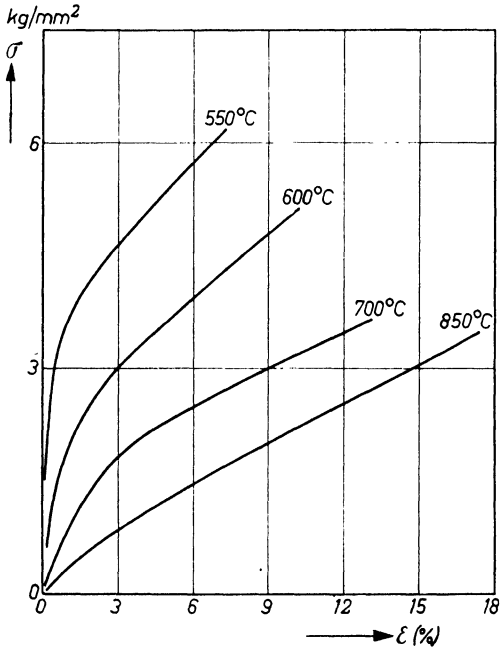


Fig. 3. Stress-strain curves at various temperatures, at a prescribed rate of strain, as obtained by PATEL and ALEXANDER [1956].

by PEARSON, READ and FELDMANN [1957]. The deformation of whiskers resembled that of macroscopic samples, in that again an initial ("delay") period in creep, and a sharp yield point in a tensile test occurred. Work-hardening was virtually absent, however. Both flow stress and yield stress decreased in the whiskers very rapidly with increasing temperature, approximately according to a T^{-5} law.

The occurrence of a "delay" or inital time in creep and a sharp yield point in the stress-strain curve are obviously interrelated; the delay period reflects the existence at a given load of a certain maximum rate of stress which must not be surpassed lest plastic flow be possible.

A certain energy-barrier appears to exist for plastic deformation, which can only be overcome with the aid of thermal vibrations. The work-hardening during a creep test and during a normal compression or tension test are also directly related to one another. Clearly either the generation or the motion of dislocations is impeded. We shall return to this question in section 5.

3. Dislocations, Slip Lines and Etch Pits

As is well known and has already been mentioned earlier (Chapter V), after etching with a suitable etchant etch pits appear on (111) or (100)-faces of Ge-crystals, which, to judge from their distribution over the surface and other characteristics, have originated at sites where a dislocation intersects the surface. The question whether all dislocations give rise to such etch pits, and whether all these etch pits correspond to dislocations, has not yet been answered satisfactorily. From certain experimental evidence about the effect of dislocations on the diffusion of impurities (PENNING [1958a], see also Chapter XXXI), as well as from a variety of other observations, it can be concluded with a fair degree of certainty, that, at least in germanium, the mutual correspondence between dislocations and "dislocation etch pits" is good to within a factor of about two.

The polished surfaces of slightly deformed germanium crystals usually show slip lines that are traces of the {111}-glide planes. Subsequent etching reveals the typical pattern of discrete etch pits which lie in rows parallel to octahedral planes. Even when, such as after slight deformation, the slip line pattern can only be very faintly resolved, the corresponding etch pit density has usually already increased very considerably over that of the undeformed material. In carefully prepared, undeformed germanium single crystals the dislocation density as derived from etch pit countings is very small compared to the density in metals; densities less than 10^4 cm^{-2} are obtainable, and under extreme precautions the production of crystals with a dislocation density of the order of one per cm^2 is possible. After a plastic strain of 0.5% at 600 °C the etch pit density has increased to 10^6 cm^{-2} (PATEL and ALEXANDER [1956]): the etch pit density was found to increase approximately exponentially with the strain. After a few percents of strain, however, it reaches a *saturation value* of the order of 10^8 cm^{-2}, whereas the distribution of etch pits then becomes homogeneous (Fig. 4).

Etch-pit counting is obviously not only a suitable tool for obtaining the dislocation density, it serves also a purpose in the detection of an eventual occurrence of small amounts of plastic glide, and in delineating the slip planes along which that glide has preferentially taken place, under circumstances where direct measurement is not possible.

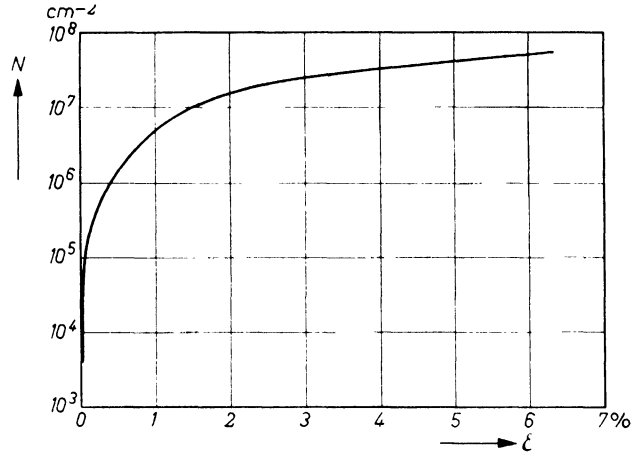


Fig. 4. Relation between density of etch pits on (111) face and strain of germanium crystal (after PATEL and ALEXANDER [1956]).

4. Deformation at Very High Temperature. Thermal Stresses

Systematic observations on the characteristics of the deformation of germanium or related substances at temperatures near their melting point are not available. One category of indirect experiments may be mentioned here however, as it bears directly on this point and shows the unusual possibilities for the observation of dislocations and other results of plastic flow in these materials by the etching technique.

Germanium single crystals pulled slowly from the melt by the so-called Czochralski technique for preparing single crystals (see e.g. ROTH and TAYLOR [1952]), usually contain 10^3 to 10^4 dislocations/cm^2. Etching reveals that these dislocations are not distributed at random: specific patterns of etch pits, depending on the crystallographic orientation, the rate of cooling and the speed of pulling, etc., can be observed, of which Figs. 5 and 6 give examples. It was assumed by BILLIG [1956] and PENNING [1958b] that these etch pits originated from dislocations formed and moved by the *thermal stresses* set up in the

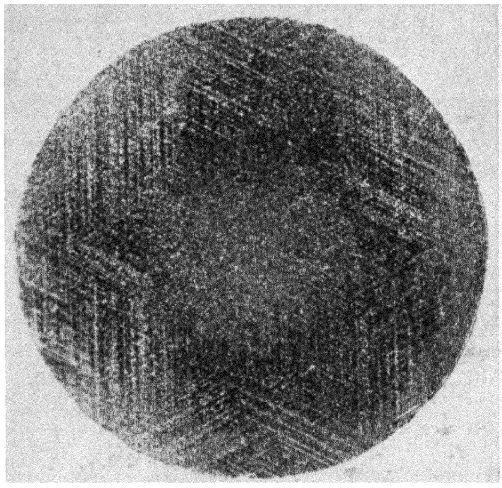


Fig. 5. Pattern of etch pits, obtained after etching a (11$\bar{1}$) cross-section of a cylindrical [110] crystal that had been cooled from 850 °C in a flow of gas (after PENNING [1958b]).

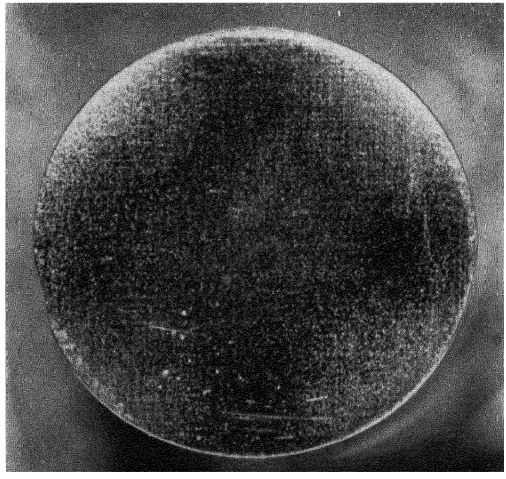


Fig. 6. As Fig. 5, now for a [100] cross-section of a (100) crystal (after PENNING [1958b]).

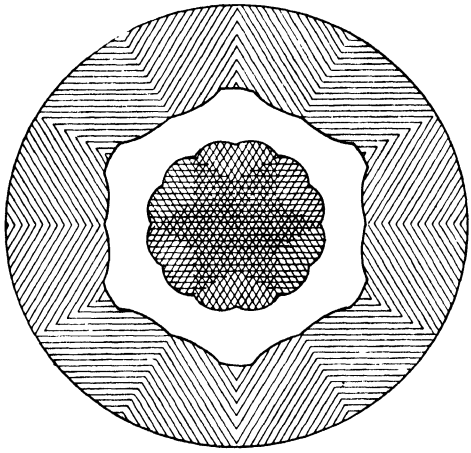


Fig. 7. Theoretically derived pattern of pits for the situation of Fig. 5 under the assumption of complete plastic relaxation (after PENNING [1958b]).

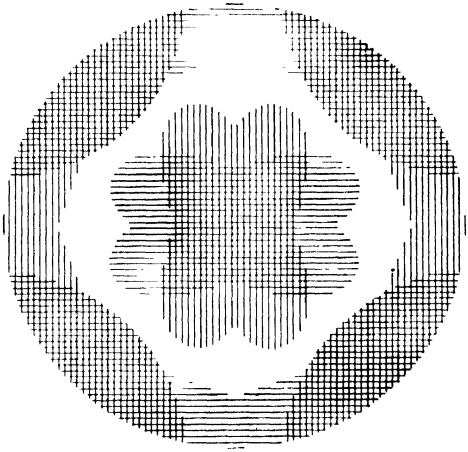


Fig. 8. Theoretically derived pattern of etch pits for the situation of Fig. 6 under the assumption of very little plastic flow (after PENNING [1958b]).

solidified rod of material during the cooling. A non-uniform temperature gradient exists in the cooling material, which results in internal stresses: the outer layers are at a lower temperature than the core, and therefore contract more than the latter. These stresses can be partly released by the generation of dislocations, since the temperature is high. The density of etch pits over the cross-section of the rod is not uniform; about halfway the centre and the circumference it reaches its minimum. Moreover, at various points of the cross-section different glide systems have evidently been operative as can be judged from the occurrence of differently directed rows of etch pits in Figs. 5 and 6.

PENNING [1958b] tried to explain the observed patterns on the basis of either of two alternative approximations. In the first one it was assumed that but a small part of the thermal stress (that is, the elastic stress that would be set up by an inhomogeneous temperature gradient in a perfectly elastic material) is released by plastic flow. Then, a deformation pattern results that is primarily determined by this *elastic* stress distribution and the geometry of the 12 available glide systems, as mentioned in section XXIX–1. The alternative assumption was that *complete plastic* stress release had taken place. The pattern of etch pits that would then be expected to occur could be derived from the principles of the mathematical theory of plasticity: the plastic strain will be distributed in such a way over the available glide systems that the total energy expended ($\int \sigma \varepsilon dV$) is a minimum. The flow pattern then no longer reflects the original elastic stress distribution, of course.

The two patterns of etch pits, to be theoretically expected on both grounds respectively, have been reproduced in Figs. 7 and 8, for the crystal orientation corresponding to that of Fig. 5 and 6 respectively. Mutual comparison of these four figures shows that the first assumption, of only a fractional release of stresses by plastic flow, yields much better agreement with the observations than the assumption of complete plastic relaxation. (Note for instance the appearance of distinct pit density minima at four symmetrical points around the circumference in Figs. 6 and 8, and the occurrence of triple glide around the circumference in Fig. 5 but not in Fig. 7.) During the time available in the first stages of rapid cooling, plastic flow, that is dislocation generation and motion, could only take place to a limited extent. In other words, even at a temperature near the melting point, the mechanism of plastic flow in germanium is still characterized by an inherent inertia (see section 5.2).

This has an important bearing on the theoretical interpretation of the low-temperature characteristics of plastic flow, as we shall discuss in the following section.

5. Dislocation Theory of Plastic Creep of Nearly Perfect Crystals

5.1. THEORY OF CREEP

An explanation of the observed characteristics of plastic creep of germanium crystals has been advanced by VAN BUEREN [1959]. The theory is not restricted to diamond structure materials, but should describe the creep of all crystals containing less than a certain maximum density of structural imperfections. It has indeed been found that the creep characteristics of LiF crystals and crystals of some refractory oxydes are very similar to those of germanium (cf. Chapter XXVII), and the considerations of this section should accordingly apply to these crystals also.

It is assumed that a number of dislocation sources to a volume density N is present in the crystal to be deformed. On the application of a stress (at an elevated temperature) these sources can emit dislocation loops, but only in time *intervals* of length

$$t_0 = \theta + \frac{\delta}{v}. \tag{3}$$

Here v stands for the (uniform) velocity of the dislocation loops, and δ for the distance such a loop must have travelled before the *back stress* caused by it at the site of the source has dropped to a level sufficiently low to enable a new loop to be formed. This distance is, in the Ge-samples tested, easily estimated (from the applied stress) as about 10^{-4} cm. The quantity θ takes care of the possible presence of an *incubation time* for the action of the source under a stress, which may be associated with the requirement of thermal activation.

The total dislocation density in the material is easily found to increase quadratically with time t:

$$\Lambda = \frac{N}{t_0} \cdot \pi v t^2, \tag{4}$$

since the expansion of the loops *and* the number of loops both increase

proportional to t. Hence, the strain rate is given by

$$\dot{\varepsilon} = Abv = \frac{\pi N b v^2}{t_0} t^2 , \tag{5}$$

which means that the initial part of the creep curve is a *cubic* parabola. The applicability of such a cubic time law has been verified in a large number of experiments, as mentioned in section 2.

The expansion of the loops cannot continue indefinitely, since ultimately they reach the surface of the crystal, and a *stationary state* develops corresponding to the constant creep rate observed in the second part of the creep curve. The time which must elapse before this stationary state begins is of the order of d/v, where d is the (transverse) dimension of the specimen (slip path of the dislocations), and the stationary creep rate is found by inserting $t = d/v$ into (5).

At the time $t_s = d/v$ or thereabouts the initial and stationary parts of the creep curve must be matched, and from this one obtains the following analytical expressions for the whole creep curve:

$$\begin{aligned} \varepsilon &= \frac{\pi N b v^2}{3 t_0} t^3 ; \qquad \left(t < \frac{d}{v} , \text{ initial cubic part}\right) \\ \varepsilon &= \frac{\pi N b d^2}{t_0} \left(t - \frac{2d}{3v}\right) ; \quad \left(t > \frac{d}{v} , \text{ stationary linear part}\right). \end{aligned} \tag{6}$$

By comparison with these expressions the parameters $\dot{\varepsilon}_s$ and t_p introduced by Penning and de Wind (section 2) and explained in Fig. 1 can now be interpreted as

$$\dot{\varepsilon}_s = \frac{\pi N b d^2}{t_0} , \tag{7}$$

$$t_p = \frac{2d}{3v} , \tag{8}$$

and

$$z_p = \varepsilon(t_p) = \frac{8\pi N b d^3}{81 v t_0} . \tag{9}$$

Here, z_p is a third parameter used by Penning and de Wind, from which they show that it is not independent, but, when divided through $\dot{\varepsilon}_s t_p$, forms a remarkably invariant quantity against variations of

stress, etc. On the average $\dot{\varepsilon}_s t_p / z_p$ was observed to be 6.7. From (7), (8) and (9) one derives immediately that in theory

$$\frac{\dot{\varepsilon}_s t_p}{z_p} = 6.75 \,. \tag{10}$$

Occasionally the initial part of the creep curve was regarded as an *exponential* curve. This does not apply very well to the germanium experiments however, since the theoretical value of the quantity $\dot{\varepsilon}_s t_p / z_p$ would then be $e = 2.7$, which deviates strongly from the observed numbers.

It is seen that the parameter t_p is entirely determined by v, the dislocation *velocity* (and by the size of the specimen), whereas $\dot{\varepsilon}_s$ depends on $t_0 = \theta + \delta/v$. Hence it is possible, by analyzing the creep curves taken at various values of stress and temperatures, to derive the stress and temperature dependence of both v and θ, the second quantity denoting the delay period, if present, of the dislocation sources. It is then found that at sufficiently high stress levels θ is insignificant as compared to δ/v, and $\dot{\varepsilon}_s$ and t_p are both determined by the dislocation velocity only. This corresponds to formula (2) of section 2.1. Below a certain transition stress, however, in eq. (3) θ begins to dominate over δ/v, and the stress dependence of $\dot{\varepsilon}_s$ (but *not* of t_p) changes. As mentioned earlier, all this is in accord with the observations.

Numerical results, the influence of the original dislocation density, the presence of impurities and the size of the sample, are discussed in full in the papers cited, to which we refer. There, also, the stress dependence of both θ and v is shown to be exponential, which is presumably caused by the presence of a stress-dependent energy of activation, for the motion as well as for the generation of dislocations. We come back to this in the next section.

About the nature of the sources not much is known at present. From the experiments on germanium it is clear that they are *not* connected with the dislocations present before the creep began, because the density of these dislocations has practically *no* influence on the creep curve. Presumably, we have to do here with *surface sources*, which must be present to a density of the order of 100 cm^{-2} in order to explain the observed creep rate.

Only when the "original" dislocation density increases above 10^6 cm^{-2}, some influence of it is noted, in that the creep rate becomes proportional to this density, and the cubic time law changes in the

sense of a more exponential decrease, indicating that some degree of dislocation *multiplication* occurs. The latter process has been observed visually by GILMAN and JOHNSTON [1959] in lithium fluoride.

5.2. GENERATION AND MOTION OF DISLOCATIONS IN THE DIAMOND LATTICE

According to the considerations of Chapter XXIX, a dislocation of arbitrary type in the diamond lattice is usually characterized by a sequence of free bonds along its axis. These free bonds could sometimes recombine into double bonds or otherwise interact (Fig. XXIX–6), or, alternatively, the core structure might locally undergo "diffusion"-rearrangement, as was depicted in Fig. XXIX–7. Thirdly, the dislocation might become extended into partial dislocations and a stacking fault; the orientation of the latter could vary from place to place. All these processes result in a *local* decrease of the mobility of the dislocation line. It is clear that two kinds of "motion inhibitors" can be distinguished, corresponding respectively to *electronic* and to *diffusional* rearrangement of the dislocation structure. Free bonds do not move as such through the crystal, they break and reform continuously. This is an electronic process and, although energy is needed to perform the breaking, the process will occur fairly rapidly at moderate temperatures. This applies a fortiori to electronically rearranged dislocations. It must therefore be expected that a dislocation with free bonds, and still more one with double bonds, needs an activation energy to perform even the simplest glide motion, but nevertheless can move rapidly. The energy must be supplied at definite places along the dislocation line, *continuously* during the motion, because each time the dislocation has moved one atomic distance, the electronic configuration has again returned to the low-energy level.*

Diffusionally rearranged dislocations are not only locally immobilized but a much longer time is usually required for such a dislocation to overcome its mobility blocks by thermal activation, thus, to obtain F.R. source action (longer in the ratio $Q/k\theta_D \approx 100$ than for the electronically rearranged dislocation, where Q is the activation energy and θ_D the Debye temperature).

Diffusion locks occur presumably only in stationary dislocations,

* Note added in proof: R. K. HAM (Phil. Mag., to be published) has an alternative explanation for the stress dependence of the activation energy, viz. by invoking the *nucleation of pairs of kinks* (cf. section XVII-2.2.1) as the thermally activated process.

because of the long time required for their formation. Therefore diffusion rearrangement retards primarily the *generation* of new dislocations, whereas the *mobility* is affected particularly by electronic processes. A quantitative statement as to this difference cannot yet be made at present.

Both kinds of physical locks (as contrasted to the chemical impurity locks) occur at discretely spaced sites along a dislocation line. The linear concentration of the electronic locks is fairly independent of the temperature insofar as the geometrical configuration alone determines the number of free bonds. On the contrary, the density of diffusion locks will presumably depend markedly on the temperature; the linear equilibrium concentration along the axis will in good approximation be proportional to $\exp(-E_r/kT)$, where E_r is the energy gained in the diffusional rearrangement of an element of atomic size of the dislocation core.

Consider now a dislocation with locks spaced at a mean distance l_1 along its axis. Let the energy that must be supplied to overcome the dragging action of one inhibitor, thus for instance the energy to move an electron from one bond to the next, be W_1. In the presence of an applied shear stress σ, the dislocation tends to move and drags the locks along with it. The potential energy gained in the stress field in moving one atomic distance is $\sigma b^2 l$ per segment containing one dragging point. The energy to be supplied per dragging point is thus

$$W_1 - \sigma b^2 l_1 . \tag{11}$$

That is, the thermal activation for *motion in* the direction of the stress of such a *continuously* impeded dislocation is

$$Q_1 = W_1 - \sigma b^2 l_1 . \tag{12}$$

There also exists a finite probability that the dislocation, by action of the thermal vibrations, moves partially *against* the direction of the applied stress. The activation energy for that process is given by

$$Q_1' = W_1 + \sigma b^2 l_1 . \tag{13}$$

Expressions of the same type hold for the energy to activate into the mobile state a dislocation segment with (diffusional)locks, so that it can move and generate new dislocations. Only, the numerical values of the quantities W_2 and l_2 pertaining to this case may differ from

those in (12). We shall therefore write for the thermal activation energies for the *generation* of dislocations:

$$Q_2 = W_2 \pm \sigma b^2 l_2 \,. \tag{14}$$

Since the diffusion process needed to perform the transformation from the immobile into the mobile state is essentially an atomic process, its rate is determined by the atomic vibration frequence and the incubation time for source action can be written as:

$$\begin{aligned} \theta &= \nu_0^{-1} \left[\exp\left\{ - \frac{W_2 - \sigma b^2 l_2}{kT} \right\} - \exp\left\{ - \frac{W_2 + \sigma b^2 l_2}{kT} \right\} \right]^{-1} \\ &= \tfrac{1}{2} \nu_0^{-1} \exp\left(\frac{W_2}{kT} \right) \left[\sinh\left(\frac{\sigma b^2 l_2}{kT} \right) \right]^{-1} . \end{aligned} \tag{15}$$

In practice the argument of the sinh is equal to or larger than unity, and we can write

$$\theta = \tfrac{1}{2} \nu_0^{-1} \exp\left(\frac{W_2}{kT} \right) \exp\left(- \frac{\sigma b^2 l_2}{kT} \right), \tag{15a}$$

which forms a possible theoretical basis for the observed exponential stress dependence of the *incubation time* of the sources (formulae (1) and (2)).

The exponential stress dependence of the dislocation *velocity* can be understood in a similar way. According to (12) and (13), each dislocation segment containing one (electronic) locking point moves forward over a distance b in the time

$$t_1 = \left(\frac{k\theta_D}{h} \right)^{-1} \exp\left(\frac{W_1}{kT} \right) \left[\sinh\left(\frac{\sigma b^2 l_1}{kT} \right) \right]^{-1},$$

from which it follows that the dislocation velocity should be given by

$$v = \frac{hb}{k\theta_D} \exp\left(- \frac{W_1}{kT} \right) \exp\left(\frac{\sigma b^2 l_1}{kT} \right) \tag{16}$$

(again the sinh has been replaced by an exponential function). This can be compared to the observed exponential stress dependences of the dislocation velocity. To obtain numerical agreement in germanium, one must put

$$W_1 = 1.7 \text{ eV} , \quad l_1 = 3 \times 10^{-6} \text{ cm} ,$$

whereas W_2 lies somewhere in the neighbourhood of 2 eV and l_2 in that of 10^{-5} cm.

Dislocation locking resulting in inhibited source action can also be caused by chemical impurities such as oxygen, as has indeed been observed in germanium, inter alia by PENNING and DE WIND [1959], and by PEARSON *et al.* [1957] in silicon. The concentration of oxygen can only be determined with difficulty. It seems probable, however, that below a volume concentration of 10^{16} cm^{-3} the dislocations in germanium and silicon are no longer notably impeded. Assuming a total number of dislocation sites of the order of 10^{11} cm^{-3}, this implies that the oxygen atoms do not show a strong preference for dislocations. This stands in contrast to the large energy (of the order of $W_2 \approx 2$ eV) that is apparently needed to release a dislocation from them, once they are attached to it. To solve this paradox one must assume that the O-atoms are *chemically* bound to dislocations, say at the free bonds along their axes. Since an oxygen atom presumably also occurs in a

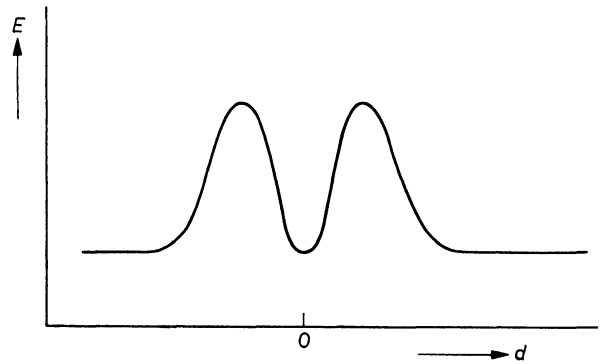


Fig. 9. Proposed (schematic) dependence of the potential energy of an oxygen atom on the distance from a dislocation. To reach the latter or to leave it, the atom has to surmount a barrier; its energy in and outside the dislocation is, however, the same. In equilibrium the O-atoms therefore show no preference for dislocation sites.

bound form in the undislocated lattice, the free energy of the crystal is not noteworthy decreased by the absorption of oxygen to dislocations. To pull a dislocation loose from its "accidentally" absorbed oxygen atoms, however, the chemical bonds must be broken, and a large activation energy is required. This argument is illustrated in Fig. 9, in which the energy of an O-atom is depicted as a function of its position with respect to a dislocation.

5.3. WORK-HARDENING IN GERMANIUM AND SILICON

The phenomenon of work-hardening, apparent in the last part of the creep curve and in the flow after yielding, is theoretically less well understood so far in germanium and silicon than in metals. Although a resemblance between the stress-strain curves of the two groups of material is superficially present (Fig. 10), the underlying mechanisms are presumably rather different. Two obvious distinctions have already been mentioned before: The strong temperature dependence of the flow stress and work-hardening in germanium against

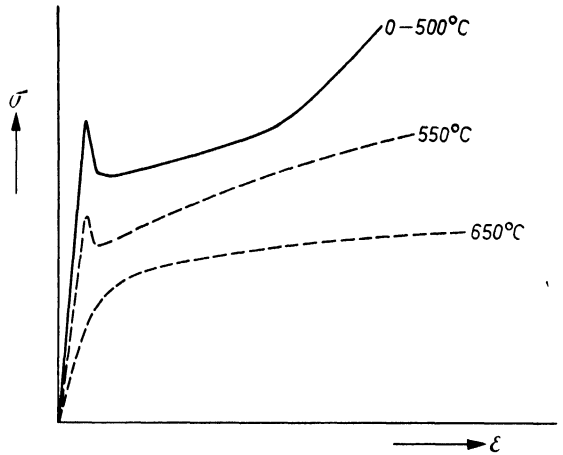


Fig. 10. The stress-strain curve of an impure metal (full-drawn) showing a yieldpoint (such as iron) compared to that of germanium (schematic), at different temperatures (dotted lines).

a practical temperature insensitivity in metals, and the large difference between the dislocation densities in the two kinds of substances, which show nevertheless comparable work-hardening. In the semiconductors, an upper density limit of about 10^8 cm^{-2} seems to exist, lower by a factor of 10^4 than in metals. It can therefore be regarded as a plausible assumption that in semiconductors work-hardening is only to a small degree caused by the interaction between dislocations, namely *only* in so far as the *long-range* stresses are concerned. These apparently limit the dislocation density to the rather small value mentioned. The work-hardening itself must rather be intimately connected with the restrictions of the mobility of dislocations, inherent to their own nature, and *not* by the geometrical interaction with other dislocations, such as in metals (see Chapter VIII).

HAASEN [1957] has proposed a theory of the flow stress after yielding, which in some aspects resembles the ideas presented in the foregoing section. Dislocations are supposed to be able to move only when a certain diffusion process takes place. This is due to the fact that a common 60°-dislocation, according to this author, should always contain a microscopical crack at its axis (compare Chapter XXIX) which has to diffuse along with it. An exponential dependence of the flow stress on the temperature, as has been observed by PATEL and ALEXANDER [1956], can indeed be explained by this theory, in a way analogous to the derivation of the temperature dependence of micro-creep, described in section IX–3. The result is:

$$\sigma_f = \frac{vkT}{b^2 D_0} e^{U/kT}, \tag{17}$$

where v is the velocity of a dislocation, assumed here (at high stress levels and temperature) to be constant and D_0 and U refer to the effective diffusivity of the crack. Taking $D_0 = 1$ cm^2/sec and $U = 0.6$ eV, fair numerical agreement with the observations was obtained.

Comparison of this theory with the theory discussed in the previous section, shows that here we have to do with a kind of diffusional locking by microcracks, which *remains* operative during the motion of the dislocation. These cracks can, of course, also be considered as a kind of core-rearrangement, although this is now continuously present along the whole dislocation line. It is the latter fact that explains the stronger (exponential) dependence of the *flow* stress on the temperature as compared to that of the *yield* stress. The two dependences were observed to be different indeed. It is possible to explain the temperature dependence of the flow stress wholly in terms of the concepts used before in this section, by noting that the spacing of electronic locks, l_1, will itself depend on the temperature, according to an exponential law in which the activation energy U may well be different from that to move the locks. By inversion of formula (7) and using (16) a formula resembling (17) is then obtained and no additional assumption is required.

6. Mechanical Twinning in the Diamond Lattice

Twinning of germanium crystals and some other diamond-structure materials has been observed under prismatic punching at a relatively low temperature (300–600 °C). It must be realized that this method

of loading is rather special, owing to the simultaneous presence of a large hydrostatic pressure and of shear stresses (CHURCHMAN, GEACH and WINSTON [1956]). Although the twinning plane was usually an octahedral plane, occasionally higher-index habit planes like (123) were noted. The twinned layers are very thin (lamellae); whether in compounds real twinning can take place still awaits confirmation. Growth twins are often encountered in germanium and silicon, the twinning plane is then invariable (111).

The stacking order of (111)-layers across the boundary of a *coherent* twin with (111)-habit plane can, in the notation of Chapter XXIX, be written as

$$\ldots\, aa'\, bb'\, cc'\, aa'\, bb'\, aa'\, cc'\, bb'\, aa' \ldots$$

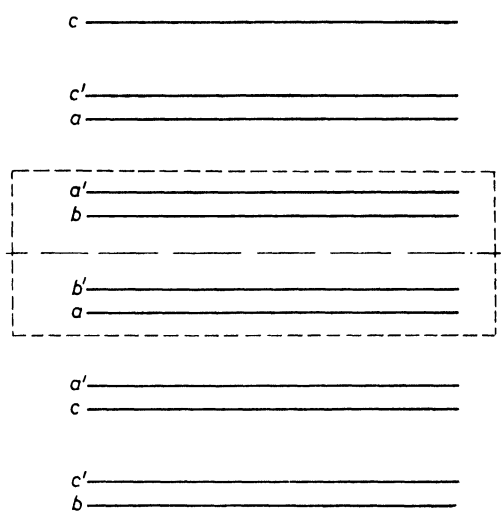


Fig. 11. Stacking of layers across a coherent (111)-twin interface in the diamond lattice, as compared to the stacking in the untwinned lattice. Each (111)-layer has two components, indicated by the same symbol with and without a dash.

This is shown in the model of Fig. 11 (after HAASEN and SEEGER [1958]). These authors point out that a *stacking fault a' bb' a* is obtained when a dislocation moves between the planes *b'* and *c* which shifts the *c* plane originally present into an *a* position. Such a dislocation must have a [112] Burgers vector and a $(11\bar{1})$-glide plane, it is thus evidently one of the partial dislocations of type I described in Chapter XXIX. Actual twinning takes place when partial dislocations of this type move

successively on neighbouring $(11\bar{1})$-planes. This will certainly not happen accidentally, but could be the outcome of a pole mechanism as described in Chapter XIII, with a screw-type pole dislocation that has a component of the Burgers vector of suitable length perpendicular to the glide plane, thus along a $[11\bar{1}]$-direction. Such a dislocation is not known, however, in the diamond structure, and a pole mechanism of the kind described does not seem very feasible, therefore.

CHURCHMAN *et al.* [1956] remarked that the {111}-planes are the only possible *coherent* twin planes in the diamond structure. The best "semi-coherent" planes are, interestingly enough, the {123} twinning planes, which have also been occasionally observed. BULLOUGH [1957] has attempted to explain the observed twinning characteristics in terms of the properties of planar arrays of large numbers of dislocations, which he and Bilby (BULLOUGH and BILBY [1956]) called *surface dislocations*. Bullough made the somewhat restrictive assumption that a twin boundary may be represented as a so-called glissile surface dislocation, a continuous set of parallel, like, glissile dislocations in *one* plane (thus a kind of large-angle grain boundary). The lattice strain across a twin interface being a pure rotation, the dislocations must be edge dislocations and the boundary is a symmetrical tilt boundary. Starting from a presupposed type of surface dislocations, it is now a geometrical problem to determine those lattice planes across which a rotational strain can be built up from these dislocations which agrees exactly with a twin relationship of the two crystal halves. When the edge dislocations which constitute the surface dislocation are of the normal type ⟨110⟩ {111}, the {123}-planes are found to be the lowest-index planes in the diamond structure for which this is possible, the next possible zone is {345}.

Twinning on {111}-planes can also be explained in this way by taking as boundary dislocations ⟨112⟩ {111} partial dislocations, thus the same partials as were mentioned above.

In the body-centred cubic structure a similar treatment leads to complete agreement with the observed twinning characteristics. The reponsible edge dislocations must then be taken of the type common to that structure.

The results of this theoretical treatment all rest on the assumption that the twin boundary is a glissile surface dislocation with only one type of dislocation in it. Turning our attention now to the actual *creation* of a dislocation array, it is found that a preference for such

simple arrays might be explained in terms of their exceptionally easy formation. Consider for instance the behaviour of large numbers of pairs of edge dislocations of opposite sign formed at some stress concentration in the crystal (Fig. 12). These dislocations can spread and line up into parallel arrays, e.g. in analogy to the mechanism of kink band formation in metals, described in section XIII–5 (FRANK and STROH [1952]). When the arrays are on {123}-planes and the

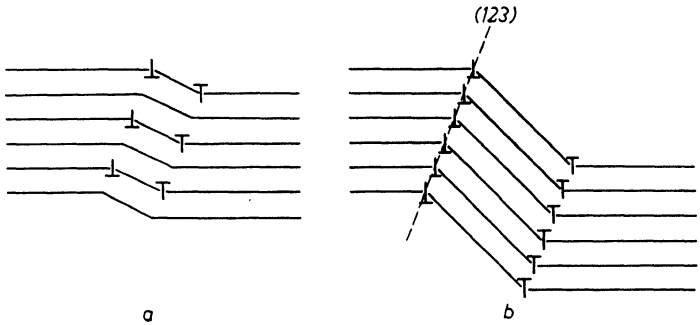


Fig. 12. When a group of edge dislocations of opposite sign are present, these can line up in two ways; *a*) to form a kink band; and *b*) to form a coherent twin boundary. The latter possibility exists only in a limited number of definite orientations, such as on a {123}-plane.

number of dislocations is large enough, formation of a twin boundary might be the final result rather than of a kink band. Presumably a dislocation distribution which conforms to the formation of such a twin boundary represents a minimum-energy arrangement, and for this reason the spreading dislocations are induced, as it where, to arrange themselves in a twin boundary.

The twin lamellae thus formed are very thin, and although now the condition for successive motion of twin dislocations on neighbouring glide planes is less stringent, some kind of (pole?) mechanism is still wanted to thicken the twin to macroscopic size, since the simultaneous glissile motion of very large numbers of dislocations seems improbable.

7. Plastic Deformation and the Electrical Properties of Germanium and Silicon

7.1. CONDUCTION OF ELECTRICITY

Plastic deformation interferes strongly with the electrical properties of semiconductors, much more so than in metals. This is due to the

fact that lattice imperfections do not only influence the mobility, but also the concentration of charge carriers.

As in metals, several different kinds of lattice defects are created simultaneously during plastic deformation. Since the latter has necessarily to be carried out at elevated temperatures, where an appreciable diffusivity of some of these defects is already observed, and as the result of the deformation itself usually depends on the temperature, recovery experiments such as those described in Chapter XV are not feasible here. The various imperfections can only be distinguished in other, much less effective ways, which renders investigations of this kind very difficult to interpret.

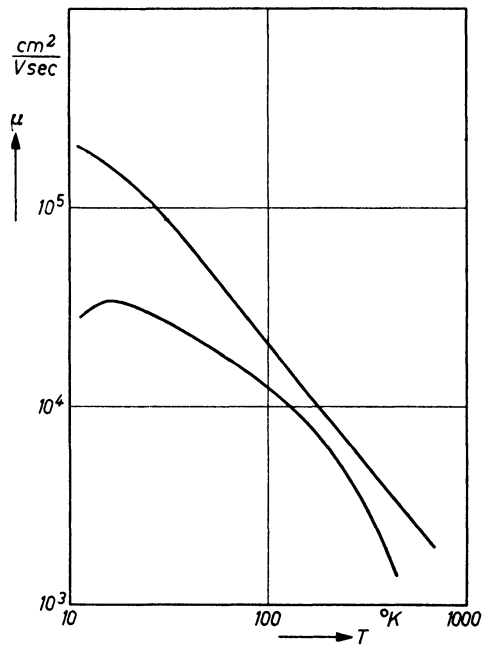


Fig. 13. Decrease of Hall mobility μ after deformation of n-germanium as a function of temperature (after TWEET [1955]). Upper curve: unstrained crystal; lower curve: same crystal strained 10%.

The most important results of the complex of phenomena occurring during plastic strain of silicon and germanium are the following:

1. The mobility of the charge carriers (deduced from the Hall constant and the resistivity) decreases. This decrease amounts at room temperature to a factor of about 2 after 10% strain in germanium and

increases with decreasing temperature (Fig. 13). The effect is, of course, due to *scattering* by the imperfections introduced.

2. The carrier concentration (as deduced from the Hall constant) varies in the sense: less electrons, more holes; that is, *acceptor centres* are created. n-Type germanium is converted into p-type after sufficient deformation (Fig. 14). The energy of activation of the acceptor

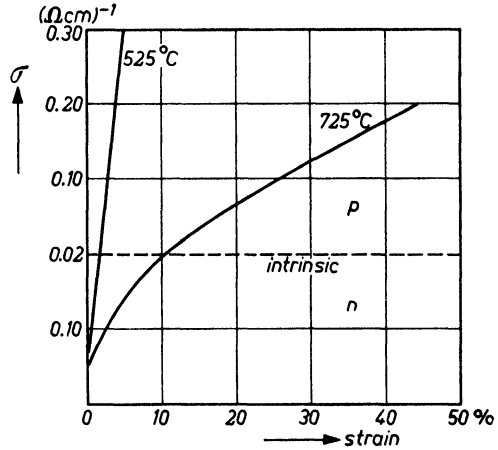


Fig. 14. Change of conductivity σ of n-germanium after plastic compression. Conversion to p-type occurs after a few per cents of strain (after GREINER and ELLIS [1955]).

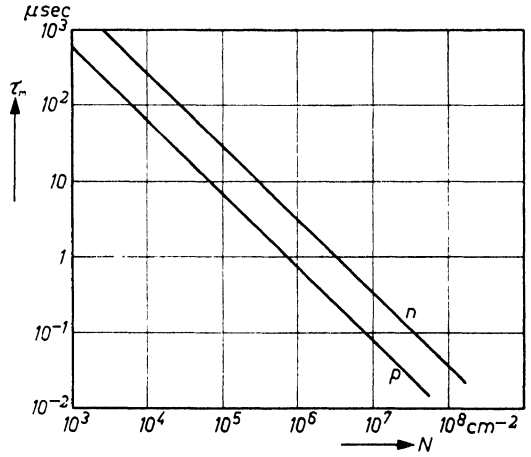


Fig. 15. Effect of dislocations on the lifetime τ_m of minority carriers in germanium. An inverse proportionality exists between dislocation density N and lifetime. (After HAASEN and SEEGER [1958]).

centre responsible for this amounts to about 0.1 eV in germanium, as follows from the temperature dependence of the Hall constant in deformed crystals.

3. The lifetime of the minority carriers decreases by several orders of magnitude, especially at low temperatures (Fig. 15). Evidently *recombination centres* are also formed. Various authors reported largely different efficiencies of these centres, however.

The following discussion is based on an excellent review paper by HAASEN and SEEGER [1958]. The subject being still in a preliminary state of investigation, a consistent description of all facets of the complex phenomena cannot yet be given. (Compare also SEEGER [1959].)

7.2. FORMATION OF ACCEPTORS BY DEFORMATION AND QUENCHING

An important series of measurements was started by PEARSON, READ and MORIN [1953] and has been extended by TWEET [1955]. These authors measured the resistivity and the Hall constant as a function of the temperature of germanium specimens plastically bent at 550 °C or higher. Tweet observed also the temperature variation of the Hall constant after annealing subsequent to the deformation and found that part of the acceptor centres (of 0.1 eV activation energy) introduced during the latter treatment annealed out in about one hour time at 700 °C. A certain residual density of acceptor centres, not characterized by a definite activation energy, remained present, however. Tweet assumed the latter to be associated with *dislocations*, which indeed can be expected to act as acceptors (see Chapter XXVIII), and annealing has certainly no effect on the dislocation density as can be concluded from the etch-pit countings. The recoverable centres could, on the other hand, very well be germanium *vacancies* (not interstitials, as these must rather be expected to act as donors). High-temperature deformation only leads to dislocation-type acceptors, of course, and does not lead to conversion, which the experiments of Pearson *et al.* confirmed.

As a check on Tweet's hypothesis, the electrical properties of *quenched* germanium crystals are of interest. LETAW [1956] has discussed these in great detail. Quenching also leads to the creation of acceptor centres, and from the dependence of their concentration on the quenching temperature, the energy of *formation* of these centres can be derived. This was found to be about 2 eV. Presumably, the

quenched-in defects are vacancies. Assuming that self-diffusion, which is characterized by an energy of activation of about 3 eV in germanium, follows a vacancy mechanism, the migration energy of the vacancies is then about 1 eV. That is, the jumping frequency of vacancies follows from the expression:

$$\nu \approx \nu_0 \, e^{-\frac{1(eV)}{kT}}. \tag{18}$$

The point defects introduced by quenching were observed by Letaw to anneal out in a few minutes at 350 °C, whereas the theoretical jumping frequency of simple vacancies at this temperature amounts to about 10^5 sec^{-1}. That is, each defect must evidently perform about 10^7 jumps before it is absorbed, for instance at a dislocation.

The observed annealing times for the recovery of the acceptors introduced during *plastic deformation* are much longer, viz. of the order of 10^3 to 10^4 sec at 700 °C. This time corresponds to about 10^{11} jumps of a vacancy. We are thus left with the remarkable fact that the "vacancies" in plastically deformed crystals, where a large number of effective traps is presumably present, must travel over much greater distances (about 10^{-2} cm) to become absorbed, than the vacancies in much less disordered quenched crystals (about 10^{-4} cm). The obvious way out of this contradiction is to assume that the defects introduced by straining differ in nature from those introduced by quenching. For this reason Tweet's hypothesis was replaced (simultaneously by HAASEN and SEEGER [1958] and by PENNING and HORNSTRA [1958]), by the assumption that after deformation predominantly *double* vacancies (divacancies or vacancy pairs) are generated, whereas after quenching only *single* vacancies are present. The vacancy pairs would diffuse much more slowly than single vacancies (in sharp contrast to the situation in close-packed metals, Chapter XV). This must indeed be expected on theoretical grounds, for among the nearest neighbours of any atom in the diamond lattice there are none which are themselves nearest neighbours. Therefore, a double vacancy must perform a very complicated "wriggling" motion when migrating as such.

It must further be supposed that the energy of formation of a vacancy pair is rather large compared to that of a single vacancy (difference of about 1 eV). Otherwise, in thermal equilibrium at high temperatures, and thus also before and after quenching, double vacancies should be present in large numbers. Now, indeed, after thermal

quenching part of the acceptors present does *not* anneal out at 350 °C. This can be due either to an appreciable fraction of the defects being present as double vacancies (LOGAN [1956]) or to the association of the vacancies during the actual annealing at about 350 °C. The latter possibility should probably be preferred, when it is assumed in addition that the gain in energy obtained when two vacancies combine into a pair (the association energy), is considerable (>1 eV). Although a vacancy pair is thus characterized by an appreciably larger energy of formation than a single vacancy, it nevertheless still represents a large energy of association. This situation can occur in germanium because a single vacancy itself has already a large energy of formation.

In order of magnitude (accuracy $\frac{1}{4}$ eV), one should expect the following values to apply, as deduced from the characteristics of the annealing processes, etc., see also section 8 and next chapter:

Single vacancy:	Energy of formation:	2 eV,
	Energy of migration:	1 eV,
	Acceptor level:	0.02 eV above valence band;
Double vacancy:	Energy of formation:	$2\frac{3}{4}$ eV,
	Energy of association:	$1\frac{1}{4}$ eV,
	Energy of migration:	$1\frac{3}{4}$ eV,
	Acceptor level:	0.1 eV above valence band;
Interstitial:	Energy of formation:	$\geqslant 2\frac{1}{2}$ eV,
	Energy of migration:	$\frac{3}{4}$ eV,
	Two donor levels:	0.225 and 0.55 eV above valence band.

The reason why double vacancies are preferentially formed during plastic deformation is connected with the typical "double-layer structure" of the diamond lattice. To form a jog in an edge dislocation with a {111}-glide plane, the extra halfplane has to move up (or down) across *two* atomlayers, viz. those two which constitute the double layer of the {111}-plane considered. Non-conservative motion of this jog will, in complete analogy with the mechanism described in section II–5, leave a row of defects in its wake. Each defect, however, consists in this case of *two* vacancies or *two* interstitials, one originating in each component of the double layer.

Also imaginable, of course, are jogs that correspond to a jump of the dislocation axis across one component layer only. The situation

that then develops can in all its aspects be compared to a "diffusion-rearranged" dislocation (see Fig XXIX–7). Such a "half-jog" shows very little mobility, not only as regards the non-conservative motion together with the parent dislocation, but also for conservative motion more or less *along* the parent dislocation. The mobility of a dislocation containing half-jogs is thus very seriously impaired, because these jogs cannot move in either way without simultaneous defect diffusion. Such dislocations, presumably, take little or no part in the gliding process. Reversely: since during glide only the dislocations with complete jogs are active, vacancy *pairs* and interstitial *pairs* are generated instead of the single defects.

The thermal generation of defects from stationary jogs, and the absorption of defects at dislocations, are presumably only slightly affected by the special dislocation geometry.

More about the acceptor action of vacancies and vacancy pairs will be said in section 8.

In Chapter XXVIII we have already described Read's theory of the statistics of occupation by electrons of the free bonds along a dislocation line. This can be studied experimentally by measuring the Hall constant and the resistivity of materials deformed at a high temperature, or of materials which have been well annealed after deformation, in order that they contain only dislocations as imperfections. Moreover, since it is only in n-germanium that dislocations are expected to affect the carrier concentration (at a low temperature), n-type material should be used. That dislocations indeed act as acceptors has been proved experimentally by PEARSON *et al.* [1953]. The acceptor level in bent n-type germanium crystals due to dislocations lies about 0.2 eV below the bottom of the conduction band. Any donor activity by dislocations has never been observed.

Grain boundaries, being arrangements of large numbers of dislocations, have been observed to possess very strong acceptor activity; polycrystalline n-germanium samples showed local hole conduction near some of its grain boundaries. At low temperatures the boundaries behaved instead as ohmic conductors, independent of the presence of impurities (REED *et al.* [1959]). The situation is not yet very clear. Coherent twin boundaries are not built up from dislocations and accordingly show the normal bulk behaviour. BARDSLEY [1960] has reviewed the influence of grain boundaries on the electrical properties of semiconductors.

7.3. SCATTERING BY LATTICE IMPERFECTIONS

Scattering of electrons and holes by point defects in the diamond lattice is presumably less important than in metals, because of the small effective mass of the charge carriers. Further, only the lattice strain is of influence here, first-order electrostatical effects being absent at moderate temperatures.

This is not true for dislocations, because of their strong acceptor activity. The contribution of the lattice strains around dislocations to the scattering of charge carriers is relatively small. It was computed in first-order approximation for an edge dislocation, taking into account the dilatational strains only (DEXTER and SEITZ [1952]). These strains contribute to the resistivity to the following extent:

$$\Delta\varrho = \frac{3\pi}{32}\left(\frac{1-2\nu}{1-\nu}\right)^2 \frac{E_1 b^2}{\hbar k T}\,\frac{N m^*}{n e^2}\,. \tag{19}$$

Here E_1 is the so-called deformation potential constant, relating bond shift and dilatation in the Bardeen-Shockley theory of the deformation potential (see also Chapter XXVIII); ν is Poisson's ratio, N and b are density and Burgers vector of the dislocations respectively, whereas m^*, n and e denote effective mass, number and charge of the carriers respectively. Only at or below hydrogen temperature does the elastic dislocation scattering in deformed germanium become comparable to the thermal resistivity. Otherwise, it is negligible compared to the impurity scattering.

However, at low temperatures in n-germanium a dislocation bears a line charge. This results, as explained in section XXVIII–2, in the formation of a fairly impenetrable cylindrical space charge region around a dislocation, which again causes appreciable scattering, in first instance proportional to the fraction of the material volume occupied by the space charge cylinders. The most interesting characteristic of this scattering is its pronounced orientation dependence.

Experimentally, the scattering of electrons and holes by imperfections forms rather a neglected field of study in germanium and silicon. Much more in the centre of attention, because of its practical interest, is the effect of imperfections on the *lifetime* of minority carriers.

7.4. DISLOCATIONS AND LIFETIME

Whether all physical point defects can act as recombination centres or not is not known at present. In all experiments (except one discussed in section 8) bearing on this point, sufficient dislocations were simultaneously present with the point defects to eclipse the possible activity of the latter.

Chemical impurities, such as copper in germanium for instance, often act as efficient recombination centres. A discussion of this falls outside the purview of this book, however. Dislocations affect the lifetime of minority carriers to a great extent, as follows from many investigations (OKADA [1955]; WERTHEIM and PEARSON [1956]; MC KELVEY [1957]). In germanium the lifetime of minority carriers has been found to be inversely proportional to the dislocation density:

$$\tau_m \sim N^{-1}. \tag{28}$$

This inverse proportionality exists in deformed as well as in undeformed crystals; the proportionality factor depends on the type of conductivity, thus on the nature of the minority carriers, and on the temperature. Expressing τ_m in seconds, N in cm^{-2}, the factor is of the order unity (about 2 in n-germanium, about 0.7 in p-germanium at room temperature).

Fig. 15 shows the average results of various observations. However, left out are observations by KURTZ *et al.* [1956] resulting in a lifetime which was about 100 times larger (n-germanium). These authors determined the dislocation density by X-ray diffraction methods however, whereas the others deduced it from etch pit countings.

Small-angle boundaries also promote recombination in germanium. Since they consist of arrays of dislocations, this is not surprising. From the misorientation angle, the dislocation density in a simple tilt boundary can be accurately determined, and from this the effective capture cross-section of one dislocation. This was found by MCKELVEY [1957] to be 25 $Å^2$ in p-germanium and 4 $Å^2$ in n-germanium. As the thermal velocities of the minority carriers are known, the proportionality factor mentioned above can now be checked theoretically. Numerical agreement within 10% is obtained. HASIGUTI and MATSUURA [1957] confirmed this, but found an exceptionally *long* lifetime when measured *along* the dislocation wall. This was interpreted in terms of the formation of a thin p-layer along the boundary in n-germanium

(see previous chapter). The difference between n-germanium and p-germanium can be understood as due to the charge present on the dislocations in the former material, as contrasted to the neutrality of the dislocations in the latter. In n-germanium the space charge region around a dislocation line, repulsive for holes but attractive for electrons, reduces the recombination efficiency. MORRISON [1956] has attempted to compute on the basis of Read's model the effective time constants describing the penetration of this barrier and the capture of holes, respectively. He finds the latter to be very large.

In p-germanium the lifetime of electrons depends strongly on the temperature; it is approximately proportional to $T^{5/2}$. This remains unexplained at the moment, since from Read's model such a strong variation with the temperature of the degree of occupation of dislocation-free bonds by electrons can certainly not be derived.

8. Irradiation and Quenching of Diamond Structure Materials

The physical and mechanical properties of germanium and silicon irradiated by neutrons, deuterons, γ-rays, α-particles and electrons have been studied by many authors, but the results cannot yet be considered to yield a complete theoretical picture. The electrical properties of germanium have been found to vary considerably by irradiation: donors as well as acceptors are introduced; at room temperature only the latter remain and accordingly n-type germanium becomes less conductive and finally even converts to p-type. It was proposed by LARK-HOROVITZ [1955], see also FAN and LARK-HOROVITZ [1957], that of the Frenkel defects which are certainly formed during irradiation, the vacancies act as acceptors, the interstitials as donors. The latter presumably diffuse easily and anneal out already below room temperature, and solely the vacancy acceptors remain. There are some unexplained difficulties, however; *two* acceptor levels have been found (CLELAND and CRAWFORD [1954]), one shallow, a few hundredths of an eV above the edge of the valence band, and another, about 0.1 eV higher. The question how many acceptor levels should be assigned to a germanium vacancy has not yet been satisfactorily answered. We shall see in the next chapter that strong evidence exists as to the presence of only one acceptor level per vacancy. However, an argument in favour of two acceptor levels per vacancy has been derived from a comparison between the observed number of acceptor

levels created and the theoretical density of Frenkel defects formed after irradiation (SEITZ and KOEHLER [1956]). Whether the displacement theory itself is reliable to within a factor 2 is a matter of doubt, though (see Chapters XIV and XV). Either the observations of Cleland *et al.* must be interpreted in terms of two-level vacancies, or two types of defects (vacancy pairs?) are formed by irradiation. In this respect it is interesting to note that on pile-irradiated silicon and germanium, after etching with the usual etchant to visualize dislocations, a peculiar superficial network structure appears, with a mesh width of 0.1 to 0.5μ (CHANG [1957]). This structure anneals out after warming up, and an activation energy for this process of about 1.6 eV could be derived (to be compared with the results of Fletcher *et al.*, discussed below). The cause of this phenomenon is not known; it might very well be associated with the presence of the other type of defect.

The annealing kinetics of the changed conductivity of electron irradiated n-germanium is also abnormal (FLETCHER and BROWN [1953]). Recovery between room temperature and about 300 °C was found to be very complex; an activation energy of 1.7 tot 1.8 eV could be deduced in a rather indirect way. At least two annealing processes are operative in this temperature region; one takes place below 150 °C and another begins at 250 °C. It is tempting to associate these two steps with the migration of interstitials and of vacancies respectively. From Letaw's recovery experiments to be mentioned below followed an activation energy for vacancy migration which is much lower than 1.7 eV. This discrepancy may be resolved, possibly, by rejecting the latter value in view of the complex recovery kinetics.

More direct evidence on the migration of point defects can be derived from the study of the annealing of *thermal acceptors*, introduced by *quenching*. We have already mentioned the observations by LOGAN [1956] and their discussion by LETAW [1956]. In germanium quenched rapidly from a high temperature, between 10^{14} to 10^{15} acceptors are introduced into the material, which have an energy of formation of about 2.0 eV. Most of these thermal acceptors anneal out with an activation energy of 0.96 eV, although part of them could only be removed by heating above 600 °C. The activation energy for the annealing of the latter centres is 1.7 eV. It was deduced in section 7.2 that the easily migrating defects are vacancies, whereas those more difficult to move are *divacancies*.

Letaw also speculated on the probability for association of vacancies.

He found that two vacancies will undergo *repulsive* interaction before combining into a pair. The height of the repulsive barrier should be 0.7 eV, the total energy of formation of a divacancy, however, would be *less* than that of one single vacancy, namely only 1.1 eV. In the light of the exposition of section 7.2 this conclusion does not seem tenable.

About the depth of the acceptor level(s) introduced by quenching not much more is known than that one level, and no more, is present in the lower half of the energy gap.

Some new evidence on the acceptor activity of vacancies in germanium can be obtained in an indirect way by observing the diffusion of various impurity elements, such as copper and nickel, through germanium. We shall discuss this in the next chapter.

BEMSKI [1957] has studied the effect of quenching on the lifetime τ_m of minority carriers in silicon. Assuming $1/\tau_m$ to be proportional to the density of possibly quenched-in recombination centres, he found that such centres, as obtained after quenching from above 400 °C, should possess an energy of formation of about 0.6 eV. The lifetime could be recovered by annealing the quenched specimens between 200 °C and 400 °C; the activation energy for this annealing is 0.8 eV approximately, the number of jumps carried out by the centres (assumed to be of atomic dimension) comes out as about 10^8. Since the silicon used will presumably have contained at least 10^4 dislocations per cm^2, the trapping sites were probably dislocations. Only a fraction of the recombination centres anneals out in this way, however. It seems reasonable, in view of the evidence obtained for germanium, to identify these centres with vacancies. Electron-bombarded silicon was also investigated (BEMSKI and AUGUSTYNIAK [1957]); there the recovery of lifetime was characterized by an activation energy of 1.3 eV, in n-type as well as in p-type crystals. The annealing kinetics were again complex.

Hence, typical but analogous differences, the origin of which is unresolved so far, seem to exist between the recovery behaviour of quenched and of irradiated silicon and germanium, respectively.

Many other observations of the properties of neutron- and e- and γ-irradiated Ge and Si have been published. It is quite clear that in these materials a strong interaction, presumably of an electric nature, exists between *physical* defects and *chemical* impurities, which complicates the recovery behaviour considerably.

BARDSLEY, W. (1960); Progress in Semiconductors **4**, 154.
BEMSKI, G. (1957); Proc. Inst. Radio Eng. **46**, 990.
BEMSKI, G. and W. M. AUGUSTYNIAK (1957); Phys. Rev. **108**, 645.
BILLIG, E. (1956); Proc. Roy. Soc. **A235**, 37.
BUEREN, H. G. VAN (1958): Physica **24**, 831; (1959): Physica **25**, 775.
BUEREN, H. G. VAN, J. HORNSTRA and P. PENNING (1958); Suppl. al Nuovo Cimento VII, No. 2.
BULLOUGH, R. (1957); Proc. Roy. Soc. **A241**, 568.
BULLOUGH, R. and B. A. BILBY (1956); Proc. Phys. Soc. **B69**, 1276.
CARREKER, R. P. (1956); J. Metals (Transact. Suppl.) **8**, 111.
CHANG, J. (1957); J. Appl. Physics **28**, 385.
CHURCHMAN, A. T., G. A. GEACH and J. WINSTON (1956); Proc. Roy. Soc. **A238**, 194.
CLELAND, J. W. and J. CRAWFORD (1954); Phys. Rev. **95**, 1177.
DASH, W. C. (1956); J. Appl. Physics **27**, 1193.
DEXTER, D. L. and F. SEITZ (1952); Phys. Rev. **86**, 964.
FAN, H. Y. and L. LARK-HOROVITZ (1955); Defects in Crystalline Solids (The Phys. Society, London).
FAN, H. Y. and L. LARK-HOROVITZ (1957); Semiconductor Conference at Garmisch-Partenkirchen.
FISHER, J. C. (1955); Transact. Am. Soc. Metals **47**, 451.
FLETCHER, R. C. and W. L. BROWN (1953); Phys. Rev. **92**, 585 and 591.
FRANK, F. C. and A. N. STROH (1952); Proc. phys. Soc. **B65**, 295.
GALLAGHER, C. J. (1952); Phys. Rev. **88**, 721.
GILMAN, J. and W. G. JOHNSTON (1958); J. appl. Phys. **29**, 110.
GREINER, E. S. (1955); J. Metals (Transact. Suppl.) **7**, 283.
GREINER, E. S. and W. C. ELLIS (1955); Bell Lab. Rep. **34**, 403.
HAASEN, P. (1957); Acta Metall. **5**, 598.
HAASEN, P. and A. SEEGER (1958); Halbleiterprobleme IV (Vieweg, Braunschweig).
HASIGUTI, R. R. and E. MATSUURA (1957); J. Phys. Soc. Japan **12**, 1347.
KURTZ, A. D., S. A. KULIN and B. L. AVERBACH (1956); Phys. Rev. **101**, 1285.
LETAW, H. (1956); J. Phys. Chem. Solids **1**, 100.
LOGAN, R. A. (1956); Phys. Rev. **101**, 1455.
MCKELVEY, J. P. (1957); Phys. Rev. **106**, 910.
MORRISON, S. R. (1956); Phys. Rev. **104**, 619.
OKADA, T. (1955); J. Phys. Soc. Japan **10**, 1110.
PATEL, J. R. and W. ALEXANDER (1956); Acta Metall. **4**, 385.
PEARSON, G. L., W. T. READ and W. L. FELDMANN (1957); Acta Metall. **5**, 181.
PEARSON, G. L., W. T. READ and F. J. MORIN (1953); Phys. Rev. **93**, 666.
PENNING, P. (1958a): Phys. Rev. **110**, 586; (1958b): Philips Res. Reports **13**, 79.
PENNING, P. and G. DE WIND (1959); Physica **25**, 760.
READ, W. T. (1957); Acta Metall. **5**, 83.
REED, B., O. A. WEINRICH and H. F. MATARÉ (1959); Phys. Rev. **113**, 454.
ROTH, L. and W. E. TAYLOR (1952); Proc. I.R.E. **40**, 1338.
SEEGER, A. (1959), in: Solid State Physics in Electronics and Telecommunications, Vol. I (Acad. Press, London, New York).
SEITZ, F. and J. S. KOEHLER (1956); Solid State Physics **2**, 304.
TREUTING, R. G. (1952); J. Metals **4**, 1044.
TREUTING, R. G. (1955); J. Metals (Transact. Suppl.) **7**, 1027.
TWEET, A. G. (1955): Phys. Rev. **99**, 1245; (1957): Phys. Rev. **106**, 221.
VOGEL, F. L. (1956); J. Metals (Transact. Suppl.) **8**, 946.
WERTHEIM, K. and G. L. PEARSON (1956); Bull. Am. Phys. Soc., November, p. 331.

CHAPTER XXXI

DEFECT DIFFUSION AND INTERNAL FRICTION IN GERMANIUM AND SILICON

1. Diffusion in Germanium and Silicon

Self-diffusion experiments in germanium have lead to an activation energy of 2.96 eV. This is well comparable with the observed diffusivities of most substitutional impurities, e.g. iron and many other metals, arsenic etc., in germanium, the activation energies of which all lie between 2 and 3 eV. Assuming the substitutional diffusion to proceed by a vacancy mechanism, and accepting the value, quoted in the preceding chapter, of 1 eV for the migration energy of the latter, the formation energy of a vacancy is derived as about 2 eV. This is in excellent agreement with the value of 2.0 eV deduced by LOGAN [1956] from the results of his quenching experiments (Chapter XXX).

Diffusion measurements of foreign elements in germanium are relatively easy, as long as the impurity gives rise to acceptor or donor levels. Then, simply by measuring locally the conductivity or Hall effect, the thermoelectric power or a related quantity (e.g. by probing the specimen with finely pointed contacts), concentration variations of the impurity in question can be observed.

This no longer holds for electrically *inactive* impurities, such as hydrogen, oxygen, nitrogen, carbon, or germanium in silicon and vice versa. If no suitable radioactive tracer isotopes exist, the direct determination of either solubility and diffusivity is virtually impossible, because of the very small concentrations concerned. Reliable data are consequently very scarce at the moment. It seems that the activation energies for diffusion of *these* atoms are slightly lower than those of the previously mentioned electrically active atoms; they lie between 1 and 2.5 eV in germanium. Hydrogen seems even to diffuse with an energy of only about 0.5 eV.

A completely different situation exists where the diffusion of *copper*,

nickel or *lithium* in germanium or silicon is concerned (FULLER *et al.* [1954]). These metals have been found to migrate extremely rapidly through the lattice, at least when this contains a sufficient number of dislocations. Accordingly, it was proposed by VAN DER MAESEN and BRENKMAN [1955] that these small impurities should migrate *interstitially*. Since, however, copper and nickel act as *acceptors* in germanium, it was further assumed that these ions can also occupy substitutional positions, in which they are presumably less mobile (Chapter XXVII). The following observations strengthen this conclusion. Solute copper gives rise to three acceptor levels, respectively located at 0.04 and 0.32 eV above the valence band and 0.26 eV below the conduction band. Nickel gives two acceptor levels, 0.30 eV from the conduction band and 0.23 eV from the valence band. This is most easily explained by assuming that copper atoms can be triply ionized in the germanium lattice and nickel atoms twice. (Since the solubilities are very low ($< 10^{-4}\%$), complexes are certainly not responsible for one or more of these levels.) Moreover, copper and nickel act as very efficient recombination centres. It is improbable that the strong acceptor action of Cu and Ni could occur if the atoms were situated interstitially; apparently, in equilibrium a large part of the copper and nickel atoms is present on substitutional sites. They occasionally convert to interstitial sites, however, which gives rise to the large mobility. In other words, there exists a *chemical equilibrium* between interstitial and substitutional copper, or nickel, in the germanium and silicon lattices. It is noteworthy that in dislocation-free materials the diffusivity is quite small (FULLER and DITZENBERGER [1957]). To the essential role played by dislocations in the establishment of the defect concentrations we shall devote the next section.

2. Precipitation of Copper and Nickel in Germanium

An interesting sidelight on the diffusion of Cu and Ni in Ge is obtained by studying the *precipitation* of these metals from supersaturated solid solutions.

The solid solubility curves of copper and nickel in germanium are retrograde below about 850 °C, the solubilities being maximum near to this temperature. Samples rapidly quenched from slightly below this temperature therefore contain a supersaturated solution and in such samples the kinetics of precipitation can be studied by annealing.

This has been done independently by TWEET [1957] and by PENNING [1958a]; the precipitation of lithium in germanium was studied by MORIN and REISS [1957]. The precipitation of copper and nickel can easily be followed by measuring the carrier density, for instance by observing the Hall constant as a function of temperature. It is found that the concentration of unprecipitated copper, that is, of copper acceptor centres, measured as a function of the annealing time, varies in the way depicted in Fig. 1: the density decreases exponentially with

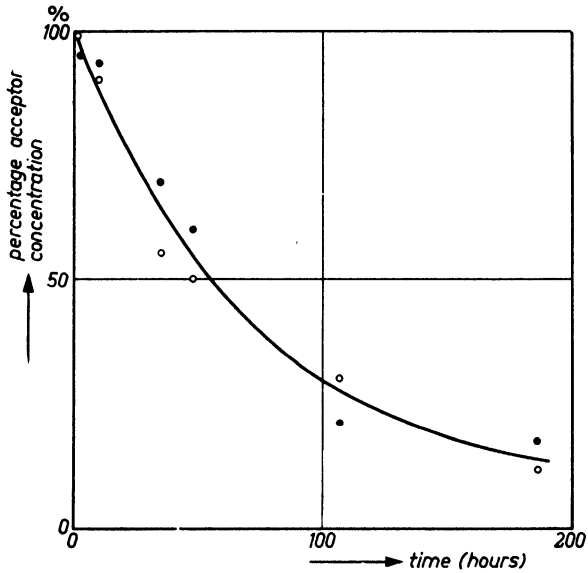


Fig. 1. Decay of acceptor density on annealing at 500 °C of two germanium samples (dots and circles) with originally 10^{15} Cu-acceptors per cm^3. The dislocation density was in both cases about 10^4 cm^{-2}. The full-drawn curve corresponds to an exponential decrease characterized by a decay-time of 95 hours (after TWEET [1957]).

time. The time constant in the exponent amounts to about 100 hours at 500 °C, it decreases rapidly with increasing temperature. A similar exponential decay of acceptor density is found during the precipitation of nickel. However, as Penning has observed, the disappearance of the 0.23 eV acceptor levels of nickel is accompanied by the creation of new acceptor levels, less than 0.06 eV below the conduction band, in such numbers that during the first part of the annealing period the *total* acceptor density remains approximately constant (Fig. 2). Later, the density of these unknown acceptors decreases together with that of the

nickel acceptors, with exactly the same time constant. To explain this, Penning assumed the new acceptors to be *vacancies*, namely those which are left behind in the lattice when a substitutional Ni-atom

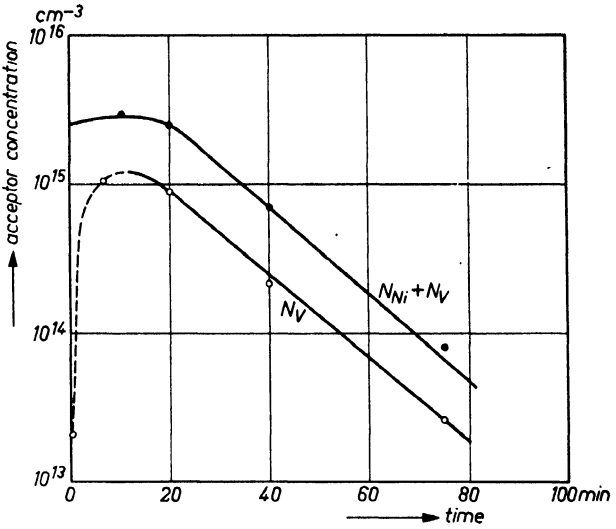


Fig. 2. Variation of the density of nickel acceptor levels, N_{Ni}, and of vacancy acceptor levels, N_V, during anneal of a nickel saturated germanium sample at 500 °C (after PENNING [1958a]).

jumps into an interstitial site (in order that it can diffuse rapidly to its precipitation nucleus). The interstitial nickel is assumed not to give rise to any donor or acceptor level, the vacancies to one acceptor level each.

An at first sight more plausible assumption, when regarded from a physical point of view, would be that a vacancy has two acceptor levels, and an interstitial Ni-atom one donor level. We shall advance arguments later to show that this assumption leads to improbable and not observed consequences, however.

To describe the phenomena quantitatively, the equilibrium between interstitial nickel ions, vacancies and substitutional nickel ions must be considered in detail. One has the equilibrium reaction

$$N_s \rightleftarrows N_i + V\,, \tag{1}$$

where N_s denotes substitutional Ni, etc. During the first part of the anneal, a quasi-equilibrium is reached:

$$\frac{[N_s]}{[N_i]\,[V]} = \text{constant} . \tag{2}$$

Here, the various concentrations themselves need *not* be equilibrium concentrations; usually they are much larger. During the further anneal the following processes now take place simultaneously:

a) Transition of Ni from substitutional to interstitial sites, together with the formation of vacancies. The rate of this reaction is governed by the equation:

$$\frac{d\,[N_s]}{dt} = -\,k_1\,[N_s] + k_2\,[N_i]\,[V] \tag{3}$$

from which (2) follows directly (in equilibrium).

b) Annihilation of vacancies, either by absorption at dislocations or through capture by an interstitial nickel atom. For the vacancy concentration we have:

$$\frac{d\,[V]}{dt} = k_1\,[N_s] - k_2\,[N_i]\,[V] - G\,[V] . \tag{4}$$

Here it is assumed that the rate of annihilation of vacancies at dislocations is always proportional to [V]. The proportionality constant G then contains the density of dislocations in linear form. This is allowed when $[V] \gg [V_0]$, the equilibrium concentration of vacancies at the annealing temperature. A third equation is obtained from the assumption that the interstitial Ni-atoms diffuse so rapidly (compared to the whole annealing process) to places where they can precipitate, that the interstitial Ni-concentration is always equal to the *equilibrium* concentration (*interstitial solubility*) at the annealing temperature. (It was already silently assumed that substitutional nickel atoms do not diffuse at all.) Thus one has:

$$[N_i] = [N_{i0}] . \tag{5}$$

A solution of (3) to (5) reads, when $N = [N_s]$ for $t = 0$; $(\alpha_2 > \alpha_1)$:

$$\begin{aligned} [V] &= \frac{N k_1}{\alpha_2 - \alpha_1}\,\{e^{-\alpha_1 t} - e^{-\alpha_2 t}\} ; \\ [N_s] + [V] &= \frac{N}{\alpha_2 - \alpha_1}\,\{\alpha_2\, e^{-\alpha_1 t} - \alpha_1\, e^{-\alpha_2 t}\} ; \end{aligned} \tag{6}$$

where $\alpha_{1,2}$ are the roots of the equation

$$\alpha^2 - \{k_1 + k_2 [N_{i0}] + G\} \alpha + k_1 G = 0 . \tag{7}$$

α_1 describes the rate of precipitation, α_2 the much more rapid rate of redistribution between substitutional and interstitial atoms. G is numerically intermediate between the two and describes the rate of annihilation of vacancies at dislocations. Now, indeed, for short annealing times, $[N_s] + [V]$ is approximately independent of the time, and the *total* number of acceptor levels thus remains constant (Fig. 2). Only for $t > 1/\alpha_2$ this number begins to decrease. The time variation of the vacancy acceptor density as shown by the lower curve of Fig. 2 is well described by the quantity [V] of formula (5).

For $\alpha_2 \gg \alpha_1$, as is indeed the case, one has

$$\begin{aligned} \alpha_2 &= k_1 (1 + K) ; \\ \alpha_1 &= \frac{G}{1 + K} ; \end{aligned} \tag{8}$$

where

$$K = \frac{[N_s]}{[V]} = \frac{k_2}{k_1} [N_i] = \frac{k_2}{k_1} [N_{i0}] = \frac{[N_{s0}]}{[V_0]} . \tag{9}$$

K is thus a constant at a given temperature, as it describes the ratio between certain equilibrium concentrations. The precipitation rate α_1 is completely determined by this equilibrium constant K and the (structure sensitive) vacancy-annihilation constant G.

These considerations apply equally well to copper and to nickel. Only in the former case the interpretation of the observations is somewhat more difficult, since copper ions have themselves already a second shallow acceptor level associated with them which lies about as deep as the vacancy level. When copper and nickel are present simultaneously, the rates of precipitation of both are slowed down. This is because, instead of K defined above, now the modified constant

$$K = K_{Cu} + K_{Ni} \tag{10}$$

must be used, and α_1 becomes accordingly smaller. That the interaction between Ni and Cu would lead to this and analogous phenomena was first predicted by Vink, and has been elegantly confirmed by

PENNING [1958a]. The value of K_{Cu} appears to be considerably larger (about 10) than K_{Ni} (about 3), so that the influence of Ni on copper-saturated germanium is much less than vice versa.

During the precipitation of supersaturated nickel, the vacancy concentration is much larger than the equilibrium concentration. Therefore, copper may be taken into solid solution in much greater quantities under these circumstances than in pure germanium. We refer to the original work for a detailed discussion of these interesting phenomena

A point that interests us here particularly is the structure sensitivity of the precipitation. TWEET [1957] has observed that at elevated

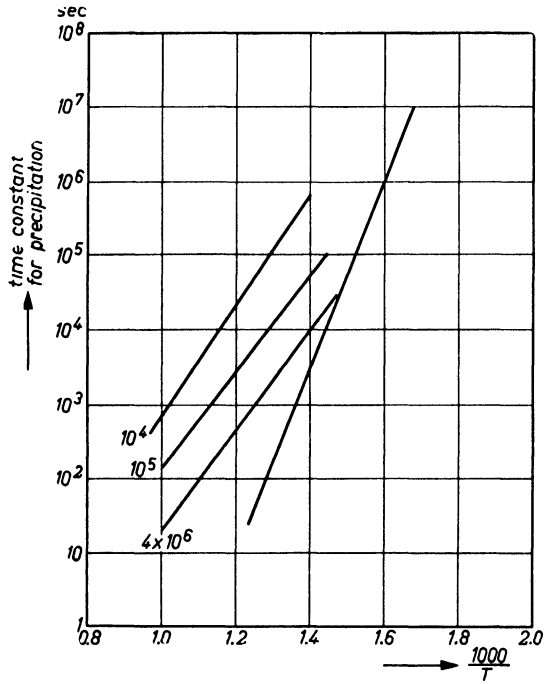


Fig. 3. Time constant for precipitation of copper in germanium as a function of the inverse absolute temperature. The various curves refer to samples with different dislocation densities (crude averages only are indicated in the diagram). At high temperatures the precipitation rate varies more slowly and depends upon the dislocation density (after TWEET [1957]).

annealing temperatures the temperature variation of the precipitation rate of copper in germanium is less than at low temperatures (Fig. 3).

The transition is discontinuous, the activation energy below it is 2.7 eV, above it only 1.3 eV. The transition temperature decreases when the density of dislocations in the material is less. On the other hand, the high-temperature precipitation rate decreases with decreasing dislocation density; both quantities were found to be proportional.

The latter result follows immediately from Penning's formula (8) when it is assumed that G is proportional to the dislocation density N_d, because then α_1 is proportional to N_d too.

At low temperatures the redistribution between substitutional and interstitional copper (or nickel) ions proceeds more slowly than the annihilation of vacancies at dislocations. When the dislocation density is sufficiently high, nevertheless the latter process still remains rate determining, but when there are only few dislocations, the rate of precipitation becomes ultimately governed solely by K and is then

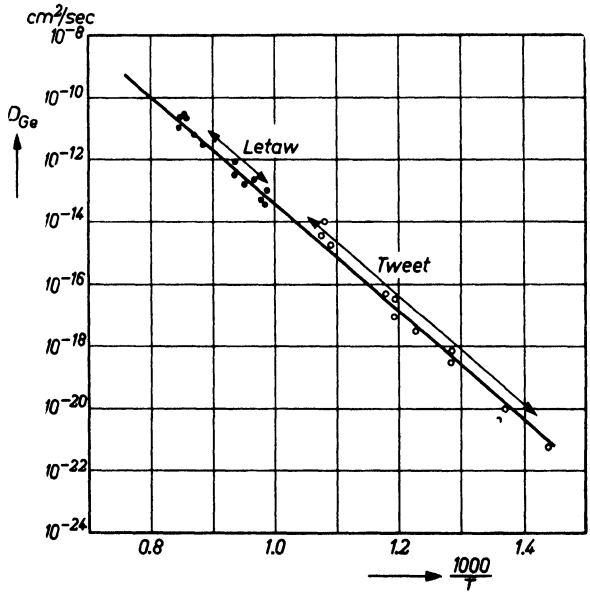


Fig. 4. Coefficient of self-diffusion derived from Tweet's measurements of the precipitation rate of copper in germanium, compared with the direct observational results obtained by Letaw. Both groups of data agree very well with a single activation energy of 2.96 eV (after PENNING [1958b]).

independent of the dislocation density. This explains the break in the temperature dependence of the precipitation rate.

The precipitation rate at large dislocation densities being given by

$\alpha_1 = G/(1 + K)$, it can, in the case of copper where K is of order 10, in good approximation be written as $G/K = G\,[V_0]/[N_{s0}]$. Now, in G the diffusion coefficient of vacancies D_v and the dislocation density must obviously be both contained in the form:

$$G = \text{const. } N_d D_v \; ; \tag{11}$$

the value of the numerical constant follows from the solution of the appropriate diffusion equation for the interaction between vacancies and dislocations. Therefore, the precipitation rate α_1 is under these circumstances directly proportional to the product $D_v[V_0]$, that is, to the coefficient of *self-diffusion*. From the dependence of the rate of precipitation of copper in germanium on the dislocation density, as has been observed by Tweet, values can thus be obtained for the coefficient of self-diffusion at relatively low temperatures, where actual diffusion measurements are impossible. (The ground of the apparent paradox lies in the establishment of a relatively stable quasi-equilibrium situation during the precipitation process, which is characterized by a very much enhanced vacancy concentration.) The results can be compared to high-temperature self-diffusion data, as were for instance obtained by LETAW *et al.* [1956], and fair agreement is obtained (Fig. 4, PENNING [1958b]). It will be clear that the activation energy for self-diffusion can be very accurately determined in this indirect way.

3. Diffusion and Dislocations

The role of dislocations in the precipitation of impurities lies evidently primarily in their interaction with *vacancies*, although, as was already observed earlier, dislocations also act as preferred sites for the precipitation of the impurities themselves in germanium. We shall see below how these two influences are connected.

Dislocations very easily become *decorated* with fairly large copper precipitates. This enabled DASH [1956] to render these imperfections visible by infrared microphotography. Some of the most beautiful examples of decorated dislocation-arrays that exist (e.g. active Frank-Read sources), were obtained by this technique (see Chapter V).

The large diffusion rate of copper and nickel in germanium when dislocations are present, together with the obvious preference of copper for precipitation at dislocation lines, leads at first glance to the idea

that the copper atoms can diffuse rapidly *along* the dislocation lines ("pipes", see Chapter XVIII).

This enhanced diffusion would, even in materials with relatively few dislocations, still dominate over the bulk diffusion. A detailed explanation along these lines of the structure sensitivity of the diffusion of impurities leads to great theoretical difficulties, however, especially when it is realized that, in order to obtain the necessary large overall migration velocities, the migration of copper along the dislocation lines must proceed at a rate comparable to or faster than the thermal velocity. This difficulty could only be avoided by assuming that the "dislocation pipes" are many tens of atom distances wide.

FRANK and TURNBULL [1956] suggested that the effect of dislocations on the diffusion coefficient of copper and nickel is not to be found in an enhanced migration of the copper atoms along the dislocations, but in the ease with which lattice *vacancies* can be *created* at dislocations. These authors remarked that the interstitial-substitutional equilibrium discussed in the previous section (formula (1)) will be shifted much to the substitutional side when large numbers of vacancies can be rapidly provided on the spot. This latter condition is indeed realized near dislocations. In the absence of these imperfections acting as sources for vacancies, the rapid diffusion of interstitial copper in germanium does not have much effect on the overall penetration of substitutional copper, since the interstitial equilibrium concentration is everywhere very much smaller than the substitutional equilibrium concentration. Only where enough vacancies are present, or can be provided at a sufficiently high rate, the transfer of copper from interstitial to substitutional sites can take place on a large scale. As the interstitial concentration is always very rapidly replenished, the local substitutional copper concentration will build up to its large equilibrium value. In this way the occurrence of precipitations near dislocations throughout the material can be understood.

The bulk diffusivity of the impurities in materials in which dislocations are present to a sufficient extent is largely determined by the ratio of the periods a copper atom spends interstitially and substitutionally respectively, thus by the ratio between the concentrations $[N_s]$ and $[N_i]$. Again assuming the substitutional diffusivity to be negligible in comparison with the interstitial diffusivity, D_i, one derives for the overall volume diffusion coefficient of copper atoms in *imperfect* material:

$$D_{\mathrm{imp}} = \frac{[\mathrm{N_i}]}{[\mathrm{N_s}] + [\mathrm{N_i}]} D_{\mathrm{i}} . \tag{12}$$

When enough vacancies can be provided, the ratio between concentrations can according to the argument of section 2 be replaced by the ratio between equilibrium concentrations (solubilities). From the known experimental data Frank and Turnbull inferred that $[\mathrm{N_{s0}}]/[\mathrm{N_{i0}}] \approx 100$ for copper in germanium at 1000 °K. The interstitial solubility is then only about 10^{-7}%.

The effective diffusivity of substitutional copper in material *without* dislocations is found as follows. In order to be able to migrate, the substitutional Cu-atoms must "dissociate" into interstitial Cu and vacancies. The former diffuse very rapidly, the latter slowly, and it is therefore the vacancies which determine the effective rate of migration. This *"dissociative diffusion"* is governed by the effective diffusion coefficient:

$$D_{\mathrm{diss}} = \frac{[\mathrm{V}]}{[\mathrm{N_s}] + [\mathrm{V}]} D_{\mathrm{v}} = \frac{D_{\mathrm{self}}}{[\mathrm{V}] + [\mathrm{N_s}]} . \tag{13}$$

It is this coefficient that also occurs in the penetration factor α_1, of the previous section. It is very much smaller than D_{imp} of formula (12).

4. Internal Friction in Germanium and Silicon

We conclude this chapter with a brief discussion of some characteristic damping properties of the elemental semiconductors. Much that concerns the damping of mechanical vibrations has already been mentioned in Chapter XVII.

Most damping investigations on germanium and silicon have until now been carried out at very high frequencies ($\approx$ 100 Mc/sec), that is, in the ultrasonic range. The attenuation, due to dislocations, of ultrasonic compressional and transverse waves in long bars was studied by GRANATO and TRUELL [1956]. It was found to be approximately proportional to the square of the frequency. As discussed in Chapter XVII, the energy dissipation per cycle is then proportional to the frequency. After diffusion of copper into the sample the damping decreased to about a quarter of its value (TEUTONICO *et al.* [1956]). This was ascribed to the influence of the precipitation of copper on the dislocations.

Similar observations were made on silicon, with analogous results.

According to the theory of GRANATO and LÜCKE [1956], discussed in Chapter XVII, this high-frequency attenuation should be ascribed to the damped resonance of dislocation segments. The frequency dependence indicates that one already closely approaches the resonance frequency in the 100 Mc range, as is indeed to be expected. At lower frequencies (kc range) one would expect this damping to become frequency independent. Measurements by KESSLER [1957] at 20–50 kc showed indeed such a constant damping to be present; however, at a 10 times higher level than expected from the extrapolation from the high-frequency range. Observations by BAKER *et al.* [1953] had shown an equally strong damping at 111 kc/sec. Therefore, the damped resonance hypothesis still awaits confirmation.

KESSLER [1957] studied the frequency, temperature and amplitude dependence of the internal friction in germanium in the 10 kc range in detail. He observed a peak in the logarithmic decrement at 40 kc and a temperature of 380 °C. The peak temperature shifted with the frequency, indicative of a relaxation mechanism. This was found to be characterized by an activation energy of about 1.1 eV and frequency factor of $10^{13}\,\mathrm{sec}^{-1}$. At temperatures above 500 °C the damping increased anew, independently of the frequency, exponentially with the temperature. This increase could be described by the same activation energy as mentioned above.

Kessler explained his relaxation peak in terms of the stress-induced diffusion of *vacancies* present within the stress field of dislocations (a kind of "Kösterdamping", see Chapter XVII). Small oscillations of dislocation segments cause changes in the distribution of vacancies over their "atmospheres" and by these changes energy is dissipated. Although good numerical agreement is obtained between the observed activation energy and the energies of formation and migration of vacancies determined in other ways, the high concentration of free vacancies that must be assumed to be present in the germanium samples together with their low diffusivity form serious objections against the proposed interpretation. Moreover, the absence of structure sensitivity on the damping is not easily accorded with a dislocation-vacancy mechanism. The situation is at present unresolved.

A different type of relaxation was observed by SOUTHGATE [1957] in silicon at about 100 kc/sec. This author observed a strong, fairly narrow relaxation peak near 1000 °C, and a much wider maximum

near 650 °C (Fig. 5). In germanium similar peaks have been observed at 700 °C and 410 °C (same frequency range). The first peak showed all characteristics of a relaxation phenomenon (activation energy 2.55 eV),

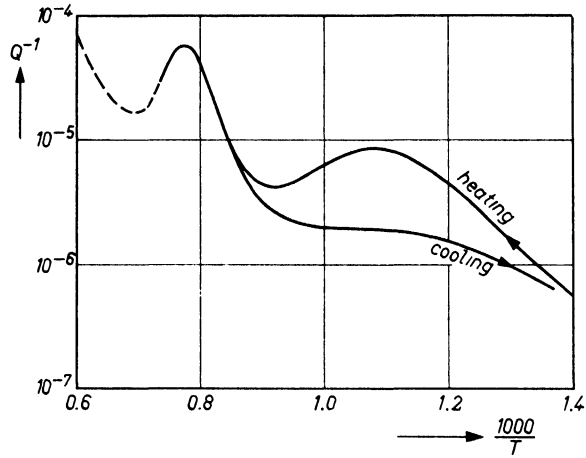


Fig. 5. Internal friction at 100 kc/sec of silicon (after SOUTHGATE [1957]). The well defined high-temperature peak is presumably due to the stress-induced diffusion of solute oxygen atoms. The much broader low-temperature maximum which is less reproducible and differs in height during heating and cooling, is caused by an electronic relaxation process.

and was attributed by its discoverer to stress-induced diffusion of interstitial solute oxygen atoms. If this interpretation is correct, this damping can be directly compared to the "Snoek-damping" produced by solute carbon in α-iron. Oxygen has indeed been shown to be present in large concentrations in silicon prepared from a melt contained in silica crucibles. However, it is not a priori certain whether oxygen actually occupies interstitial positions. A theoretical analysis of some infrared absorption bands associated with oxygen (HAAS [1957]) agrees best with the observations if it is assumed that the oxygen atoms show a preference for substitutional positions. However, the optical absorption has been studied at very low temperatures only, and a transfer of oxygen from substitutional to interstitial sites with rising temperature is not impossible. In view of the observed strong anisotropy of the damping, in the sense that excitation along a cubic direction does *not* result in oxygen internal friction, the interstitial oxygen model seems indicated. SEEGER [1959] has proposed as an alternative explanation the stress-induced motion of divacancies (which

indeed become mobile at about 700 °C in germanium, see section XXX–7.2). This is again a Snoek-type damping.

The second damping peak is much less reproducible than the high-temperature peak. SOUTHGATE [1959] has advanced an explanation in terms of an *electronic* rather than an atomic dissipation mechanism, hat is, of a new mechanism which has no counterpart in metals.

The rapidly varying stress induces periodic shifts in the energy band tance and thereby periodic alterations in the statistics of carrier tribution. These are brought about by the mechanisms of excitation d recombination respectively, and part of the recombination energy lost as heat. From an analysis of the damping characteristics, a value of 1.3 eV for the "deformation potential" in silicon (see Chapter XXVIII) was deduced.

REFERENCES

BAKER, W., L. SLIFKIN and W. MARX (1953); J. Appl. Phys. **24,** 1331.
DASH, W. C. (1956); J. Appl. Phys. **27,** 1193.
FRANK, F. C. and D. TURNBULL (1956); Phys. Rev. **104,** 617.
FULLER, C. S. and J. A. DITZENBERGER (1957); J. Appl. Phys. **28,** 40.
FULLER, C. S., J. A. DITZENBERGER and V. B. WOLFSTIRN (1954); Phys. Rev. **93,** 1182.
GRANATO, A. and K. LÜCKE (1956); J. Appl. Phys. **27,** 583 and 789.
GRANATO, A. and R. TRUELL (1956); J. Appl. Phys. **27,** 1219.
HAAS, C. (1957); unpublished.
KESSLER, J. O. (1957); Phys. Rev. **106,** 646 and 654.
LETAW, H., W. M. PORTNOY and L. SLIFKIN (1956); Phys. Rev. **102,** 636 and J. Phys. Chem. Solids **1,** 100.
LOGAN, R. A. (1956); Phys. Rev. **101,** 1455.
MAESEN, F. VAN DER and J. A. BRENKMAN (1955); J. Electrochem. Soc. **102,** 229.
MORIN, F. J. and H. REISS (1957); J. Phys. Chem. Solids **3,** 196.
PENNING, P. (1958a); Philips Res. Reports **13,** 17.
PENNING, P. (1958b); Phys. Rev. **110,** 586.
SEEGER, A. (1959); Solid State Physics in Electronics and Telecommunications, Vol. 1 (Acad. Press, London, New York).
SOUTHGATE, P. D. (1957); Proc. Phys. Soc. **B70,** 804.
SOUTHGATE, P. D. (1959); to be published.
TEUTONICO, L. J., A. GRANATO and R. TRUELL (1956); Phys. Rev. **103,** 832.
TWEET, A. G. (1957); Phys. Rev. **106,** 221.

NAME INDEX*

A

AARTS 285, 483
ABELES 123
ACHTER 426
ADAMS 173, 177
AIROLDI 290
ALBERT 458
ALEXANDER 613, *617*, 619, 630
AMELINCKX 68, 76, 133, 223, *435*, *504*, 507, 532, 541, *578*
ANDRADE *229*, 245
ANDREW 322
APKER 563
ARDLEY 186, *217*
ARONIN 271, 484
ASCHNER 514
ASSMUS 466
ATKINSON 334
AUGUSTINYAK 289, 644
AUST 444
AUSTERMAN 289
AVERBACH 199, 325, 499, 641
AVRAMI 457
AYTEKIN 224, 235

B

BACHELLA 290
BACON *338*
BAILEY 332
BAKER 657
BALL 441
BALLUFFI 299, 302, 412, 419, 431
BANGERT *387*
BARDEEN 412, 589
BARDSLEY 639
BARNES 310, 432
BARRER 414
BARRETT 27, 128, 136, 143, 221, 245, *254*, *473*
BARTLETT 307
BASSANI *541*, 544
BASSETT 133
BAUERLE 32, 41, *278*, 290, 305
BEAN 544
BEAR 183
BEAUJARD 223, 464
BECK *221*, 455, 467
BEEMAN 333
BELL, C. R. 446, 597
BELL, R. L. 254
BEMSKI 644
BENNEMAN 33
BERGHOUT 285, 392
BERNER 164, 178
BESHERS 373
BEVER 298
BILBY *59*, 109, *184*, *194*, *210*, *253*, 448, *632*
BILLIG 73, 620
BILLINGTON 3, 263
BIRCHENALL 414
BIRNBAUM 361
BISHOP *247*
BLADE 458
BLANK 231
BLATT *123*, 286
BLEWITT *284*, *302*, *311*, 373
BLIN 334
BLOEMBERGEN 317, *320*
BOAS 135, *249*, *298*, 568
DE BOER 554
BOHR 264
BOLK *418*
BÖMMEL 397
BONTINCK 68, 76, 507, 578
BORDONI 361, 365, 382

* Page numbers in italics indicate places where the authors' work is more extensively discussed.

Born *10*
Borrmann 335, 337
Bosman 415
Boulanger 237, 368, 395
Bowles 476
Bradshaw 290
Bragg 323, 436
Brammer 333
Brandon 139
Breckenridge 547
Brenkman 647
Brenner 507
Brinkman *273*, 291, 303, *306*, *415*, 484
Brommer 415
Brooks 445
Broom 226, *284*, 300
Bross *124*, 283
Brouwer 520
Brown, A. F. 138, 165
Brown, W. F. 343
Brown, W. L. 643
Buckley *240*
van Bueren *70*, *125*, 155, 204, 249, *284*, *292*, *300*, 332, 464, 580, *614*, 622
Bullough 104, 632
Burgers, J. M. *48*, 77, *100*, 144, 436
Burgers, W. G. 77, 427, 436, *455*, *468*
Burke 445, 455, 461

C

Cabrera *130*
Caffyn 541, 577
Cahn, J. W. 450
Cahn, R. W. 136, 164, *222*, *254*, 446, 463, 470
Campbell 203
Carpenter 135
Carruthers 126, 296
Carreker 617
Castner *559*
Chalmers 72, 197, *444*
Chang, H. C. 499
Chang, J. 643
Chang, R. 237
Chaudron 221, 458
Chen 136, 498
Childs 389
Churchman 613, 631
Clare 458
Clarebrough 298
Clark *198*
Cleland 642, 646
Clothier 287
Cochardt 185, 395
Cohen 199, *474*
Coltman *284*, 311
Compaan 412
Compton 514, 548, 564
Conwell 591
Coogan 242
Cooper 288
Corbett 289, 302
Cottrell 2, *57*, 66, 71, 109, 119, 161, 173, 177, *183*, 191, 194, *200*, *210*, 222, 224, 235, 253, *273*, 284, 307, 310, *374*, 382, 448, *478*, 490, 496
Couling *426*
Coulomb 62, 312, 455
Crawford 3, 642
Crookes *547*, 549
Crooks 514
Crussard 60, *186*, 212, *227*, 237, 251

D

Damask 309, 485
Darby 431
Darling 378
Darwin 128, 323
Dash 76, 129, 133, 505, 610, 654
Davis 233
Debye, P. (Sr) 17
Debye, P. (Jr) 591
Dehlinger 144
Dekeyser 68, 76, 223, 435, *504*, 507, 578
Denney 289, 478
Deruyttere 166
De Sorbo 291, 478
Detert *450*, 466
De Wit 80
Dexter 334, *574*, 577, 597, 640
Diehl *165*, 179
Dienes 3, 35, 39, *260*, 307, 309, 311
Dietrich 345
Dietze 63, *118*, 145, *345*
Dijkstra 388
Dimitrov 458
Ditchburn 581
Ditzenberger 647
Dixon 291, 306, 484
Donth *365*
Doo 419
van Doorn *557*

DORENDORFF 130
DORN *200*, 234
DOYLE 388
DRUYVESTEIN *285*
DUGDALE 289, 431, 483
DUNN *444*, 455

E

EDWARDS 243
EGGEN 288
EGGLESTON 289, 305
EINSPRUCH *361*
ELAM 135
ELLIS 635
ENTWISTLE *240*
ERICKSON 195
ESHELBY 60, *70*, 80, 94, *104*, *116*, 212, 255, 507
ETZEL 40, 514, 544
EWING 462
EYKELENBOOM 415
EYRING 408

F

FAN 642
FARA 412
FAST 378, 388, 466
FEDERIGHI 279, 290
FELBECK 499
FELDMAN 508
FELDMANN *613*, 618, 628
FELTHAM 233
FESHBACH 105
FINKELSTEIN 388
FISCHBACH 542
FISHER 2, 116, *200*, 216, *423*, *444*, 460, 494
FISKE 289
FLETCHER 643
FOREMAN 63, 76, 118
FOURNET 333
FRANK *45*, *70*, 77, 130, 257, 432, *435*, 452, 476, 502, 505, 507, 633, *655*
FRANKL 574
FRANZ 480
FRENKEL 143
FRIAUF 514, 549
FRICKE 335
FRIEDEL 2, 62, *70*, *155*, 162, 216, 236, *249*, 312, 368, 485, *494*
FRÖHLICH 544
FUCHS 269
FUJINO 575
FUJITA 490
FULLER 597, 647
FULLMAN 452, 471
FUMI 32, 306, *513*, 544
FUSSFELD 376

G

GALLACHER 613
GARBER 286
GARSTONE 170
GAY 326, 331
GEACH 613, 631
GERMAGNOLI 290, 306
GEROLD *333*
GERRITSEN 283
GIBSON 35, 263
GILLIES 311
GILMAN 71, 129, 145, 245, 256, 275, 495, 498, 579, 625
GLASSTONE 408
GLEN 284, 300, 484
GOERING 390
GOLAND 35, 263
GONSER 274, 479
GOODFELLOW 541, 577
GORDON *377*, *574*
GORMAN 288
GORUM 568
GRAF 430
GRANATA 291
GRANATO *368*, 377, *656*
GRANT 499
GRAY 2
GREENOUGH 167
GREETHAM 170
GREINER 635
GRIFFITH 492
GROSS 355
GRÜNDIG 514
GRUENEISEN 10
GUARD 139
GUGGENHEIM 27
GUINIER 333, 430, 480
GURNEY 34, 39, 513, 533
GYULAI 541

H

HAAS 658
HAASEN 138, 155, 186, 200, 205, 245, 300, 334, 603, 608, *613*, *630*, 631, *636*

HAAYMAN 524
HAHN, G. T. 499
HAHN, R. *387*
HALL, E. O. 255, 452
HALL, G. L. 114
HALL, W. H. 128, 329
HAM 211, *226*
HANTELMANN 514
HARDY 477
HARGREAVES 298
HARPER 388
HARRISON 262, 514
HART 196, 203, 250
HARTLY 541
HARTMANN 249
HARTWIG 337
HASHIMOTO 133
HASIGUTI *305*, 385, 641
HAVEN 39, 412, 514, *547*, 557
HAYFIELD 245
HAYNES *426*
HAZLETT 216
HEAD 491
HEAL 477
HEDGES 133, 578
HEIDENREICH 51
HENDRICKSON *198*, 429
HENNIG 285
HERRING 237, 412, 589
HEUMANN *421*
HIKATA *377*
HILL *247*
HIRAI 385
HIRSCH 67, 128, 133, 148, 227, 234, 280, *322*, 333, 343, 441, 451
HOFFMAN *425*
HOGARTH 597
HOLDEN 309
HOLLOMON 73, *460*
HOLMES 287, 312, *373*
HONEYCOMBE 165, 170
HORN 343
HORNE 133
HORNSTRA 77, 437, 442, 464, *599*, 614, 637
HOUSTON-MCMILLAN 483
HOVE 549
HUANG 12
HULL 490
HUNTER, L. P. *335*
HUNTER, S. C. 124, 191
HUNTINGTON 16, 36, *260*, *303*

I

IBE 466
INGHAM 289
IRMLER 337
ISHIGURO 575
ITOH 552

J

JACKSON 507
JAGODZINSKI 480
JAOUL 136, 251
JARVIS 285
JASWON 63, 118, 197, 382
JOFFÉ *569*
JOHNSON, R. P. 413
JOHNSON, W. A. 457
JOHNSTON 2, 71, 129, 147, 275, 498, *579*, 625
JONES 27
JONGENBURGER *123*, *285*, *302*
JOST 414

K

KAISER 611
KANZAKI 114, *569*, 579
KÄNZIG *559*, 574
KATZ 514
KAUFFMAN 276, 290
KAUFMAN 476
KAWAMURA 320
KÊ 366, *377*, *382*, *388*
KECK 611
KELLY 178, 186, 200, 245, 326
KELTING 514
KEMP 296
KENNEDY 233
KERSTEN *343*
KESSLER 657
KEYES 414
KIMURA 73, 279, 312
KINCHIN *260*, 271, 304
KIRKENDALL 418
KITTEL 397
KLEIN 568
KLEMENS 125, *296*
KNELLER 345
KOBAYASHI 581
KOCHENDÖRFER 135
KOCKS 248

KOEHLER *34*, 41, *53*, 100, *260*, *275*, 278, *287*, 299, 369, 484, 581, 643
KONITZER 558
KÖSTER 343, *377*, *387*, 480
KOTTMANN 421
KRAMER 200
KRAMERS 409
KRÖGER *516*, 560
KRONBERG 532, 609
KRÖNER 80, 94, *104*, 249, 442, 480
KRONMÜLLER 346, 388
KROPSCHOTT 286
KUHLMANN-WILSDORF 73, 76, *138*, *224*, 275, *279*, 312
KULIN 641
KURNICK 569
KUROSOWA 513
KURTZ 641

L

LACOMBE 129, 223, *422*, 455, 464
LAIDLER 408
LAMB 458
LANG 337
LANGE, C. F. 611
LANGE, H. 138
LARK-HOROWITZ 642
LAUBENSTEIN 288
LAVES 480
LAZAREV 290
LE CLAIRE 390, *405*
LEE 431
LEIBFRIED *12*, 63, 71, 119, 138, 155, 245, 255, 272
LEIDER 578
LESHI 470
LETAW 34, 636, 643, 654
LEVI 312
LEVY 312, 361
LICHTMANN 246
LIDIARD *543*, *548*, 576
LIEH HUANG 234
LITTLETON 513, 528
LIU 464
LOGAN 638, 643, 646
LOMER, J. 127, 296
LOMER, W. 66, 161, 203, 307, 334, 390
LONSDALE 340
LOUAT 187, 195
LOVE *80*
LOW 129, 139, 195
LÜCKE 119, 138, 166, 353, 369, 377, *450*, *657*
LUTTS 224
LYTTON 234

M

MA 393
MACHERAUCH 249
MACHLIN 129, 429, *491*, 498
MACKENZIE 143, 476
MADDIN 73, 136, 279, 312, 431
MADER *138*, *164*, 170
MAESEN, VAN DER 647
MAGNUSSON 283
MAKIN 311
MANINTVELD 285
MANN 33, 112, 300
MAPOTHER 514, *547*, 549
MARCHIN 129
MARKHAM 558
MARTIN 289
MARX 288, 657
MASING 136, 164, 224, 240, 458, 471
MASLENIKOV 246
MASON *362*, *368*, *383*, 397
MATARÉ 639
MATSUURA 641
MAUKSCH 240
MAURER 40, 514, *544*, *548*, 564
MCCAMMON 491
MCGEE 289
MCKELVEY 641
MCKEOWN 289
MCLEAN *236*, 331, *435*
MCREYNOLDS 289
MEECHAN 291, *303*, 484
MEHL 457
MEIJERING *388*, 415, 467
MENDELSSOHN 297
MENTER 133
VAN DER MERWE 505
METZGER 306, 312
MICHELL 298
MILGRAM 35, 263
MISEK 389
VON MISES 246
MITCHELL 133, 564, *578*
MOELLER 568
MOLENAAR 285
MOLLWO 560
MORIN 597, 636, 648

MORRISON 642
MORSE, P. M. 105
MORSE, R. W. 396
MOTT 27, 34, 56, 70, 149, *160*, 169, *213*, *225*, *231*, 249, 332, 367, *384*, *435*, 446, 461, 489, 513, 528, 533
MÜLLER 138

N

NABARRO *56*, *63*, 71, *80*, 100, *112*, *124*, 191, *213*, *231*, 367, 507
NACHTRIEB 414
NEWKIRK 337
NEYNABER 333
NIBLETT *362*
NICHOLAS 77
NOGGLE 287, 312
NOWICK 224, 306, 312, 349, *359*, *377*, *390*, 432, 542, *574*
NUOVO 365, 382
NUTTING 139
NYE *107*, 223, 325, *568*

O

ODA 575
OGILVIE 249, 450
OKADA 641
OKKERSE 274, 427, 479
VAN OOSTERHOUT 524
OROWAN 144, *215*, 234, 254
OTSUKA 320
OVCHARENKO 290
OVERHAUSER 288, 299
OWEN 199

P

PALMER 283
PANETH 34
PANSERI 279, 290
PARKER 147, 216, 446, 568
PARTRIDGE 492
PASHLEY 133
PATEL *613*, *617*, *619*, 630
PATERSON 128
PAXTON 173, *183*
PEACH 53, 59
PEARLSTEIN 289
PEARSON, G. 508, 613, 618, 628, 636, 641
PEARSON, S. 290
PEARSON, W. 382
PEASE *260*, 271, 304
PEIERLS 27, 63, *116*
PENNING 71, 73, 203, 597, *614*, 619, *620*, 628, 637, *648*
PERRYMAN 224
PETCH 250, 493
PFANN 146
PICK *562*
PIPPARD *397*
PITSCH 332
PLATEAU 238
POHL 554
POLANYI 144
PORTNOY 654
POUND 319
POWERS 388
PRATT 514, 540, *568*, *575*
PRESTON 480
PRY 285
PUGH 309
PURCELL 317, 321

Q

QUINNEY 246

R

RACHINGER 226, 238
RAFFELSIEPER 224
RANDOLPH 289
RATHENAU 415
READ 2, 59, 77, 188, 193, 280, *435*, 508, *593*, 613, 618, 628, 636
REBSTOCK 136, *164*
REDDING 310
REDMAN *284*
REED 639
REHBINDER 246
REINACKER 393
REISS 597, 648
REMAUT 541
RICHTER 216
RIECK 330, 466
RIGDEN 578
ROBERTS 310
ROMEYN 524
ROSENBERG 127, 296, 491
ROSENBLATT 289, 309
ROSENHAIN 462
ROSWELL 306, 312
ROTH 620

ROTHERHAM 382
ROWLAND 320
ROZIN 388
RÜCHARDT 581
RYDER 242

S

SAADA 160
SACHS 240, 247
SACK *366*
SALKOVITZ *366*, 373
SAMPSON 299
VAN SANTEN 39
SCHAMP 514
SCHEIL 393
SCHILLER 388
SCHINKEL 415
SCHMID 135, 143, 568
SCHOECK 163, *176*, 185, 189, 195, 230
SCHONFELD 568
SCHRÖDER 167
SCHULTZE 503
SEEGER *33*, 55, *61*, *69*, *80*, *123*, *136*, *143*, *155*, *163*, *189*, 216, 226, 259, 274, 288, *299*, 303, 311, 332, *361*, 376, 388, 474, 514, 530, 608, *613*, *631*, *636*, 658
SEGALL 492
SEITH 414, *418*
SEITZ *2*, 34, *40*, *68*, *260*, *278*, 299, 305, 414, 484, 507, *514*, 530, 541, 543, *544*, 551, *563*, 564, *572*, 578, 581, 597, 640, 643
SHEPHARD *200*, 234
SHERBY 234
SHEWMON 416
SHOCKLEY *51*, *415*, *435*, *589*, 592, *597*
SHOJI 240
SIEGEL 484
SILCOX 133, 281
SIMMONS 299, 302
SLICHTER 321
SLIFKIN 654, 657
SMAKULA 568
SMALLMAN 133, 281, *329*
SMEKAL 324
SMIGELSKAS 418
SMIT 249
SMITH 8, 104
SMOLUCHOWSKI 289, 309, *426*, 436, 447
SNOEK 189, *386*, 390
SONDHEIMER 26
SOSIN 287, 303
SOUTHGATE *657*
SPERRY 221, 456
SPRACKLING 568
STASIW 528
STAUBWASSER 138, *166*
STEHLE *125*, 305, 334
STEINBERG 200
STEVELS 359
STOKES 177
STROH 257, 489, 493, 633
SUITA 552
SUITS 129
SULLY 230
SUMINO 189
SUZUKI, H. 188, 216, 254
SUZUKI, T. 451, 570

T

TAFT 563
TAINSH 296
TAMHANKAR 238
TARPINIAN 289
TAYLOR, G. I. 100, 144, 168, *246*
TAYLOR, W. E. 620
TEGHTSOONIAN 72
TEISSIER DU CROS 450
TELTOW 514
TEUTONICO 656
TEWORDT 33, 34, 300
THOMAS, D. A. 499
THOMAS, G. 430
THOMPSON, D. O. *373*
THOMPSON, N. 77, 233, 242, 489
THOMSON 2, 541
THORNTON 234
TICHELAAR 414
TIEDEMA 427, 464
TIEN-SHIH-LIN 200
TIPTON 568
TITCHENER 298
TOMIZUKA 431
TOSI *513*
TRAUTZ 474
TREUTING 613
TRUELL 361, 656
TSIEN 389
TUCKER 299
TURNBULL 291, *425*, 432, *445*, 455, 460, 478, *655*
TWEET *634*, *648*

U

Ueta *557*, 574
Uyeda 133

V

Varley 270, *553*, 561
Vennik 541
Verbraak 468, 476
Verdini 365
Verma 504
Verrijp *378*
Verwey 524
Vineyard 3, 35, 39, *260*, 311
Vink *516*, 651
Vogel 606, 613
Vogt 589
Volger 565
Volmer 503
Volterra *80*
Vook 299, 302
Vreeland 2, 199, 203

W

Waber 568
Wadsworth 242, 489
Wagner 413, 514
Walker 289, 302, 314
Walter 447, 455
Walther 421
Wang 389
Warren 325
Washburn 147, 242, 446, 568
Wassermann 216
Watkins *319*
Webb 449
Wechsler 290
Weertman *234*, *366*, 373
Weinberg 447
Weingarten 110
Weill 238
Weinrich 639
Weiss 342
Wert *299*, 388
Wertheim 641
West 298
Westmacott 133, 281
Whelan 128, 133, 343, 430
Whipple 423
White 298
Wiedersich 185
Wilks *362*
Williams 245
Williamson 128, *329*
Wilsdorf 76, *138*
Wilson 27, 128
de Wind *615*, 628
Winston 613, 631
Winterberger 279, 291
Witt 514, *543*, 549
Wolf 163
Wolfstirn 647
Wood, D. S. *198*
Wood, G. K. *63*, 118
Wood, W. A. 226, 238, 331, 451
Woodruff 560
Woolley 240
Wu 389
Wyatt 231
Wyon 129, 238

Y

Yang 389
Yukawa 428

Z

Zener 116, *349*, *366*, 385, 474
Ziman 27, 125

SUBJECT INDEX*

A

Abnormal after-effect 245
Absorption line (nuclear spin resonance) 317
Absorption tail 577
Accelerated creep 238
Accelerated ordening *308*, 431, 483
Acceptor *585*, *636*, *642*
Activation volume 74, 232, 414
Activity coefficient 407
Additive colouring 555
Age-hardening *477*
Alloy-hardening 483
Amplitude dependence of internal friction *367*
Anchoring of dislocation 66, *311*
Andrade creep *230*
Anisotropic centre *559*
Anisotropic resistance *286*
Annealing *220*
Annealing twin *468*
Anomalous transmission *335*
Antiphase region 218
Antiplane strain 89
Anti-Schottky defect *513*
Anti-structure disorder *516*
Associated defects *33*, 279, 413, *513*, *549*, *643*
Asterism 221, *325*
Attenuation 353, 396

B

Back stress 153
Bad crystal *45*
Band gap 584
Band of secondary slip 136
Barkhausen jump 344
Barrier for dislocations 66, 149, *155*, *169*, 205, 229, 234, 243
Bauschinger effect *240*, 568
Beltrami equations 86, 108
Bend planes *256*
Berg-Barrett method 337
Binding energy *187*, 203, 211, 398
Bipotential equation 89, 108
Birefringence *568*
Bleaching *554*
Bloch wall *344*
Boat form 600
Bordoni damping *361*
Born approximation 264
Boundary dislocation 606
Bravais lattice 7
Breakdown 551
Break-loose of dislocation *196*, *374*
Brinkman spike 274, 290, *479*
Brittle fracture *492*
Brittleness 310
–, hydrogen 493
–, low temperature 496
Brittle-to-ductile transition 498
Burgers circuit *45*, 107
Burgers vector *47*, 110

C

Carrier concentration 589, 592, *634*
Cascade *262*
Cauchy relations 15
Cell 226, 331, 451
Centre (colour –) 322, *511*, 532, *553*
Centre of compression, dilatation 92, 111
Chair form 600
Charged dislocation 540, 564
Charged jog 530
Cleavage 494, 568
Climb *52*, *63*, *220*, *232*, *304*, 313, 378, 463, 573, 580
Close pair 273
– – recombination *303*

* Page numbers in italics indicate places where a subject is treated more in detail.

Clusters of defects 35, 280, *310*, 562, 581
Cobalt transformation *474*
Coercive force *343*
Coherent precipitation 481
Coherent twinboundary *251*, 451
Collision cross-section 42, *260*
– – for deuterons 265
– – for electrons 270
– – for neutrons 261
– – for protons 265
Coloration by irradiation *572*, 581
Compatibility equations 84
Complex 545, 549
Composite dislocation 604
Compton effect 270
Concentration-pressure chart 521
Condensation of vacancies 72, 505, 580
Conduction band 518
Conduction loss 359
Conductivity, extrinsic 543, 586
–, intrinsic *544*, 569, *583*
Constriction 66, 162, 175, 231
Continuum theory of dislocations 104
Controlled valency 524
Coordination number *11*
Core-rearrangement *602*
Correlation factor 413, *527*, *548*
Cottrell-Lomer barrier 66, *169*, 234, 243
Coulomb interaction 263
Crack *489*
Creep 52, 141, *229*, 310
Creep of germanium *614*, *622*
Critical nucleus 501
Critical shear stress for slip *141*, 247, 310
– – – for twinning 254
Cross slip *138*, *163*, 173
Crowdion *34*, *271*, 311
Crystal anisotropy, magnetic 345
– –, elastic *15*
Crystal classes, systems 7
Crystal growth 73, *500*
Crystal momentum 588
Cusp 444
Cyclotron resonance 590

D

Damped eigenvibrations 360
Damped resonance *356*
Dangling bond *592*
Darkening 553
Debye-Hückel equation 540
Debye model *18*
Debye temperature, frequency *18*, 533
Decoration *129*, 564, 578, 654
Defect accumulation 490
Deformation band 136, 161, 256, 327
Deformation hysteresis *378*
Deformation loss 359
Deformation potential 589, 659
Degeneracy 23, 587
Degree of association 547, *549*
Delay period 154, *198*, 203, 614
Delayed yield *198*, 568
Density of sources 286
Depletion zones 312
Development (photographic) 564
Dichroism *557*
Dielectric loss *545*, *574*
Diffusion *405*, 525
– coefficient *405*
– –, partial 412
– –, inter- 418
–, effect of pressure on 414
– in ionic crystals *547*
– of point defects 52, 225, 236, 272, *305*, *404*, *411*
– in homopolar crystals *646*
Diffusion, chemical 412
–, dissociative 438, 656
–, grain boundary *384*, *422*
–, dislocation 307, *428*
–, thermal *415*
–, enhanced 321, *429*, 480
Diffusionless phase change 472
Di-interstitial 36, 305, 639
Dilatation 54, *86*, 184
Dipole moment 544
Direct exchange *407*
Dislocation, concept *45*
–, composite 132
–, charged 540, *593*
–, edge *48*
–, extended 50
–, half *47*
–, helical 68, 507, 578
–, mixed 49
–, multiple 504
–, partial *47*, 65, 607
–, perfect *47*
–, prismatic 68, 280
–: R-dislocation 73, 281
–, rotational 111
–: source of defects 40

–, Somigliana *110*
–, translational 111
–, twinning *252*
–, Volterra 47, 110
Dislocation acceptor *594*
– climb 52, *63*, *220*, *232*, *304*, 378, 463
– core *55*, *116*, *329*
– damping *312*, *361*
– density 64, 74, *146*, 156, *328*, 345, 575, 653
–, diffusion through 307, *428*
– dissociation 50, 66
– distribution *107*, *329*, 440
– energy 55, 119
–: electromagnetic analogy 59
–: electron scattering 128
–: electrical resistivity 124
– etchpits 129, 579, *619*
–: flexibility 213
–, force on *56*, 112, 144
–: geometrical theory *45*
– hysteresis *372*
– interaction *57*, *145*
– –, chemical *188*, 207
– –, elastical 59, *186*
– –, electrical *191*
– –, hydrostatic *59*
– jog 61, *64*, *225*, 421
– kink *205*, 364
– line tension *56*, *74*, *200*, *368*
– locking 66, *311*
– loop 73, 133, *275*
– magnetoresistance *126*
–: mathematical theory *64*, *99*
–, motion of *62*, 217, 242, 539, 579, 625
– multiplication 625
– network 47, 72, 147, *327*, 579
– notation 77
–: optical detection *129*
–, origin of *72*
–: paired dislocations 218
–: partial splitting *47*, 532, 607
– pile-up 49, *160*, 234, 249, 329, 489
– pipes 429, 659
– in polar crystals *528*
– reaction 66
– relaxation *361*
– resistivity *124*, *286*
– resonance *368*
– row of acceptors *592*
– in specific lattices 78
– stresses and strains *54*, *64*, *99*
–, superextended 218
– in valence crystals *599*
– velocity 70, 130, 190, 255, 498
– width 63, *118*
–: X-ray scattering *127*
Dislocation-free crystal 610
Dislocation-induced internal friction 488
Disordering phenomena 286, 305, 431
Disorientation angle *427*
Displacement by irradiation *260*
Displacement function 104
Displacement spike 260, *272*
Dissipation of energy 267, *350*, 544
Dissociation energy 34, 279
Dissociative diffusion *656*
Distorsioni 99
Distortion tensor 106
Divacancy 34, 277, 306, 610, *637*
Domain of order 218
Donor *585*
Double bond *603*
Double diffraction 333
Double force 92, 109
Double jog 572
Double layer 600
Double vacancy 610, 637
Drift velocity 26, 409
Ductile fracture 492
Dynamic recovery 175

E

Easy glide *166*, 311
Effective mass 20, 588
Einstein equation 406
Einstein model *17*, 38
Einstein relation 527
Einstein temperature, frequency 10, *17*
Elastic constants *14*
Elastic modulus *14*, *84*, *349*
Elastic singularity 80, *90*
Electrical resistivity
– – of perfect lattice 25
– – of defect lattice *123*, 283
– – of cold-worked metal *284*
– – of irradiated metal *287*
– – of quenched metal *290*
– – of dislocation *124*, *286*
– – of point defect *126*, *287*
Electrolytic conduction *526*, 541
Electron diffraction 332, 343
Electrostatic force 12
Elementary structure 139, 170

Endurance limit 488
Energy of association of point defects *34*, 638
Energy of formation of point defects *31*, 37, 41, 69, 276, 305, *636*
Energy of migration of point defects *35*, 290, 305, 638
Enhanced diffusion 321, *429*, 480
Enhanced nucleation 431
Entropy of activation 461
Epitaxy 305, 452, 505
Equilibrium conditions 85
Equilibrium of defects *515*
Exchange force 13
Exciton 535, *555*, 573
Exciton band 559, 578
Extended dislocation 50
Extrusion *488*

F

F, F(M) band *555*
F-centre *535*, *551*, 574
F'-centre 562
Fatigue *487*
Fermi-Dirac statistics 586
Fermi-energy 13, 23, 32, *395*, 525, 586
Ferromagnetic hysteresis 395
Fick's law 304, 424
Fission fragments 267
Flow stress 141, *148*, *212*, 310
Focussing collision *263*, 272
Force dipole *92*, 109, 115, 185
Fracture *492*
Fragmentation *227*, 451
Fragmented slip bands 163, 177
Frank-Read source *70*, 111, 147, *153*, 254, 504
Free energy of perfect crystal *16*
– – of defect crystal *36*
Free surface energy 455
Frenkel defects 28, *36*, *512*, *563*
Fundamental equation of dislocation theory 107

G

Gaspermeation 412
Gaspockets 309
Glide *52*, *63*, *135*
von Göler-Sachs equation 456
Good crystal *45*
Grain boundary 150, 183, 427, *435*
–: dislocation model 427, *436*
–, diffusion through *384*, *422*, *461*
– energy of *442*
– large angle 446
–, motion of *446*
–, interaction of *447*
– migration 236, 462
– relaxation *383*
– sliding 246, 385, 467
Grain growth *454*
Griffith crack 492
Growth steps and spirals *502*
Growth twin 471
Guinier-Preston zone 216, 335, 393, 430, *470*

H

H-band, centre 561
Haasen-Kelly effect 244
Halfdislocation 51
Halfjog 638
Hall effect 512, *590*
Hardening, precipitation 211, 310
– by order 216
–, inter- and intragranular 250
–, solid solution 214
Heat conductivity, lattice *19*
– –, electronic *25*
Helical dislocation 68, 507
Heyn stress 240
High-temperature dislocation damping *366*
Hole 24, 517, 533, *584*
Hollow dislocation core 504
Hooke's law 84
Hysteresis 343, *355*

I

Impurities, association with point defects *40*
–, effect on damping 386
–: binding energy *187*, 203, 398
Impurity atmosphere *186*, *195*
Incipient vacancy *530*, 565, 572
Incompatibility tensor 108
Incubation period *455*
Incubation time 614
Infrared absorption 611
Initial period *615*
Internal friction *348*
– – by dislocations *361*

– – by grain boundaries *383*
– – by impurity atoms *386*, 658
– – by phase transition *392*
– – by point defects 382, *385*, 657
– – in ionic crystals *574*
– –, magneto-elastic 394
– – in germanium and silicon *656*
– –, magneto-acoustic 395
– –, thermo-elastic 394
– –: influence of irradiation 373
Internal stress 168, *213*, *224*
Interstitial atom *32*
Interstitialcy migration 304, 549
Interstitial-vacancy pair 36
Intrusion 490
Ionic conductivity 525
Ionic radius 13
Ion-ion repulsion 12, 33
Iron-carbon diagram 209
Irradiation *41*, *260*, 478
– by γ-rays 270, *642*
– by deuterons 263
– by electrons *268*, 642
– by neutrons *261*, *642*
– by photons *551*
– by protons 263
– disordering 271
Irreversible strain 375
Island model 436

J

Johnson mechanism of diffusion 412, 430, 550
Joffé effect *569*
Jog 61, *64*, *225*, 421
–, effective charge of 192, *334*
–, energy of 175, 312
Jump frequency 387, 415, *546*

K

K-band, centres 554
Kinkband *136*, 149, 165, 256, 327, 633
Kink in dislocation 205, 363, 381, 500
Kinking of hexagonal crystal 140, *255*
Kirkendall effect 390, *407*, *417*
Köster damping 387, *391*, 657
Kurdjumov-Sachs relation *473*

L

Lattice curvature 222, *325*
Lattice structure *8*, *29*, 77
Law of constant resolved shear stress 143
– of Dulong and Petit 17
– of mass action 516
Lifetime 512, *591*, 597, *640*
Linear visco-elastic solid 349
Lithium transformation 474
Local recovery 464
Lock 626
Logarithmic creep 230
Logarithmic decrement *352*
Lüders band 183, 203, 249

M

M- band, centre *556*
Madelung number 12, 513
Magnetic aftereffect 394
Magnetic properties *343*
Magnetic saturation 343
Magneto-acoustic resonance 395
Magneto-elastic damping 394
Magnetoresistivity *126*, 292
Martensitic transformation 252, 452, *470*
Matthiesens rule 283
Microbeam 326, *330*
Microcrack 492
Microcrack in dislocation 603
Microcreep *194*, 382
Minority carriers 591
Mobility 26, 405, 512, 538, 564, *589*, *633*
Moiré effect 133
Mosaic structure 128, *322*
Motional narrowing 321
Moving-jog mechanism 64, 155
Multiple slip 137

N

N-band, centre *556*
Natural whiskers 508
Neutrality condition *519*
Neutral jog 530
Neutron damage 42, *261*, 580
Neutron diffraction *337*
Neutron transmission *340*
Nishiyama relation 473
Normal mode 16
Nuclear spin resonance *316*
Nucleation of grains *208*, *455*, *460*
–, heterogeneous 462
–, homogeneous 456
–: critical size 460
Number of carriers 533, 589

O

Octahedral position 29
Order 189, 271
–, short range *216*, 309
–, long range *216*
Order parameter 481
Order of twin 252
Order-disorder transformation *481*
Ordering kinetics *482*
Ordering phenomena 286, 305, 389, *393*, 431
Oriented growth theory 467
Oriented nucleation theory 467
Over-ageing 213

P

Pair production 271
Paramagnetic resonance 322
Particle size *324*, *329*
Partition function *16*
Peierls force 34, *63*, *117*, 145, 169, 235, *363*, 512
Pencil glide 139, 568
Perfect crystal *7*
Permeability *343*
Phase lag 350, 544
Phase transformation 393, *470*
Photoconductivity 563
Photoelectric effect 270
Photolytic darkening *563*
Pile-up 49, *160*, 234, 249, 329, 489, 572
Plane strain 88
Plasticity spike 484
Plasticity of homopolar crystals *612*
Point defect internal friction 385
Point group 7
Poisson distribution 370
Poisson equation 540
Polarization 12, 32, 513, *557*
Pole mechanism 253
Polycrystalline metals *246*
Polygonization *219*, *325*, 451, 463, 569
Polytypism 506
Pores 417
Porosity zone 418, 419
Precipitation *221*, 335, 343, 393, *477*, 647
Pre-yield creep *198*
Primary extinction *323*, 337
Primary knock-on *260*
Primary recrystallization 454
Primitive vector 7
Projection topography 337
Pulse attenuation method 360

Q

Quadrupole interaction *318*
Quality factor 351
Quenching *279*, 305, 636

R

R-band, centre *556*
Radiation damage 43, 260, *275*
Radiation damping 70
Radiation growth 309
Radiation hardening 275, *311*
Radius of gyration 333
Range of particles *267*
Rate of formation of dislocations 156
– – – of point defects 158
Reactor neutrons 337
Rearrangement, electronic *602*, *625*
–, diffusion *602*, *625*
Recombination 591, *636*
Recovery, mechanical 52, 220, *223*, 232, 376, 457, 463
Recovery, electrical 283, *300*
Recovery of damping *377*
Recovery stages *302*
Recrystallization 52, 221, 376, 392, *454*, 489, 569
–: driving force 461
– texture *465*
– in situ 227, 451
Rehbinder effect 246, 569
Relativistic particles 266
Relaxation 31, 114, *348*, 414, 513, 577
– function 354
– spectrum 354
– time 24, *318*, *349*, 545, 589
Relaxed modulus *350*, 385
Replacement collision 271, 484
Resistivity increase of irradiated metals 289
–: of cold-worked metals 285
–: of quenched metals 291
Resonance scattering 338
Retrograde mass flow 421
River 495
Rocking curve *325*
Rotation 86
Row of vacancies 47, 58, 75

Rowland transformation 468
Rutherford scattering *264*

S

Satellite line 319
Schottky defect 28, *36*, *512*, 564
Schottky disorder *516*
Screening *32*, 191, 264
Secondary displacement *262*
Secondary extinction *324*, 340
Secondary recrystallization 455
Segregation of impurities 449, 505
Sensitizer 565
Shockley half-dislocation 47, 474
Simple dislocation 600
Single slip 135
Sintering 237, 465
Size effect 245, 569
Slip *49*, *134*, *144*
Slip line, band 135, *311*, 366, 619
Slip system *49*, *135*
Small-angle scattering *332*
SN curve 487
Snoek damping *386*, 391, *658*
Snoek ordering 189, 195, 217
Solubility 208, 388, 650
Source function 108
Space charge 530, 572
Space charge cylinder *593*
Space group 7
Specific heat, lattice *17*, *295*
– –, electronic *24*
Spike 284, 484
Spiral growth 503
Spiral prismatic dislocation 507
Spotty arc 330
Stacking fault 50, 65, 128, *474*, *607*
– –, energy of *162*, 175, 226
– –, resistivity of *124*
Standard linear solid 349
Static after-effect 361
Steady state creep *233*, 331, 615
Stored energy *298*, 329
Strain 14, 31, *53*, *81*, 348
Strain ageing *208*, 377, 388
Stress 14, 31, *53*, *81*, 348
Stress-dependent activation energy 174, 206
Stress function *88*, 100, 101
Stress-induced diffusion 392
Stress-induced ordering *389*
Stress magnification 493
Stress-strain curve 141, *164*, 310, 618
Stress-strain relations 84
Striae 136
Structural disorder 459
Structure factor 459
Structure feuilletée 450
Structure sensitive precipitation 652
Structure sensitive properties 1, 316
Subboundary 223, 237, *450*, 463
Subgrain 327
Substitutional-interstitial equilibrium *649*
Superstructure 218, *480*
Surface dislocation 632
Surface effect *245*
Surface energy, tension 445, 462, 501
Surface hardening 577
Surface source *570*, *624*
Susceptibility *343*
Swelling 309

T

Tangle *313*
Temperature-activated hysteresis *382*
Temperature spike *272*, 484
Tertiary recrystallization 455
Tetragonal deformation *185*
Tetrahedral position 29
Texture, deformation 248, *464*
–, fiber 466
–, recrystallization *465*
Thermal acceptor *637*, *643*
Thermal conductivity *290*
Thermal dilatation 10
Thermal equilibrium of point defects *36*
Thermal neutrons 337
Thermal resistivity 127
Thermal spike 43, 260, *272*, 484
Thermal stress 73, 277, *620*
Thermal vibrations 10, *16*
Thermo-elastic damping 394
Thin foil technique 133
Threshold energy for displacement *260*
Tilt boundary 326, *438*
Tracer 548
Transformation *476*
Transient cold-work internal friction *376*
Transient creep *230*
Transition temperature of flow stress *175*
– – of order *217*, 481
Translation plane 251
Transmission electronmicrography *133*, 332

Transported energy 418
Transport number *548*
Trap 307, 575
Trivacancy 35, 278
Tunnel effect 562
Twinning 140, 163, 246, 251, *468*, 475, 565, 613, *631*
Twinning mill 254
Twist boundary *438*

U

U-band, centre 563
Uphill diffusion 406

V

V-band, centre 535, *560*
Vacancy acceptor 598
Vacancy creep *236*, 368
Vacancy pair 34
Valence-type bond 13
Van der Waals attraction 12
Varley mechanism 270
Verhakung 144
Vibration frequency *16*, *38*
Vibration loss 359
Visco-elastic aftereffect 382
Viscous damping *357*
Void 35, 62

W

Weiss domain 344
Whisker growth 506
Whisker strength 506, 618
Widmanstätten structure 452, 481
Wigner effect 310
Work-hardening 141, *148*, *160*, 250, 310, 346, 568, 617, 629
Work-softening *177*

X

X-ray diffraction 128, *322*

Y

Yielding 163, *182*, 196, 211, 245, 249, 617, 629
Yield stress 64, 195, 217

Z

Z-band, centre 562
Zero-point energy 10
Zones *479*