THE BOOK WAS DRENCHED

UNIVERSAL
LIBRARY

OU_168153

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 669/Pr96 Accession No. 32602

Author Chalmers, Bruce ed.

Title Progress in metal physics v.4

This book should be returned on or before the date last marked below.

PROGRESS SERIES

METAL PHYSICS

4

PROGRESS IN METAL PHYSICS

4

Editor

BRUCE CHALMERS, D.Sc., Ph.D.

LONDON

PERGAMON PRESS LTD

1953

Published in Great Britain by Pergamon Press, Ltd., 2, 3 & 5 Studio Place, London, S.W.1
Printed in Great Britain by The Pitman Press, Bath, Somerset

CONTENTS

		PAGE
FOREWORD TO VOLUME I		vii
ACKNOWLEDGEMENTS		viii
1 INTERNAL FRICTION IN METALS		1
	A. S. Nowick	
	Hammond Metallurgical Laboratory, Yale University, New Haven, Connecticut, U.S.A.	
2 THE MECHANISM OF OXIDATION OF METALS AND ALLOYS AT HIGH TEMPERATURES		71
	Karl Hauffe	
	Institut für physikalische Chemie der Humboldt Universität, Berlin NW 7, Bunsenstrasse 1	
3 GASES IN METALS		105
	C. R. Cupp	
	Department of Metallurgical Engineering, University of Toronto, Toronto 5, Ontario, Canada	
4 THE THEORY OF SINTERING		174
	G. A. Geach	
	Associated Electrical Industries Ltd., Research Laboratory, Aldermaston Court, Aldermaston, Berks.	
5 THEORY OF DISLOCATIONS		205
	A. H. Cottrell	
	Department of Physical Metallurgy, The University, Edgbaston, Birmingham 15	
6 DIFFUSION IN METALS		265
	A. Le Claire	
	Atomic Energy Research Establishment, Harwell, Didcot, Berks	
7 NUCLEATION		333
	J. H. Hollomon and D. Turnbull	
	Research Laboratory, General Electric Co., The Knolls, Schenectady, N.Y. U.S.A.	
AUTHOR INDEX		389
SUBJECT INDEX		397

FOREWORD TO VOLUME 1

THE study of the physical properties of metals has developed through a number of stages. The first was that in which the mechanical properties were correlated empirically with the heat treatment to which the metal had been subjected and, sometimes, to the chemical composition. At this stage the successful treatment of metals was an art, in the sense that experience rather than understanding led to the most satisfactory results. The next stage, in which the internal structure of the metal was examined, was based originally on the use of the microscope and it was found that many experimental facts could be explained in terms of effects that were of the right size to be seen under magnifications of less than about two thousand. The development of the x-ray diffraction techniques allowed phenomena of a smaller order of magnitude to be examined, and much of the existing information was found to be comprehensible in terms of the geometry of the crystal structure of the various phases that were visible under the microscope.

More recent development can perhaps best be discussed by a division of the field into what may be termed 'statics' and 'dynamics'. Under the former heading is the study of the conditions which govern the structure of a metal or alloy when it is in thermodynamic and mechanical equilibrium. The theories of the phases that are present in equilibrium and of the elastic constants have made remarkable progress in terms of rapidly developing theory of the part played by electrons in the metal.

Under the heading 'dynamic' effects we may include both the conditions governing the approach to equilibrium in respect of the phases that are present, in which diffusion plays an important part, and the response of a metal to forces which are sufficient to cause non-recoverable or plastic mechanical deformation.

These and associated subjects have advanced so rapidly that it has become difficult for research workers in one part of the field to remain up to date in other branches. It is the purpose of this volume, which is the first of an annual series, to present authoritative reviews of the present state of knowledge in specialized aspects of the field that includes both physical metallurgy and metal physics. It is not intended that any one volume should form a complete textbook on these subjects. It is hoped rather that a few subjects of current interest should be discussed rather fully so as to cover, in the course of several years, all the more important aspects in which progress is being made. In order to make the series reasonably self contained it is proposed that the necessary 'historical' background should be included the first time a particular subject is discussed. Subsequent articles on such subjects will generally only cover the more recent progress.

B. CHALMERS

ACKNOWLEDGEMENTS

THE authors and publishers wish to express their thanks to those scientists and authorities who have given permission for the use of copyright material, and particularly to the American Chemical Society for the quotation on p. 141 and Fig. 12 in Dr. Cupp's contribution; the American Institute of Mining and Metallurgical Engineers for the quotations on pp. 122, 152 and 159, and for Figs. 7, 13, 14, 15, 16, 17, 18, 19 and 20, in Dr. Cupp's contribution; the American Institute of Physics for the quotation on p. 134 and Fig. 9 in Dr. Cupp's contribution; the American Society of Metals for quotations on pp. 107, 122 and 156, and for Fig. 3 in Dr. Cupp's contribution; *Comptes Rendus* for Fig. 8 in Dr. Cupp's contribution; the Electrochemical Society for the quotation on p. 128 and Figs. 4 and 5 in Dr. Cupp's contribution; the Institute of Metals for the quotations on pp. 157, 161, 162 and 164, and for Fig. 21 in Dr. Cupp's contribution and Fig. 25 in Professor Cottrell's contribution; the Iron and Steel Institute for the quotation on p. 165; *Metal Abstracts* for the quotation on p. 116; Messrs. Philips Electrical Ltd. for the quotations on pp. 114, 145 and 155; the *Philosophical Magazine* for Figs. 4 and 5 in Prof. Cottrell's contribution; *The Physical Review* for the quotation on p. 128 and Figs. 1 and 2 in Dr. Cupp's contribution and Fig. 22 in Prof. Cottrell's contribution; the Physical Society for the quotations on pp. 130, 139, 150 and 151, and for Figs. 6, 10 and 11 in Dr. Cupp's contribution; The Royal Society for the quotation on p. 127; the Royal Society of Arts for the quotation on p. 163; the University of Chicago Press for the quotations on pp. 134, 136 and 137; John Wiley Inc. for the quotations on p. 107. While every effort has been made to obtain permission in every case where copyright material has been used, it is hoped that any inadvertent omission from this list will be excused by those concerned.

1

INTERNAL FRICTION IN METALS

A. S. Nowick

THE capacity of a vibrating solid to convert its mechanical energy of vibration into heat, even when so well isolated that energy losses to its surroundings are negligible, is called internal friction. The most common manifestation of internal friction is the damping, or loss of vibration amplitude, of a freely vibrating body. Interest in internal friction in metals undoubtedly extends back to prehistoric times, to the earliest cymbals and bells. Modern investigation has followed two different channels. On the one hand there has been widespread interest in damping for engineering applications; depending on the particular application, one may wish to obtain either a very high or a very low rate of damping. In work of this kind, the effect of heat treatments and alloying elements on the damping characteristics of engineering materials is investigated and commonly reported in terms of a quantity called "damping capacity." Vibration amplitudes employed are commensurate with those that occur in practice. An adequate review of this phase of the subject would in itself require a sizeable article. A second approach to internal friction, with which the present review is concerned, is the use of damping as a tool to study internal structure and atomic movements in solids. This is a relatively new field of investigation and has been developing rapidly. Interest is focused on the origin and mechanism of internal friction rather than its desirability, and the vibration amplitudes employed are very small.

A perfectly elastic material will not produce damping since, under oscillatory conditions, stress and strain are always in phase, and consequently no mechanical hysteresis or loss of vibrational energy can take place. Damping is, therefore, a result of the non-elastic behaviour of materials. The fact that it can be observed at stress levels far below those at which plastic flow occurs, shows that there is really no such thing as an "elastic range." In recent years there has been a great increase in understanding of those physical processes that result in non-elastic behaviour, and, therefore, in damping, of metals at low stress levels. Review articles on this subject, written in the period 1939–1943,[1–3] emphasize three main sources of damping: thermoelasticity, ferromagnetic effects, and the internal stresses associated with cold work, or dislocations. Of these, the last was least understood, in the sense that the mechanism by which these internal distortions act to produce damping was not established. During the war and early

post-war period intensive investigations of internal friction, particularly by ZENER and co-workers, brought to light other sources of internal friction, including stress relaxation across grain boundaries and stress-induced ordering. These last two mechanisms together with thermo-elastic internal friction fitted into a scheme of non-elastic behaviour which ZENER[4] called *anelasticity*, the development of which culminated in his book[5] in 1948. In the treatment of anelasticity it is shown that not only internal friction but several related static effects (recoverable creep, stress relaxation, and elastic after-effect at low stress levels) result from this type of non-elastic behaviour. Only internal friction that does not depend on the amplitude of vibration is considered in this treatment. The impact of this important work has been so great as to lead to the widespread misconception that all internal friction at low amplitudes can be explained in terms of anelasticity. The internal friction originating in cold work is an appropriate example; theories attempting to explain observed effects in terms of the scheme of anelastic phenomena,[6–8] have not met with notable success. In fact, after READ[9, 10] showed that cold work internal friction in single crystals is strongly amplitude dependent, the writer demonstrated[11, 12] that this behaviour originates in a mechanism of static hysteresis, which is quite different from anelasticity.

In the present article, an attempt is first made to provide a phenomenological description of the general features of non-elastic deformation which is sufficiently general to include, as special cases, anelasticity as well as static hysteresis and amplitude dependent internal friction. The first three main sections deal with this phenomenological approach. It is shown that internal friction cannot be divorced from an equally important effect: the change in the effective elastic modulus resulting from the non-elastic behaviour, which is called "the ΔM-effect". Following the phenomenological treatment, the various experimental methods used to measure internal friction are reviewed. The remainder of the article is devoted to the physical origins of internal friction. All sources of damping involve an internal re-arrangement that takes place under stress. There are three types of re-arrangements that have been studied: *thermal* (the thermo-elastic effect), *magnetic*, and *atomic* re-arrangements. All of these are considered in the present article, although emphasis is placed upon the latter category, since most recent developments are in this field and it is the least adequately reviewed. Included in the category of atomic re-arrangements are diffusion effects, relaxation across interfaces, and the motion of dislocations.

PHENOMENOLOGICAL THEORY: HOMOGENEOUS STRAIN

In most vibration experiments the specimen under investigation is subjected to a non-homogeneous strain. A discussion of the elastic and

non-elastic behaviour of the specimen is, therefore, complicated by the dependence of stress and strain on position in the specimen. In this section, the principles of non-elastic behaviour will be introduced under the assumption that we are dealing either with a small element of a specimen within which the strain is uniform, or with a complete specimen throughout which the strain is the same at any instant. The latter condition can be realized, for a specimen of appropriate shape, by the introduction of an auxiliary inertia member relative to which the inertia of the specimen itself is negligible. For example, it is possible to set up a homogeneous shear strain by torsional vibration of a hollow tube to which is attached a heavy inertia member.

For simplicity of notation, it is assumed that the stress system within the specimen can be defined by a single component σ. Furthermore, only the principal strain ε, corresponding to the stress σ, need be considered, since other strain components do not enter into expressions for the work done by the stress. The symbol M is used to represent the appropriate elastic modulus that relates stress and strain.

As already mentioned, in order for internal friction to occur, the oscillatory stress and strain must not be in phase with each other and, in fact, in order that energy be dissipated, the strain must lag behind the stress. Let us represent this angle of lag by ϕ. Then ϕ is a convenient measure of internal friction, since the energy dissipated in a cycle of vibration goes to zero when ϕ approaches zero. It is convenient to use complex notation to represent the various phase relationships. Thus, the variation of stress with time is expressed as

$$\sigma = \sigma_1 e^{i\omega t} \qquad \ldots . \quad (1)$$

where ω is the angular frequency of vibration ($\omega = 2\pi f$, where f is the number of vibrations per second) and σ_1 is the stress amplitude. Correspondingly, the strain is

$$\varepsilon = (\varepsilon_1 - i\varepsilon_2)e^{i\omega t} \qquad \ldots . \quad (2)$$

where the quantities ε_1 and ε_2 are the amplitudes of the components of strain in phase with and 90° behind the stress, respectively. Clearly,

$$\tan \phi = \varepsilon_2/\varepsilon_1 \qquad \ldots . \quad (3)$$

Throughout this article it will be assumed that the internal friction is small, i.e. $\varepsilon_2 \ll \varepsilon_1$ and $\tan \phi \cong \phi$. This approximation greatly simplifies the relations between various measures of internal friction, and there have been very few experiments to which it does not apply. An appropriate definition for the *dynamic modulus*, M, is

$$M = \sigma_1/\varepsilon_1 \qquad \ldots . \quad (4)$$

It is also convenient to define the *complex modulus*, $\tilde{M}$, as the ratio of

stress to strain, using the complex notation of equations (1) and (2). Thus,

$$\tilde{M} = \frac{\sigma_1}{\varepsilon_1 - i\varepsilon_2} = \frac{\sigma_1}{\varepsilon_1(1 - i \tan \phi)} \cong M(1 + i\phi) \quad . \quad . \quad . \quad . \quad (5)$$

where the approximation of small ϕ has been introduced. The real part of the complex modulus is M and the ratio of the imaginary part to the real part is the phase angle ϕ, which is our present measure of internal friction. The complex modulus, therefore, contains the essential parameters pertaining to the vibration.

The occurrence of non-elastic behaviour means more than the fact that $\phi \neq 0$. In order to discuss the concept of non-elastic behaviour we must consider the law that describes the interrelation of stress and strain, not only under conditions of oscillation, but also under static conditions, e.g. the application of a constant stress. A *perfect lyelastic substance* is one for which the strain is uniquely determined by the stress at any instant, i.e. a monotonic functional relationship exists between the two quantities. This function only deviates from Hooke's law of proportionality when the strain is large. Since we are concerned entirely with small deformations, a simple law of proportionality may always be used to represent the stress-strain equation for a perfectly elastic substance. A material that shows *non-elastic behaviour* is one for which the stress-strain equation is not a simple proportionality but assumes a generalized form, in which, for example, there may be terms involving time derivatives of stress and strain or terms that represent multiple-valued functional relationships. For self consistency with the requirement of small damping, we may limit ourselves to stress-strain equations where all these non-elastic terms are very much smaller than the elastic terms. If the stress-strain equation is linear in stress, strain, and time derivatives of these quantities, it can be solved by the substitution of equations (1) and (2) to obtain the complex modulus directly. If, as a result of the non-elastic behaviour, terms which are nonlinear appear in the stress-strain equation, an exact solution can no longer be obtained by the substitution of a simple harmonic stress and strain, but requires complete Fourier series for these quantities, i.e. stress and strain may be *periodic* but not sinusoidal. An approximate solution can be obtained[13] if we ignore everything but the fundamental Fourier components of stress and strain and determine the complex modulus from these quantities. The parameters M and ϕ obtained from the real and imaginary parts of this complex modulus will depend on the amplitude of vibration, as a consequence of the nonlinear terms. We then speak of "amplitude dependent internal friction" and, in general, also of an "amplitude dependent dynamic modulus." This method, which converts a nonlinear problem into an approximately equivalent linear

problem with parameters dependent on the amplitude of vibration, is known as the *method of equivalent linearization*.[13] The approximation is a good one when the nonlinear terms are small, a condition which is certainly valid in the present case, since all nonlinear terms are among the small non-elastic terms.

We will now consider two well-known theories of non-elastic behaviour which lead to very simple stress-strain equations.

The Voigt Solid

In order to account for non-elastic behaviour of materials, VOIGT[14] made the assumption (previously introduced by STOKES[15] and others) that, in addition to the elastic stress $M\varepsilon$ there is a frictional resistance proportional to the rate of strain $\dot{\varepsilon}$. We, therefore, obtain the stress-strain equation

$$\sigma = M\varepsilon + \eta\dot{\varepsilon} \qquad \ldots . \quad (6)$$

This is a linear equation and, therefore, leads to an amplitude independent internal friction. A material that can be described in terms of equation (6) is called a Voigt solid. The complex modulus is $\tilde{M} = M + i\omega\eta$, so that the internal friction is given by

$$\phi = (\eta/M)\omega \qquad \ldots . \quad (7)$$

The internal friction of a Voigt solid is, therefore, proportional to the frequency of vibration and its dynamic modulus is a constant (independent of frequency).

The Voigt assumptions are most often used in textbooks on theoretical physics,[16] to discuss the damped oscillations of a mass point attached to a weightless spring. The purpose of such treatments is usually to show how internal friction manifests itself in free and forced vibrations. No attention is usually paid to the question of the frequency dependence of internal friction.

The Maxwell Solid

An alternative description of non-elastic behaviour is due to MAXWELL.[17] Rather than use the nebulous concept of an internal frictional force, this approach first assumes that the strain is made up of two parts:

$$\varepsilon = \varepsilon' + \varepsilon'' \qquad \ldots . \quad (8)$$

where ε' is the perfectly elastic strain which obeys Hooke's law

$$\varepsilon' = \sigma/M' \qquad \ldots . \quad (9)$$

and ε'' is the *non-elastic strain* which results from internal re-arrangement and is not proportional to the stress. Equation (9) is to be regarded as the definition of M'. MAXWELL further assumes that the non-elastic

term $\dot{\varepsilon}''$ obeys the Newtonian viscosity equation, i.e. the flow $\dot{\varepsilon}''$ is proportional to the stress, or

$$\dot{\varepsilon}'' = \xi\sigma \qquad \ldots . \quad (10)$$

The stress-strain equation is obtained by combining equations (8), (9) and (10):

$$\dot{\varepsilon} = \frac{\dot{\sigma}}{M'} + \xi\sigma \qquad . \quad (11)$$

From this equation and the assumption of small internal friction, we obtain

$$M = M'; \quad \phi = \xi M'/\omega \qquad \ldots . \quad (12)$$

Again the dynamic modulus is a constant, but now ϕ is inversely proportional to the frequency. The amplitude independence of ϕ is related to the linearity of equation (11).

It has been known for many years that the experimentally observed frequency dependence of internal friction, for most materials, is in agreement with neither the VOIGT nor the MAXWELL descriptions of non-elastic behaviour. Another shortcoming of these descriptions is the prediction, in both cases, that the dynamic modulus of the system is unaffected by the non-elastic behaviour, except for terms of second order in the internal friction which are omitted here. Experimental observations show that the occurrence of internal friction is often accompanied by a fractional decrease in the dynamic modulus, of the same order of magnitude as the internal friction.

Various methods have been suggested for the generalization of the Voigt and Maxwell solids to obtain a more widely applicable theory. These methods regard the properties of an actual solid as a superposition of those of a large collection (finite or infinite) of Voigt or of Maxwell solids each with different parameters. In this way it is possible to introduce into the description of the non-elastic behaviour of a material as many arbitrary parameters as are necessary to describe its observed behaviour. Such superposition methods have been particularly popular in the study of mechanical properties of high polymers.[18, 19] They are limited, however, to the treatment of materials whose internal friction is independent of the amplitude of vibration, and also, cannot be used for the study of static hysteresis (described in a later section). It will, therefore, be more suitable for us to introduce a different approach, which is sufficiently general to include all phenomena that lead to internal friction in metals.

The Generalized Maxwell Approach

The first assumption of the Maxwell description is that a material under stress may undergo internal re-arrangement leading to a non-

elastic strain ε'' in addition to the elastic strain ε'. This assumption, expressed by equations (8) and (9), is very generally applicable to non-elastic phenomena. On the other hand, we will drop the highly restrictive second assumption expressed by equation (10) and, for the present, leave unspecified the equation that describes the non-elastic strain ε''. Without such an equation, we cannot obtain the stress-strain relation required to complete the description of the behaviour of a material. Possible equations for ε'' will be discussed in a later section; at present, phase relationships will be discussed in general terms.

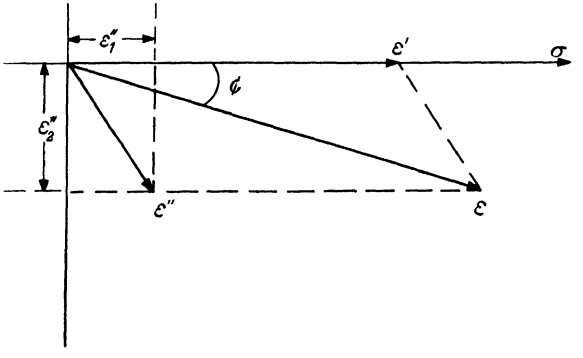


Fig. 1. Phase relationships between the stress and the elastic and non-elastic strain

If the stress (or its fundamental Fourier component) is represented by equation 1, the elastic strain is,

$$\varepsilon' = \varepsilon_1' e^{i\omega t} \quad . \quad . \quad . \quad . \quad (13)$$

The non-elastic strain ε'' will, in general, be resolvable into a component in phase with the stress and one lagging by 90°, or

$$\varepsilon'' = (\varepsilon_1'' - i\varepsilon_2'')e^{i\omega t} \quad . \quad . \quad . \quad . \quad (14)$$

where the quantities ε_1', ε_1'' and ε_2'' are appropriate amplitudes and are real and positive quantities. Fig. 1 shows the phase relationships between stress and components of strain. This diagram is not drawn to scale, since the assumption of small internal friction and non-elastic effects means that $|\varepsilon''| \ll |\varepsilon'|$. With this approximation $\varepsilon_1 \cong \varepsilon_1'$ (to first order), and equation (3) becomes

$$\phi \cong \varepsilon_2''/\varepsilon_1' \quad . \quad . \quad . \quad . \quad (15)$$

The internal friction is, to the first approximation, dependent only on the component of ε'' out of phase with stress. The result is not surprising since only the component of strain out of phase with the stress will

contribute to energy dissipation. The dynamic modulus M is given by (see equation (4))

$$M = \frac{\sigma_1}{\varepsilon_1' + \varepsilon_1''} \cong M'(1 - \varepsilon_1''/\varepsilon_1') \quad . \quad . \quad . \quad . \quad (16)$$

where M' is defined by equation (9). Thus, since the presence of non-elastic behaviour may increase the in-phase component of strain, the dynamic modulus is less than (or equal to) M', and the difference, $\Delta M = M' - M$ is given by

$$\Delta M/M = \varepsilon_1''/\varepsilon_1' \quad . \quad . \quad . \quad . \quad (17)$$

This lowering of the elastic modulus because of the existence of a non-elastic strain will be called the ΔM *effect*. This discussion has, therefore, shown that internal friction and the ΔM effect are closely related consequences of the occurrence of non-elastic strain, the first resulting from the component out of phase with stress, and the second from the component in phase with the stress. It is also clear that if the equation, thus far unspecified, that describes ε'' is nonlinear, the ratios $\varepsilon_2''/\varepsilon_1'$ and $\varepsilon_1''/\varepsilon_1'$ will be functions of strain amplitude. Specific forms for this equation will be considered in the section called "Viscoelasticity and Static Hysteresis."

Measures of Internal Friction

The relationship between various measures of internal friction, as well as of the ΔM effect, will now be considered. As long as the discussion of non-elastic behaviour is restricted to the use of quantities such as the complex modulus and phase angle ϕ, it is possible to treat in the same way, either a small element of material or a complete specimen under homogeneous strain. If quantities such as the logarithmic decrement and the width of a resonance peak are to be considered, it is necessary to deal with an entire specimen in a particular mode of vibration. For simplicity, we will discuss here only specimens vibrating under homogeneous strain, and in the next section will show how the present discussion is generalized when the strain is not homogeneous. It is required, then, that a relatively large inertia member be attached to the specimen under consideration. If θ represents the displacement of the inertia member, I its "inertia," F_e and F_a the "force" exerted upon this member by the specimen and the externally applied "force," respectively, then the equation of motion of this member is

$$I\ddot{\theta} = F_e + F_a \quad . \quad . \quad . \quad . \quad (18)$$

(In the example of the torsion of a hollow tube mentioned earlier, θ is simply the angle of twist, I the moment of inertia, and the quantities

F are torques. On the other hand, if the vibration is under a longitudinal load, θ becomes the longitudinal displacement, I the attached mass, and F a force.) The quantity θ is essentially the strain, ε, throughout the specimen, except for a geometric constant of proportionality, while F_e is proportional to the stress σ, but opposite in sign. Inasmuch as $\sigma = \tilde{M}\varepsilon$ (by the definition of $\tilde{M}$, equation (5)), with $\tilde{M} = M(1 + i\phi)$, we may also write

$$F_e = -\tilde{k}\theta; \quad \tilde{k} = k(1 + i\phi) \qquad \ldots . \quad (19)$$

where $\tilde{k}$ is a complex "force constant." Equation (18) may then be solved in the case of *forced vibrations* with an impressed angular frequency ω, i.e. when the applied "force" is taken as sinusoidal of the form $F_a = F_0 e^{i\omega t}$. Under steady state conditions, the equation will have a solution of the form $\theta = \theta_0 e^{i\omega t}$, where θ_0 is, in general, complex. (If the stress-strain equation is nonlinear, θ will actually be representable by a complete Fourier series, but as mentioned earlier, we will ignore all but the fundamental Fourier components.) When this solution and equation (19) are substituted into equation (18), an equation for θ_0 is obtained:

$$\theta_0 = \frac{F_0/I}{(\omega_r^2 - \omega^2) + i\omega_r^2\phi} \qquad \ldots . \quad (20)$$

where

$$\omega_r^2 = k/I \qquad \ldots . \quad (21)$$

The real quantity $|\theta_0|^2$ is given by

$$|\theta_0|^2 = \frac{(F_0/I)^2}{(\omega_r^2 - \omega^2)^2 + \omega_r^4\phi^2} \qquad \ldots . \quad (22)$$

This is the well-known equation for a resonance curve (a plot of $|\theta_0|^2$ versus ω) with a resonant angular frequency (at which the amplitude of vibration is a maximum) equal to ω_r. If the frequencies at which $|\theta_0|^2$ falls to half the maximum value are denoted by ω_1 and ω_2 it follows that

$$\phi = \frac{\omega_2 - \omega_1}{\omega_r} = \frac{1}{Q} \qquad \ldots . \quad (23)$$

The internal friction ϕ may, therefore, be obtained in forced vibration as the ratio of the width of the resonance curve (square of displacement amplitude versus frequency) at half maximum divided by the resonant frequency. Inasmuch as the reciprocal of this ratio is called the "Q" of an electrical circuit, the symbol Q^{-1} has been adopted by some authors[1, 20] as a measure of internal friction; from equation (23), Q^{-1} is equal to the phase angle ϕ between stress and strain. The study of internal friction by means of a resonance curve is complicated when the

internal friction is amplitude dependent. When ϕ is itself a function of $|\theta_0|$, and in fact (as commonly observed experimentally) an increasing function of $|\theta_0|$, the height of the resonance peak will be determined by the value of ϕ at maximum amplitude, but the peak is broader than for the case of constant internal friction. Although the resonance peak is broadened by the amplitude dependence of ϕ, it may still be symmetrical about $\omega = \omega_r$ (to terms of first order in the internal friction) so long as there is no ΔM effect. In general, however, amplitude dependent internal friction is accompanied by an amplitude dependent dynamic modulus M. This means that k is a function of amplitude and, therefore, that ω_r as well as ϕ, in equation (22), is dependent on $|\theta_0|$. In the usual case, where the ΔM effect increases (and, therefore, ω_r decreases) with increasing amplitude, it is readily shown from equation (22) as well as experimentally[11] that the resonance curve falls off faster on the low frequency side than on the high frequency side. If the amplitude dependence of ω_r is sufficiently large, this asymmetry may become so great that the resonance curve bends over on itself and becomes double valued on the low frequency side. Thus, discontinuities may be observed in tracing this curve experimentally.[21]

When amplitude dependent effects are studied by forced vibration methods, it is usually preferable to work at exactly the resonant frequency where, from equation (22), the amplitude at resonance $|\theta_0{}^m|$ is

$$|\theta_0{}^m| = F_0/(I\omega_r{}^2\phi) \qquad \ldots\ldots (24)$$

It is necessary to determine the maximum amplitude $|\theta_0{}^m|$ and the force F_0 in order to obtain ϕ; both ϕ and ω_r are then known at the specific amplitude $|\theta_0{}^m|$. By variation of F_0 it is then possible to obtain these values at any other displacement amplitude. Whereas, in principle this method is a very simple one for determination of the internal friction and ΔM effect as a function of amplitude, the quantities F_0 and $|\theta_0{}^m|$ are often difficult to obtain experimentally. In the piezoelectric method (to be discussed in a subsequent section) these quantities are readily calculated, so that this method has been particularly suitable to the study of amplitude dependent internal friction. In other methods of forced vibration, it is more convenient to keep F_0 constant and determine Q^{-1} (equation (23)), which requires only a knowledge of *relative* amplitudes and does not require that F_0 be known. The latter methods are, therefore, only useful when internal friction is independent of amplitude.

In addition to the possibility for studying internal friction by the determination of the resonance curve in forced vibrations, the oldest and most popular methods are related to determination of the *damping of free vibrations*, where vibrations, once excited, are allowed to continue,

with no external forces applied to the system; the rate of decrease of amplitude is then determined. In order to solve the equation of motion 18, for $F_a = 0$ (no external forces), we substitute a solution of the form

$$\theta = \theta_0 e^{i\tilde{\omega}t}; \quad \tilde{\omega} = \omega(1 + i\delta/2\pi)$$

which represents exponentially damped oscillations, if δ is a constant. (θ_0 is now real.) When this solution is substituted into equation (18), and cognizance taken of equation (19), we obtain: $I\tilde{\omega}^2 = \tilde{k}$; if real and imaginary parts are now separately equated, and second order terms in internal friction neglected, it is found that

$$\omega^2 = k/I; \quad \delta = \pi\phi \qquad \ldots . \quad (25)$$

The quantity δ represents the natural logarithm of the ratio of amplitudes in two successive vibrations and is called the *logarithmic decrement* (or simply, the "decrement"). When this quantity is small (as always assumed in this discussion) it is also equal to the fractional decrease in vibrational amplitude per cycle. In order to obtain δ from experiment, it is simply necessary to plot the logarithm of the amplitude of vibration against the number of cycles of vibration. If internal friction is independent of amplitude a straight line is obtained whose slope gives the decrement directly. However, if internal friction is amplitude dependent, this curve will not be a straight line, but its slope at any amplitude will give the decrement at that particular amplitude. Thus an entire curve of decrement versus amplitude of vibration may be obtained from a single decay curve if the precision of the data is sufficiently good.

Equations (21) and (25) also show that the natural frequency in free vibration and the resonant frequency ω_r in forced vibration are the same. This resonant frequency is affected by the existence of a ΔM effect; under conditions where non-elastic behaviour does not occur, its value would be $\omega' = (k'/I)^{1/2}$, where k' is the force constant corresponding to the modulus M'. Because of the ΔM effect, there will be a decrease in resonant frequency and, if we define $\Delta\omega = \omega' - \omega_r$, we find that

$$\Delta\omega/\omega_r = \tfrac{1}{2}\Delta M/M \qquad \ldots . \quad (26)$$

Thus, the existence of a ΔM effect is synonymous with a decrease in the resonant frequency or in the frequency of free vibration. Also, as already noted, amplitude dependent internal friction is usually accompanied by an amplitude dependent dynamic modulus. This effect is detected by a change of resonant frequency, or of the frequency of free vibration, with a change in the amplitude of vibration.

Another measure of internal friction is the fractional decrease in vibrational energy per cycle. Inasmuch as δ is the fractional decrease

in amplitude per cycle, and the vibrational energy is proportional to the square of the amplitude, it follows that

$$\delta = \Delta W/2W$$

where W is the vibrational energy and ΔW is the loss of vibrational energy in one cycle. A closely related quantity which is widely used in engineering applications is the *damping capacity* which is the per cent decrease in vibrational energy per cycle, so that: damping capacity $= 200\delta$.

The relation between the various measures of internal friction are summarized as follows:

$$\phi = \delta/\pi = \Delta W/2\pi W = Q^{-1} \quad . \ . \ . \ . \quad (27)$$

The last equality applies only when internal friction is amplitude independent. These simple relationships between the various measures of internal friction are derived under the assumption of small damping and break down when this assumption is no longer valid. For large damping, the relations between different measures of internal friction depend on the mechanism of the damping and must be derived separately for each case.[22]

The measured values of ϕ and $\Delta M/M$ will not only depend upon the frequency and possibly the amplitude of vibration, but will also depend on the type of stress employed. The relation between values obtained in shear and in longitudinal vibration, for example, will depend on the mechanism of the non-elastic behaviour.

Phenomenological Theory: Non-homogeneous Strain*

In practice, damping is rarely, if ever, studied under conditions where the mass is all concentrated in an auxiliary inertia member, and the stress and strain are homogeneous throughout the "elastic" member. Very often, the material whose elastic and non-elastic properties are to be studied, will be in the form of a bar which is vibrated in a longitudinal, transverse, or torsional mode of oscillation. Under these circumstances the stress amplitude is not constant throughout the specimen. It will be shown in this section, however, that the interpretation of vibration experiments on these more realistic specimens is completely analogous to the case of homogeneous strain, except that special attention is required when the internal friction is amplitude dependent.

It is well known[23] that a perfectly elastic system will vibrate in certain normal modes of vibration characteristic of the system. Once it is vibrating in one normal mode it will continue to do so indefinitely.

* The reader who is not interested in the details of this section may omit the section without loss of continuity.

If small non-elastic effects are introduced, it is to be expected that the system will still vibrate in the same normal mode, except that now it will diminish in amplitude due to energy losses. Such behaviour will actually occur when the stress-strain equation is linear; in the case where nonlinear terms occur, the solution of the equation of motion is no longer a single normal mode of vibration. It is useful, however, to consider only the components of the vibration in a single normal mode, just as we considered only fundamental Fourier components in the case of homogeneous strain when a nonlinear stress-strain equation occurred. This extension of the method of equivalent linearization to vibration under non-homogeneous strain has been justified experimentally[11] over a range of amplitudes in which the internal friction reaches twice its value at low amplitudes.

In most of the experimental methods a simple stress system is involved. We may, therefore, continue to describe the vibration in terms of a single component of stress σ and the corresponding principal strain ε. When the system vibrates in one particular normal mode, the strain ε may be written in the form

$$\varepsilon = C(t)G(x, y, z) \quad . \quad . \quad . \quad . \quad (28)$$

where the space-dependent function G is characteristic of the mode of vibration (an eigenfunction of the vibration equation). By means of this separation of variables, the problem of vibration in a normal mode is reduced to a one-dimensional problem similar to that of homogeneous strain, since only the co-ordinate C, which specifies the extent of excitation, is needed in order that the strain at every point in the specimen be known. A system vibrating very nearly in a normal mode, under small damping, may, therefore, be represented by an equivalent simple vibrator, whose equation of motion is of the form 18, where the quantity C is now the displacement. For a perfectly elastic system, $C = C_0 e^{i\omega t}$; on the other hand, if the specimen displays non-elastic behaviour, the quantity C will show damping in free vibration or a resonance curve in forced vibration. It is necessary to relate the damping ϕ_s of the equivalent vibrator, which represents the behaviour of the specimen as a whole and the dynamic modulus M_s, obtained from the resonant frequency of the specimen, to the corresponding properties ϕ and M of an element of specimen *material*. If the stress-strain equation is linear and the quantities M and ϕ, therefore, independent of vibration amplitude, we may anticipate that $\phi_s = \phi$ and $M_s = M$. If, on the other hand, M and ϕ are functions of the strain amplitude ε_1, within a small element of the specimen, the quantities ϕ_s and M_s will depend upon the amplitude of vibration C_1 of the specimen.

To obtain the relations between the quantities that represent the behaviour of the specimen as a whole and those that represent the

properties of the specimen material, it is convenient to employ the notation of the generalized Maxwell approach introduced in the last section. The quantity C may be divided, in the usual way, into C' and C'', so that the parameters characteristic of the specimen are:

$$\phi_s = C_2''/C_1'; \quad \Delta M_s/M_s = C_1''/C_1' \qquad . \; . \; . \; . \quad (29)$$

where $\Delta M_s = M' - M_s$ by analogy to the definition of ΔM. Now, consider an element of the specimen for which the total strain is ε and non-elastic strain ε''. The internal friction

$$\phi(\varepsilon_1) = \varepsilon_2''/\varepsilon_1' \qquad . \; . \; . \; . \quad (30)$$

is, in general, a function of the strain amplitude ε_1. If the specimen is assumed to be vibrating very nearly in a single normal mode, as in equation (28), we have

$$\varepsilon_1' = C_1'G; \quad \varepsilon_2'' = C_2''G$$

If these last two expressions are inserted into equation (30), we obtain

$$C_2''G = C_1'G \,.\, \phi(C_1G) \qquad . \; . \; . \; . \quad (31)$$

This equality cannot be strictly true unless ϕ is a constant (since dividing through G gives a space dependent function on the right and a constant on the left, if $\phi \neq$ constant), i.e. as already stated, it is not strictly correct to assume that only a single mode of vibration is involved when ϕ is not constant. Inasmuch as any function of co-ordinates of the specimen can be expanded in terms of a set of eigen-functions, we may take components of the two sides of equation (31) corresponding to the normal mode G. To do this, we multiply both sides by G and integrate over the volume of the specimen to obtain

$$\phi_s(C_1) = \frac{C_2''}{C_1'} = \frac{\int \phi(C_1G) \,.\, G^2 d\tau}{\int G^2 d\tau} \qquad . \; . \; . \; . \quad (32)$$

where $d\tau$ represents an element of volume of the specimen. Equation (32) relates the experimentally measured function $\phi_s(C_1)$ to the function $\phi(\varepsilon_1)$, which is a property only of the specimen material. The equation shows, as we have already anticipated, that only in the amplitude independent case is the internal friction measured on a specimen vibrating under non-homogeneous strain independent of the mode of vibration, i.e. dependent only on the properties of the specimen material. In exactly the same way, the result

$$\left(\frac{\Delta M_s}{M_s}\right) = \frac{\int (\Delta M/M) \,.\, G^2 d\tau}{\int G^2 d\tau} \qquad . \; . \; . \; . \quad (33)$$

is obtained.

In summary, it has been shown in this section that the only difference

in the analysis of the non-elastic behaviour of a specimen vibrating in a normal mode and one for which the strain is homogeneous throughout, is that in the case of non-uniform strain the measured values of internal friction and the dynamic modulus represent a weighted average (as given in equations (32) and (33)) of the corresponding properties of the specimen material, when these properties are amplitude dependent.

Viscoelasticity and Static Hysteresis

In the phenomenological treatment of internal friction given so far, Maxwell's first assumption (equations (8) and (9)) has been used, while his second assumption (equation (10)) has been eliminated on the grounds that it is too restrictive for a description of the non-elastic behaviour of most materials. It has been possible to introduce the ΔM effect and to show the relationship between the various measures of internal friction without substituting any other assumptions in place of equation (10). In order to obtain any specific information, such as the variation of internal friction with frequency or amplitude of vibration, it is obviously necessary to introduce expressions which give further information about the non-elastic strain.

The "Standard Linear Solid"

A very important type of non-elastic behaviour is shown by materials that have been called[5] "standard linear solids." Such materials may also be said to show a "single time of relaxation." The basis for the behaviour to be described is the existence of an internal parameter p which takes on a different equilibrium value $\bar{p}$ at each value of the stress and whose variation is accompanied by a non-elastic strain ε''. The following three conditions are assumed to apply: (1) the non-elastic strain is proportional to p

$$\varepsilon'' = \lambda p \qquad \ldots . \quad (34)$$

(2) The equilibrium value of p, denoted by $\bar{p}$, is proportional to the stress. (3) The parameter p approaches its equilibrium value $\bar{p}$, under constant stress, according to the relaxation equation,

$$\dot{p} = -(p - \bar{p})/\tau \qquad \ldots . \quad (35)$$

so that the difference $p - \bar{p}$ goes to zero exponentially, with relaxation time τ. From conditions (1) and (2), the equilibrium non-elastic strain $\bar{\varepsilon}''$ $(= \lambda\bar{p})$ is proportional to stress, and, therefore, to the elastic strain ε', or

$$\bar{\varepsilon}'' = \Delta_M \varepsilon' = \Delta_M \sigma / M' \qquad \ldots . \quad (36)$$

where the dimensionless proportionality constant Δ_M is called the *relaxation strength* and must be much less than unity if our assumption

of small damping is to be valid. When equations (34), (35) and (36) are combined they give, as an equation for ε'',

$$\dot{\varepsilon}'' = -(1/\tau)(\varepsilon'' - \Delta_M \varepsilon') \quad \ldots . \quad (37)$$

This result replaces equation (10) in the Maxwell approach and is essentially the same as that equation except for the presence of an extra term on the right-hand side. When a harmonic solution is substituted into equation (37), there results

$$\varepsilon''/\varepsilon' = \Delta_M/(1 + i\omega\tau) \quad \ldots . \quad (38)$$

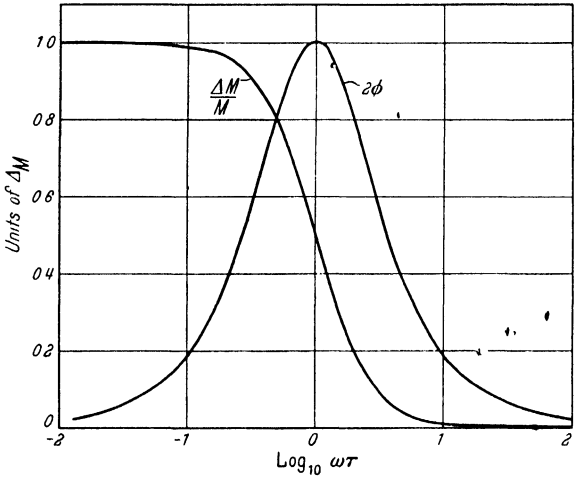


Fig. 2. Frequency dependence of internal friction and ΔM effect for a standard linear solid. The functions are plotted for unit relaxation strength, Δ_M

When equations (13) and (14) are substituted for ε'' and ε' and the real and imaginary parts of equation (38) separately equated, the internal friction and ΔM effect are obtained directly:

$$\phi = \frac{\varepsilon_2''}{\varepsilon_1'} = \Delta_M \frac{\omega\tau}{1 + \omega^2\tau^2} \quad \ldots . \quad (39)$$

$$\frac{\Delta M}{M} = \frac{\varepsilon_1''}{\varepsilon_1'} = \frac{\Delta_M}{1 + \omega^2\tau^2} \quad \ldots . \quad (40)$$

or, for the dynamic modulus itself,

$$M = M'\left(1 - \frac{\Delta_M}{1 + \omega^2\tau^2}\right) \quad \ldots . \quad (41)$$

Because of the assumptions of linear relationships (equations (34) and (36)) the internal friction and ΔM effect turn out to be independent of

vibration amplitude. The internal friction curve is a symmetrical peak when ϕ is plotted against log ($\omega\tau$), as shown in Fig. 2, with a maximum value equal to $\Delta_M/2$ at $\omega\tau = 1$. The interpretation for the occurrence of such a peak is as follows: the internal parameter p will constantly strive to attain the appropriate equilibrium value $\bar{p}$ at every value of stress, i.e. the relaxation process (equation (35)) is constantly going on. When the period of vibration is very short compared to τ, or $\omega\tau \gg 1$, practically no relaxation is possible during a cycle because the stress is changing rapidly compared to the time for achievement of equilibrium. Under these conditions ε'' is practically equal to zero, the behaviour of the material is essentially perfectly elastic, and, therefore, $\Delta M = 0$, in agreement with equation (40) and Fig. 2. The dynamic modulus M that is measured at sufficiently high frequencies is, therefore, equal to M'; it is also known[5] as the *unrelaxed modulus*, inasmuch as it is measured under conditions where no relaxation occurs. Fig. 3 (*a*) shows the proportionality between stress and strain at high frequencies. The other extreme, $\omega\tau \ll 1$, corresponds to a period of vibration so much longer than τ that there is sufficient time for achievement of very nearly equilibrium values of the parameter p at every value of stress. Under these conditions the strain at every value of stress is $\varepsilon' + \bar{\varepsilon}''$, which is a unique function of stress. As illustrated in curve *b* of Fig. 3, the stress-strain relationship is again a straight line and, therefore, ϕ is essentially zero; the dynamic modulus M, which is the slope of this straight line is lower than M', and, in fact, is equal to $M'(1 - \Delta_M)$ (from equation (41)). This value is known[5] as the *relaxed modulus*, since it is measured under conditions where relaxation is always essentially complete. Between these two extreme cases, stress and strain will not be unique functions of one another, since relaxation cannot keep up with alternations in stress. The stress-strain curve is then an ellipse whose area attains a maximum value when $\omega\tau = 1$. The slope of the major axis of the ellipse is the dynamic modulus M and always falls

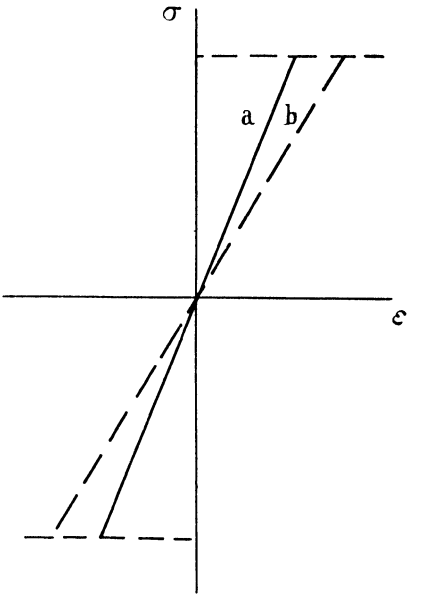


Fig. 3. Stress-strain curves in a standard linear solid: (*a*) at very high frequencies, (*b*) at very low frequencies. In both cases stress and strain are in phase and there is no damping

between the two extreme values of Fig. 3. The variation of the quantity $\Delta M/M$ with frequency is shown in Fig. 2, which is a plot of equation (40). It is interesting that only in the range where damping is high is there an appreciable dependence of the dynamic modulus on frequency. If measurements are made only at low frequencies, there is no way of distinguishing the behaviour of the specimen from that of a perfectly elastic substance whose modulus is $M'(1 - \Delta_M)$.

The standard linear solid subjected to a given type of vibration is completely described by three independent constants, M', Δ_M and τ, in contrast to the one constant, the elastic modulus, of a perfectly elastic body. The appropriate modulus depends on the type of experiment and in most cases will be either Young's modulus E or the rigidity modulus G. The relaxation strength Δ_M is a measure of the magnitude of the non-elastic effects, i.e. the height of the internal friction peak and the value of $\Delta M/M$ at very low frequencies. It is a quantity that should be small compared to unity in order that the assumptions made in deriving the relationship between different measures of internal friction be valid. Actual observed values of Δ_M have ranged from 10^{-4} to greater than 10^{-1}. The latter value may seem to violate our requirement of $\Delta_M \ll 1$, but the error in the application of the above theory to the case of $\Delta_M = 0{\cdot}1$ is not serious. The reason for designating the appropriate modulus as a subscript to Δ is that the relaxation strength is dependent upon the type of vibration experiment. For example, in many cases relaxation is produced only by shear stress. Thus, if a vibration experiment were conducted under hydrostatic stress only, we would have $\Delta_K = 0$ (K = bulk modulus). In a torsion experiment, Δ_G is involved and in tension or flexure Δ_E. However, since in tension only the shear components would produce relaxation, the relation between Δ_E and Δ_G is readily obtained[5] from elasticity theory as

$$\Delta_G = 3(G/E)\Delta_E \quad . \quad . \quad . \quad . \quad (42)$$

For the usual values of G/E, Δ_G is about 15 per cent greater than Δ_E.

The third independent constant of the standard linear solid is the relaxation time τ. This parameter is essentially the mean time for the appropriate internal re-arrangements to occur which will lead to the attainment of the equilibrium value of the internal parameter p. In the important case where a change in p occurs by atomic re-arrangement or diffusion, we have at our disposal a valuable method for obtaining the mean atomic jump time. This application of internal friction measurements will be discussed later. Whenever atomic movements of any kind are involved in the relaxation process it is to be expected that τ will be strongly temperature dependent, according to the Arrhenius equation,

$$\tau = \tau_0 e^{H/RT} \quad . \quad . \quad (43)$$

where H is a heat of activation, R is the gas constant, and T the absolute temperature. The existence of such a relationship can be used to great advantage when it is desired to obtain the internal friction peak experimentally. If this peak is to be obtained by varying the frequency it requires measurements over two or three powers of ten in ω (see Fig. 2). Such measurements usually are experimentally difficult. Inasmuch as ϕ is dependent only on the product $\omega\tau$, the same end can be achieved if ϕ is observed as a function of τ (by varying the temperature) at constant ω as by a study of ϕ as a function of ω at constant τ. A similar curve to that of Fig. 2 may, therefore, be obtained by plotting ϕ against T^{-1} when τ obeys equation (43). This curve will still have a maximum when $\omega\tau = 1$, except that τ is now the variable. Curves of ϕ against T^{-1} obtained at two different frequencies ω_1 and ω_2 can be superposed by a horizontal shift, and from the magnitude of this shift the heat of activation H can be calculated as follows: for the same value of ϕ, $\omega_1\tau_1 = \omega_2\tau_2$, where the subscripts 1 and 2 refer to the two curves. Therefore

$$\omega_1 e^{H/RT_1} = \omega_2 e^{H/RT_2}$$

or

$$\ln(\omega_2/\omega_1) = \frac{H}{R} \cdot \Delta(1/T) \qquad \ldots \quad (44)$$

so that from the experimentally observed shift, $\Delta(1/T)$, between the curves for frequencies ω_1 and ω_2, the quantity H is determined. Equation (44) is more generally valid than the case of a standard linear solid, since its derivation is dependent only on the fact that ϕ is a function of ω and τ through the product $\omega\tau$, and not on the form of that function.

Superposition of Relaxation Times

Although there are several physical mechanisms in metals and other solids that lead to non-elastic behaviour describable by equation (37), a material whose internal friction curve is given precisely by equation (39), over the *entire frequency range*, has never been found. On the one hand one may find more than one such peak, resulting from different mechanisms and located at different positions on the frequency scale. Thus, an *acoustic* or *relaxation spectrum*[5] is said to exist, somewhat in analogy to the optical spectrum. On the other hand, any single peak of the relaxation spectrum may deviate to a lesser or greater extent from the theoretical curve of equation (39). The reason for this deviation is related to fluctuations in the time of relaxation at different positions in the specimen. For example, if the relaxation process is controlled by atomic diffusion, the time of relaxation will be closely related to the mean atomic jump time. This latter quantity is capable of fluctuating

in accordance with fluctuations in the number of impurity atoms and in the state of internal stress in the immediate neighbourhood of the atoms under consideration. Usually, therefore, the term ε'' will be made up of a series of terms each obeying equation (37) but with slightly different values of τ. The observed internal friction and ΔM effect may then be a superposition of terms of forms given by equations (39) and (40), with different values of τ. The existence of a range of relaxation times, no matter how narrow, in place of a single relaxation time, will always result in a broadening of the internal friction peak. Equation (39), therefore, serves as a theoretical limit which is very closely approached in some cases (e.g. interstitial diffusion in b.c.c. lattices) while in other cases (such as grain-boundary internal friction) considerably wider peaks are observed. These examples will be discussed more fully in subsequent sections.

When the observed internal friction is the result of the superposition of mechanisms which have a range of relaxation times, but for which each value of τ obeys equation (43), there are two possibilities that must be considered: (1) the various values of τ differ from each other only in the values of τ_0; (2) different values of H are involved. The first case (probably illustrated by grain-boundary internal friction) occurs when the various values of τ are each different multiples of a unit jump time τ_1, which itself obeys an Arrhenius equation. Such a situation might arise if, in different parts of a material, relaxation is produced by the occurrence of a different *number* of atomic jumps, each jump requiring a mean time τ_1. In this case, the internal friction curve, no matter what its form, is still dependent on frequency and temperature through the product $\omega e^{H/RT}$. The arguments leading to equation (44), therefore, still apply and the value of H may be obtained from the horizontal shift of the curve of ϕ versus T^{-1} with a change of frequency. In the second case, in which H is not constant, a change of frequency produces a change in the shape of the internal friction curve, since its various components are shifted by different amounts. If the change of shape is not too great, an average value of H may be obtained from the average horizontal shift.

Anelasticity

The material that shows a single time of relaxation and one which is describable in terms of a superposition of relaxation times are the most important examples of that field of non-elastic behaviour that ZENER[4, 5] has called "anelasticity." In order that an imperfectly elastic solid be said to show anelasticity two conditions are required: (1) the stress-strain equation should be linear; (2) if a static stress is applied to such a material and then removed, no permanent set shall remain, i.e. any residual strain that exists after the stress is removed shall decay back

to zero if sufficient time is allowed. As an example, in the case of the standard linear solid, if equations (8) and (9) are combined with (37) it is possible to eliminate ε' and ε'' to obtain a linear equation in ε, $\dot{\varepsilon}$, σ, and $\dot{\sigma}$ (the stress-strain equation). (This equation may be generalized[24] by including higher derivatives of ε and σ.) Anelasticity excludes amplitude dependent effects, for as long as condition (1) applies, the internal friction and ΔM effect will be independent of amplitude. As for condition (2) in the case of the standard linear solid, it follows from the defining assumptions that there must be only one equilibrium value of strain for every value of stress. It is readily shown that the Maxwell solid does not obey the second condition for anelasticity, although it clearly satisfies the first condition. (If the stress is brought to zero after non-elastic deformation has occurred, equation (10) states that $\dot{\varepsilon}'' = 0$, i.e. there is no recovery of the non-elastic strain in the Maxwell solid.)

Coupled Relaxations

It is possible that if several sources of relaxation with different values of τ are operating simultaneously, the internal friction may not be a simple superposition of the values that would be found if each were acting separately. If the relaxation of one source affects the conditions at the second, coupling may be said to occur and the resultant relaxation may be considerably greater than that obtained by superposition. That such coupling of relaxations can occur has been noted recently[25] in connexion with an irregular arrangement of interfaces in a metal. The phenomenon of coupled relaxations, which also falls into the category of anelasticity, will be further discussed in a subsequent section, in relation to the mechanisms that produce it.

General Viscoelastic Behaviour

Although most studies of the mechanism of internal friction in metals have thus far centred about anelastic effects, the two conditions for anelasticity have been too restrictive to be as widely applicable in the study of high polymers. Workers in the latter field have found[26] that in many cases, it is not possible to give adequate description to the mechanical properties of polymeric substances in terms of linear equations. They have also found[27] that equations which call for viscous flow and which, therefore, permit the occurrence of a permanent set, are also necessary to interpret the experimental behaviour of these materials. Any material which can be described in terms of a differential equation in stress, strain, and the time derivatives of stress and strain, whether or not the equation is linear or implies the possibility of a permanent set, is called *viscoelastic*. In metal physics the occurrence of viscoelastic behaviour outside of anelasticity has not been widely reported, probably because extensive enough investigations have not

been made and also because most investigations have been conducted at low temperatures where permanent set is less likely to occur. In one investigation[28, 29] of internal friction at relatively high temperatures, it has been established that the mechanism of damping is one that corresponds to a viscous-type flow under static stress.

On the basis of present experimental evidence it seems as if any viscoelastic behaviour in metals (even the type that produces a permanent set) does not result in strongly amplitude dependent internal friction at the low strain amplitudes at which most experiments are conducted. The maximum strain amplitude ordinarily employed is of the order 10^{-5} or lower. The reason for this independence of amplitude below strains of 10^{-5} is probably that even where a nonlinear equation is involved, it can usually be approximated adequately by linear expressions, at these low amplitudes. Amplitude dependent internal friction plays a more important role in the case of static hysteresis, which will now be discussed.

Static Hysteresis

In this section, it will be shown that a material that manifests no viscoelastic behaviour may still behave non-elastically and produce internal friction and a ΔM effect, as a result of a mechanism entirely different from viscoelasticity.

Internal friction can only occur when stress and strain are not unique functions of one another in vibration, that is, when the dynamic stress-strain curve is a hysteresis loop. In viscoelastic materials, hysteresis arises as a consequence of the appearance of time derivatives in the stress-strain equation coupled with the fact that the experiment is conducted under oscillatory stresses. The occurrence of the hysteresis loop is a consequence of the dynamic nature of the experiment and the area of the loop is very much dependent on frequency. In the case of an anelastic material, for example, if an experiment is conducted statically with stress applied and released sufficiently slowly to keep the material constantly in equilibrium, the area of the hysteresis loop will go to zero; stress and strain will be uniquely related. We have already seen such behaviour in curve *b* of Fig. 3. Hysteresis may also arise, in materials that are not viscoelastic, in the following way: suppose that as the stress is increased there is a definite value of strain for each stress, this strain being attained instantaneously and, therefore, independently of the rate of loading. Now, when a particular maximum strain is attained, the stress is released. Again it is assumed that at each value of the decreasing stress the strain instantaneously adjusts itself to an appropriate value, but the strain, for a given stress during unloading, is not the same as the corresponding value during loading. A permanent residual deformation remains after removal of the stress which can only

be brought back to zero strain by the application of a reversed load. This type of hysteresis loop, illustrated in Fig. 4, is the same whether it is traversed statically or dynamically; it is entirely independent of the frequency of vibration. The stress-strain equation, therefore, does not involve time derivatives, as in the case of viscoelasticity, but is, simply, a multivalued function. This phenomenon is well known in the case of large-scale plastic deformation at low temperatures where the deformation is practically instantaneous, so that the amount of deformation is almost independent of the rate of loading. It was with regard to such hysteresis, obtainable in large-scale plastic flow, that BURGERS[30] suggested the term *static hysteresis* in contrast to the *dynamic hysteresis* displayed by viscoelastic materials. The fact that static hysteresis occurs for stresses greater than those required to start plastic flow may be important in fatigue studies, but certainly not in internal friction at very low amplitudes. It has been known for some time, however, that ferromagnetic substances, because of the phenomenon of magnetostriction, will follow a mechanical hysteresis loop that is independent of frequency and which gives rise to internal friction at low amplitudes. Recently it has been demonstrated[11, 12] that damping related to the motion of dislocations in a single crystal originates from a static hysteresis loop at strain amplitudes as low as 10^{-7}.* It, therefore, appears that static hysteresis may occur at exceedingly low amplitudes, and that it may result from atomic as well as magnetic re-arrangements. Obviously, such re-arrangements do not actually occur instantaneously, but are probably propagated with the velocity of sound. Such velocities are sufficiently "instantaneous" at frequencies normally used in vibration experiments.

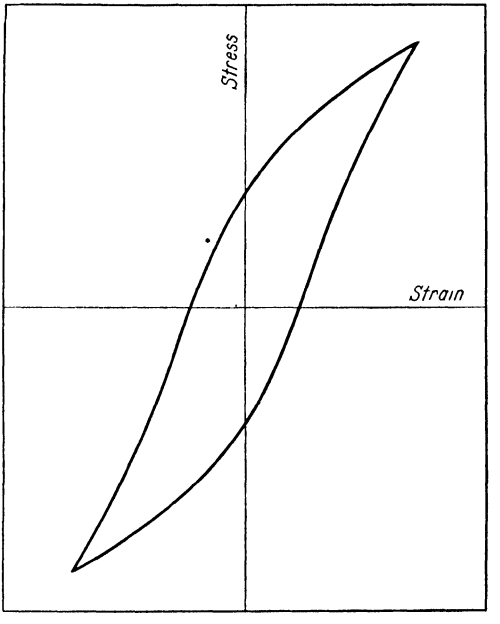


Fig. 4. Schematic illustration of a static hysteresis loop. The same loop is obtained regardless of the rate of traversal

In contrast to viscoelastic internal friction, that originating in static hysteresis seems to show a marked amplitude dependence even at very

* The writer in his original description of this behaviour called it "simple hysteresis."

low amplitudes. Such observations are not at all surprising because the static hysteresis loop will, in general, deviate strikingly from a linear relationship (as, for example, Fig. 4).

In order to calculate ϕ and $\Delta M/M$ for a static hysteresis loop, it is convenient to consider only the loop of non-elastic strain ε'' versus stress. Let us represent the slope of these curves, $d\varepsilon''/d\sigma$, by*

$$\begin{aligned} d\varepsilon''/d\sigma &= F(\sigma) \text{ (loading)} \\ &= f(\sigma) \text{ (unloading)} \end{aligned} \quad \ldots \quad (45)$$

If $\sigma = \sigma_1 \cos \omega t$, where σ_1 is the amplitude, ε'' is periodic but in general not simple harmonic. It is desired, in the usual way, to take the fundamental Fourier components of ε'' in phase with and 90° behind the stress. These components, ε_1'' and ε_2'' respectively, may be obtained by Fourier analysis (after partial integration to obtain integrals of $d\varepsilon''/d\sigma$ rather than of ε'') and inserted into the expressions (15) and (17) to give

$$\phi = \frac{2M'}{\pi} \int_0^{\pi/2} [F(\sigma_1 \cos \theta) - f(\sigma_1 \cos \theta)] \cos \theta \sin \theta d\theta \quad \ldots \quad (46)$$

$$\frac{\Delta M}{M} = \frac{2M'}{\pi} \int_0^{\pi/2} [F(\sigma_1 \cos \theta) + f(\sigma_1 \cos \theta)] \sin^2 \theta d\theta \quad \ldots \quad (47)$$

In the derivation of equations (46) and (47) it is assumed that F and f are even functions, i.e. the two halves of the hysteresis loop are identical. Such an assumption will generally be valid under steady-state conditions. In the case of the parallelogram loop, where F and f are constants, the quantities ϕ and $\Delta M/M$ are independent of the stress amplitude σ_1. In the more general case, it may be convenient to use power series expansions for the functions F and f:

$$\left.\begin{aligned} F(\sigma) &= A_0 + A_1|\sigma| + A_2\sigma^2 + \ldots \\ f(\sigma) &= a_0 + a_1|\sigma| + a_2\sigma^2 + \ldots \end{aligned}\right\} \quad \ldots \quad (48)$$

The absolute value $|\sigma|$ appears in the linear term because F and f are even functions. When these expressions are substituted into equations (46) and (47) and the appropriate integrals evaluated, the amplitude dependence of ϕ and $\Delta M/M$ may be obtained as power series in the strain amplitude

$$\left.\begin{aligned} \phi &= \frac{M'}{\pi}[(A_0 - a_0) + \tfrac{2}{3}M(A_1 - a_1)\varepsilon_1 + \tfrac{1}{2}M^2(A_2 - a_2)\varepsilon_1^2 + \ldots \\ \frac{\Delta M}{M} &= M'\left[\tfrac{1}{2}(A_0 + a_0) + \tfrac{2}{3}\frac{M}{\pi}(A_1 + a_1)\varepsilon_1 + \tfrac{1}{8}M^2(A_2 + a_2)\varepsilon_1^2 + \ldots\right] \end{aligned}\right\} \quad \ldots \quad (49)$$

* It is assumed, in this treatment, that the functions F and f are independent of the stress amplitude. This assumption is a restrictive one, and the question of its validity can only be determined when the mechanism of the non-elastic behaviour is known.

where we have substituted $M\varepsilon_1 = \sigma_1$. In general, if there is no reason to expect that the stress-strain curve is a parallelogram loop, the occurrence of static hysteresis is closely associated with amplitude dependent effects even at very low amplitude. This characteristic amplitude dependence together with the frequency independence of internal friction resulting from static hysteresis make it possible to distinguish this type of behaviour from viscoelasticity.

Summary of Phenomenological Theory

The various mechanisms that produce damping may be classified into the following categories of non-elastic behaviour:

I. Viscoelasticity (dynamic hysteresis)
   A. Anelasticity
      1. Single time of relaxation (standard linear solid).
      2. Superposition of relaxations.
      3. Coupled relaxations.
   B. Processes involving viscous flow.
II. Static hysteresis.

More complex behaviour, which may be represented as a combination of viscoelasticity and static hysteresis, is also possible but has not been investigated as yet.

Methods of Measurement

The methods most commonly used in the measurement of internal friction will be classified and the general principles discussed briefly in this section. The specific variations of each general method have been too numerous to make a detailed discussion of them possible here.

The Torsion Pendulum

This general type of method is essentially distinct from the others to be discussed, in that it is a low frequency method. The use of an auxiliary inertia member of relatively large moment of inertia compared to that of the specimen to which it is attached, usually results in a low torsional frequency of vibration, in the range 0·1 to 15 vibrations per second. The damping is determined by observation of the rate of free decay, i.e. the logarithmic decrement (equation (25)), usually by means of an optical method, where a beam of light is reflected by a mirror mounted on the inertia member and viewed on a translucent scale at an appropriate distance from the mirror. The initial excitation of vibrations is usually accomplished by making part of the inertia member out of a ferromagnetic material and suitably placing an electromagnet so as to attract this part when a switch is closed to activate the magnet.

If the frequency of vibration is less than about 2 sec^{-1} it is possible to obtain the amplitude as a function of number of vibrations either visually or by automatic (usually photographic) recording; at higher frequencies the automatic method must be used.

The torsion pendulum method is the oldest and simplest method for measuring internal friction. This method has a long history, dating back to COULOMB at the end of the eighteenth century, and having recently come into wide use in such modifications as those of CHEVENARD[31] and BOULANGER[32] (the Coulomb micropendulum), KÊ,[33] and FINE.[34] The torsion pendulum of KÊ has been imitated by several authors and used in most of the recent measurements at low frequencies. It requires a specimen in wire form, usually 6 to 12 in. long and about 0·03 in. in diameter. The Coulomb micropendulum, on the other hand, uses a specimen about $\frac{1}{2}$ in. long and 0·06 in. in diameter. Automatic recording methods have been used by several experimenters.[32, 34, 35]

The strain amplitude in a cylindrical specimen of circular cross section, vibrating in a torsion pendulum, is variable, from a maximum value at the surface to zero at the cylinder axis. The maximum amplitude of shear strain in the specimen is usually of the order 10^{-5}. The torsion pendulum is particularly useful for large values of damping up to values of ϕ of the order 0·1, while the lower limit is not much below 10^{-4}, even when the chamber is evacuated to eliminate air damping effects.[28] The study of materials that show very low damping, of the order 10^{-5}, has not been possible, thus far, with the torsion pendulum because of the background damping due to losses in the apparatus.

It is possible to obtain the amplitude dependence of internal friction from a decay curve in free vibration, as already pointed out in the last section. The dependence of vibration frequency on amplitude is much more difficult to determine by this method. To obtain the dependence of internal friction on temperature, the specimen is placed inside a vertical tubular furnace, preferably noninductively wound.

High Frequency Methods

These methods are usually characterized by the absence of an auxiliary inertia member; the specimen, in the form of a rod, is usually vibrated in one of its normal modes. Both free decay and steady-state forced vibrations may be studied. The frequency range varies depending on the type of vibration produced. For example, in torsional or longitudinal modes of vibration of cylindrical specimens, the frequency, which depends primarily on the length of specimen, usually ranges from 10^4 to 10^6 cycles per second. Flexural vibrations of a rod, on the other hand, usually are observed in the frequency range 300 to 10,000 cycles

per second, and the frequency depends on both the diameter and length of the specimen.

There are a wide variety of methods that have been used for the excitation and detection of vibrations at high frequencies. In most of them the power source is an electrical oscillator. The methods generally differ in the manner by which electrical oscillations are converted into the mechanical vibrations of the specimen. Detection consists of a

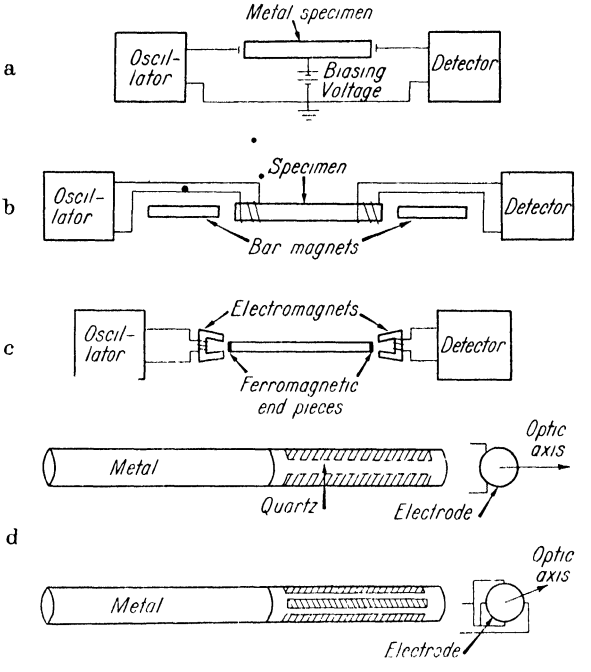


Fig. 5. Schematic diagrams of the apparatus for various high frequency methods of measurement of internal friction

means for reconverting mechanical into electrical vibrations with subsequent amplification. Most investigators have employed parallel schemes for excitation and detection, but in some cases completely different methods have been used for these two purposes. The most common high frequency methods, which are illustrated schematically in Fig. 5 (taken from SIEGEL[36]), are as follows.

(*a*) The electrostatic method[37–39] treats one end of the specimen as one plate of a condenser, applying an alternating potential difference between the specimen and an adjacent fixed plate to excite mechanical vibration. Detection of the vibration may be achieved by forming a similar condenser at the other end which is used as a condenser microphone. Frequency modulation has also been used[39] as a method of

detection. If the specimen is not a conductor, the electrostatic method may still be used if a metal foil is cemented to the specimen.[40] The electrostatic method is suitable for longitudinal and flexural vibrations; it may also be used for torsion[37] if electrodes are attached to the specimen in an appropriate manner.

(*b*) The eddy-current method[20, 41] is limited to longitudinal vibration of specimens that are good conductors. In order to drive the specimen, one end of it is surrounded by a coil and placed in the field of a permanent magnet. An alternating current from an oscillator through the coil induces eddy currents in the specimen which, in turn, interact with the magnetic field to set the specimen into vibration. If a similar coil and magnetic field are located at the other end of the specimen, the motion of the specimen in this magnetic field induces eddy currents in the specimen. These currents, in turn, induce an electromotive force in the coil, which may then be amplified. The amplitude of vibration produced in the specimen is very small, since the electromechanical coupling is weak. The method is, therefore, only suitable for the study of very low internal friction, where not too much power is needed to drive the specimen. Adequate shielding is required to prevent direct coupling between the detector and input coils. The chief advantage of the method is the fact that nothing is attached to the specimen which might produce internal strains and, therefore, additional spurious contributions to the damping.

(*c*) The magnetic method[42–44] requires that pole pieces (discs of ferromagnetic material) be attached to the specimen unless it is itself ferromagnetic. A coil, placed close to one of the discs is connected to the oscillator and the resulting alternating magnetic field interacts with this pole piece to drive the specimen. A second coil near the other disc develops an induced *e.m.f.* which is amplified and detected. The positioning of the pole pieces and coils determines the mode of vibration; the method may be used for longitudinal, flexural, and even for torsional[42] vibration by appropriate arrangement of the pole pieces. An additional advantage of the method is that it may be used for materials that are not electrical conductors. It is particularly useful for the study of large damping, because of the relatively strong electromechanical coupling. On the other hand, the possibility that the cementing of pole pieces to the specimen produces internal strains which contribute to the damping, means that this method is not especially good for cases of very small damping. An interesting variation of this method is that of FÖRSTER,[45] in which the two discs are not directly cemented to the specimen but attached to it through the supporting wires. The specimen is then driven by the oscillator through one wire and the vibrations detected through the other. The method is particularly useful for measurements at high temperatures, since the supporting wires may be

brought outside the furnace so that all the electrical equipment is at room temperature.

(*d*) The piezoelectric method of QUIMBY[46, 47] utilizes a quartz crystal, to which the specimen is cemented. Application of an alternating potential difference across the quartz sets the two cylinders into vibration as a single resonant body, or a "composite oscillator." Longitudinal and torsional vibrations are produced respectively, by means of an X-cut and a Y-cut quartz crystal. The positioning of the electrodes for these two cases is shown in Fig. 5*d*. The equivalent electrical circuit of this composite oscillator at the resonant frequency depends on the elastic and dissipative properties of both the quartz and the specimen, but the contributions due to the quartz crystal may be determined separately and taken into account. In particular, the equivalent resistance of the circuit, which may be measured in an a.c. bridge,[48] gives a composite internal friction from which that of the quartz alone may be subtracted. The contribution of the adhesive between quartz and specimen to the measured damping is small if the natural resonant frequency of the specimen is closely matched to that of the quartz crystal, so that a node of stress occurs at the junction. A range of values of maximum strain amplitude from about 10^{-8} to 10^{-5} may be obtained. READ[9] has shown how this amplitude is calculated from measurable quantities. This method has been the most important one for the measurement of amplitude dependent internal friction, because the internal friction at any amplitude may be obtained directly from the equivalent resistance at resonance, once the proportionality constant relating these quantities is known. The specimen need not be an electrical conductor in this piezoelectric method, but the requirement of close matching in the individual resonant frequencies of the quartz and specimen is an important limitation. In addition, it had been found[49] that strains at the junction resulting from differences in thermal expansion of the quartz and specimen, made significant contributions to the damping, and thus defeated attempts to measure internal friction as a function of temperature by this method. This difficulty is resolved[11] by the use of a liquid adhesive film; however, the handling of the composite oscillator becomes more difficult when a liquid adhesive is used.

In all the high frequency methods, the external contribution to the damping is limited only to losses in the supports, if the system is evacuated. The contribution to the damping resulting from the presence of air in the system is usually of the order 10^{-4}; that due to the supports can be made considerably lower if the supports are located at displacement nodes. Thus, values of ϕ as low as 2×10^{-6} have been reported[50] in evacuated systems.

The high frequency methods are suitable for the study of either free

or forced vibrations. If the damping is sufficiently small and frequency not too high (i.e. not too large a decay in 1 sec), it is often possible to measure directly the time of free decay to a particular fraction of the initial amplitude. More generally, electrical circuits may be used[51–54] to time the reduction of amplitude by any predetermined factor. In contrast to the damping in free decay, which becomes easier to measure for low damping, the resonance curve in forced vibrations becomes more difficult to measure as damping gets very small, because of the sharpness of resonance. Electrical oscillators can be constructed of such stability and with sufficiently fine frequency controls, however, that a resonance curve for which Q^{-1} (equation (23)) is 10^{-5} may be measured without difficulty. It is only necessary that one be able to read frequency *variations* of better than one part in 10^6 to obtain such a resonance curve; the frequency itself need only be known to about 1 per cent.

A method that utilizes frequencies in the megacycle range is the pulse technique, by which a short ultrasonic pulse is produced at one end of the specimen and its time of transit through the specimen and attenuation are measured. Although this method has been used successfully for the measurement of elastic constants,[55, 56] the measured attenuation may be, in a large part due to scattering by imperfections rather than to absorption, because of the short wave lengths employed. This method, therefore, has not as yet been used for quantitative measurements of internal friction.

Stress-induced Ordering: A Source of Internal Friction

A discussion of order concerns itself with the distribution of different atomic species in a solid solution on various lattice sites. If there is any deviation from a completely random distribution in the manner in which lattice sites are occupied, the alloy is said to show a degree of *order*. The concepts of long range and short range order that have been so widely discussed,[57] are special cases of this definition. The application of stress to a system may result in a change in the state of order, if one already exists, or it may produce a state of order in an initially disordered solid solution. For stress to have such an effect, it is simply required that a particular change in the state of order result in the production of strain. This requirement is simply a consequence of the thermodynamic reciprocity relations. For example, in the study of piezoelectricity, it is well known[58] that if the application of stress produces electrical charges on particular surfaces of a crystal, then a voltage across these surfaces will produce strain (electrostriction). It will be of value to derive the reciprocity relation that applies to the change of order induced by stress. Let p represent a parameter that is a measure of the state of order. Deviation from randomness means the

existence of an energy term that favours ordering. Let u represent this ordering energy: the energy decrease per unit volume, per unit increase in p at constant strain (defined so that $u > 0$), or

$$u = (-\partial U/\partial p)_\varepsilon \qquad \ldots . \quad (50)$$

where U is the total energy of the system per unit volume. If we anticipate that u is not appreciably temperature dependent, we may estimate its value by considerations at absolute zero, where

$$dU = \sigma d\varepsilon - u dp$$

The differential of $U - \sigma\varepsilon$ is, therefore, $-\varepsilon d\sigma - u dp$, and since it is a perfect differential, the reciprocity relation

$$(\partial u/\partial \sigma)_p = (\partial \varepsilon/\partial p)_\sigma \qquad \ldots . \quad (51)$$

is valid. The meaning of this relation is that if a change of order at constant stress produces strain, the application of a stress will increase the ordering energy and will, therefore, shift the equilibrium in the direction of a greater degree of order.

In this section three examples of such stress-induced ordering will be discussed. In the first two cases the state of the material for zero stress is one of randomness or disorder; in the third example, we will consider initially ordered materials. In all cases, it is expected that the behaviour of the material under stress will correspond very closely to that of the standard linear solid, since the three conditions that define such a solid are valid to a good approximation. First, it is reasonable to assume that the non-elastic strain that accompanies a small change in p is proportional to this change. Equation (34) is, therefore, valid if p is redefined as the difference between the state of order existing and that at zero stress, so as to obtain a direct proportionality between ε'' and p instead of a linear relationship. Furthermore, the change in the equilibrium state of order under stress is proportional to the change of ordering energy, u, which is, in turn, proportional to the stress (again assuming small changes). Finally, the rate of approach to equilibrium is determined by the difference between atomic jumps that lead to greater and to lesser degrees of order per unit time, which is in turn, proportional to the deviation of p from its equilibrium value. (For p equal to its equilibrium value $\bar{p}$, the rate of occurrence of atomic jumps that act to increase p and those that act to decrease p must be equal.) Equation (35) is, therefore, valid. The mechanical properties of a material in which stress-induced ordering occurs will, therefore, be those of a standard linear solid. In particular, measurement of internal friction as a function of vibration frequency will show the characteristic peak whose equation is given by (39). The time of relaxation τ, associated with stress-induced ordering must bear a simple relation to

the mean frequency of atomic jumps, since the mechanism of the change of the state of order must be the same as that for volume diffusion. It follows that

$$1/\tau = \alpha\Gamma \qquad \ldots . \quad (52)$$

where Γ is the appropriate atom jump frequency and α a dimensionless constant. Inasmuch as a change in the state of order of a material can be accomplished by a small number of atomic jumps, the magnitude of α is of the order of unity. It therefore appears that by simply locating an internal friction peak, which determines τ according to the relation $\omega\tau = 1$, it is possible to measure atomic jump frequencies in appropriate alloys.

The application of internal friction measurements to the study of diffusion phenomena is based on equation (52). Let us contrast the internal friction methods to conventional diffusion measurements. In the latter case, there must be a concentration gradient initially present in the specimen, and one obtains the change in the concentration-versus-distance curves as a function time. From these curves a diffusion coefficient may be obtained, which is related to the appropriate atomic jump frequency by the equation

$$D = \beta a^2\Gamma \qquad \ldots \quad (53)$$

where a is the lattice parameter, and β a dimensionless constant which depends on the geometry of the lattice structure and is of the order of magnitude unity. Since diffusion is a random walk problem, the mean distance that an atom travels from its starting position is proportional to $t^{1/2}$ (t = time) and, therefore, to the square root of the number of atomic jumps. Thus, in order that diffusion takes place, even over distances of the order of 0·1 mm, a mean number of atom jumps of about 10^{10} to 10^{11} is required. If such a large number of jumps is to occur within the time limit of an experiment, the jump frequency must be very great. The magnitude of Γ and, therefore, of D can be controlled by selecting the appropriate temperature, since both quantities obey an Arrhenius equation,

$$\left.\begin{aligned} D &= D_0 e^{-H/RT} \\ \Gamma &= \Gamma_0 e^{-H/RT} \end{aligned}\right\} \qquad \ldots . \quad (54)$$

where H is called the heat of activation for diffusion. Conventional diffusion measurements are, therefore, limited to high temperatures. For example, the temperature at which the jump frequency is 1 sec^{-1} is far too low for diffusion measurements, since the experiment requires a time of at least 10^{10} sec or 300 years! On the other hand, such a jump frequency can be measured very conveniently by internal friction methods when the period of vibration is of the order of 1 sec. The

restriction of conventional diffusion measurements to high temperatures has led to the following limitations: (1) only a small range of values of D can be measured (usually D can be obtained only over two to three powers of ten); (2) when diffusion is due to the movement of vacancies, it is not possible to learn more about the process by trapping non-equilibrium numbers of vacancies at low temperatures and detecting their presence by the abnormally rapid atomic movement that they produce; (3) it is not possible to study the effect of dislocations (cold work) on diffusion because experiments are conducted above the recrystallization temperature; (4) precipitation cannot be studied by direct measurement of the rate of atomic movement in supersaturated solid solutions. The use of internal friction methods for the study of diffusion phenomena, has already contributed towards getting these investigations out of the strait jacket of high temperature measurements. In those cases in which internal friction methods are applicable, contributions have been made toward the elimination of each of the limitations listed above.

Examples of Interstitial Solutions in B.C.C. Lattices

The study of stress-induced ordering by internal friction methods has been most extensive in the case of interstitial solid solutions in b.c.c. (body-centred cubic) lattices. Most investigations have been concerned with carbon and nitrogen as interstitials in α-iron; in fact, the internal friction effects of this general type were first discovered and explained for the case of α-iron. In early work[59] it was observed that the damping of steel tuning forks goes through a maximum as a function of temperature and that the location of this maximum is dependent upon the frequency of the fork. SNOEK[60] showed that interstitial solutes were essential to this damping peak by demonstrating that the effect no longer occurred when all traces of carbon and nitrogen were removed. He also first explained[61] the origin of the effect as follows: interstitial carbon and nitrogen in b.c.c. iron are not located in the largest spaces, but at the centres of the cell edges, i.e. positions of the type $(\frac{1}{2}, 0, 0)$, and at the face centres, which are equivalent positions. An interstitial atom located in one of these positions will produce a distortion that has tetragonal symmetry, i.e. most of the distortion is in the direction of the two closest iron atoms (one of the $\langle 100 \rangle$ directions). If the directions of the cube axes are designated as x, y and z, respectively, the interstitial sites may be designated as x, y or z *sites* according to the direction in which maximum distortion is produced. Under zero stress, the three kinds of interstitial sites are occupied in a random fashion, in accordance with the concentration of interstitial atoms; each type of site, therefore, has one-third of the total number of interstitials. Since an excess of atoms in any one type of site would result in an elongation in the

corresponding direction, our reciprocity equation (51) requires that the application of a tensile stress in the z direction, for example, will lower the energy for atoms in z sites as against x and y sites. Such a tensile stress will, therefore, produce a redistribution of atoms such that the z sites are occupied preferentially. A state of order is thus induced by stress where a completely random distribution existed for zero stress.

The atomic description given by SNOEK to explain the internal friction peak observed in α-iron is so complete, that a calculation of the anelastic parameters Δ_M and τ in terms of physical constants is readily carried out in terms of this description. It is convenient to define the parameter p as

$$p = n_z - n/3 \qquad \ldots . \quad (55)$$

where n is the number of interstitials and n_z the number in z sites, per unit volume. The non-elastic strain in the z direction is given by equation (34), and its equilibrium value at any stress, by equation (36), where M is now replaced by Young's modulus in the $\langle 100 \rangle$ direction, which we will call E. The quantity u, equation (50), is now the decrease in energy when one interstitial is brought from an x or y site to a z site. When equations (34) and (36) are combined, we obtain

$$\Delta_E = \lambda E' \bar{p}/\sigma \qquad \ldots \quad (56)$$

In order to complete the derivation of Δ_E and to obtain the relaxation time τ, it is necessary to calculate the net rate of change of n_z. This quantity is given by

$$\dot{n}_z = \Gamma_{yz} n_y + \Gamma_{xz} n_x - (\Gamma_{zx} + \Gamma_{zy}) n_z \qquad \ldots . \quad (57)$$

where n_y and n_x are defined analogously to n_z, Γ_{yz} is the rate (or probability per second) of transitions from y sites to z sites, and similar definitions apply to Γ_{xz}, Γ_{zx} and Γ_{zy}. A consideration of the geometry of the lattice shows that, if all jumps of interstitials occur between sites that are nearest neighbours, then an atom in a z site may move only into four positions of which two are x — and two y — sites (similarly for other transitions). Furthermore, the activation energy is no longer the same for jumps into and out of z positions. Fig. 6 shows that if the height of the barrier under zero stress is H, in accordance with equation (54), then the application of stress must raise the barrier for $z - x$ transitions by $u/2$, and lower the barrier for $x - z$ transitions by this same amount; the energy difference between the z and x states, is u, as already mentioned. These arguments lead to the conclusion that

$$\Gamma_{yz} = \Gamma_{xz} = \tfrac{1}{2}\Gamma_0 e^{-(H-u/2)/kT} = \tfrac{1}{2}\Gamma e^{u/2kT} = \tfrac{1}{2}\Gamma(1 + u/2kT)$$

and

$$\Gamma_{zy} = \Gamma_{zx} = \tfrac{1}{2}\Gamma(1 - u/2kT)$$

where Γ is the rate of jump of interstitials under zero stress, and $u/kT \ll 1$ because the applied stress is assumed small. These terms may now be inserted into equation (57) to obtain

$$\dot{n}_z = -\frac{3}{2}\Gamma\left[(n_z - n/3) - \frac{2}{9}\frac{un}{kT}\right] \quad . \quad . \quad . \quad . \quad (58)$$

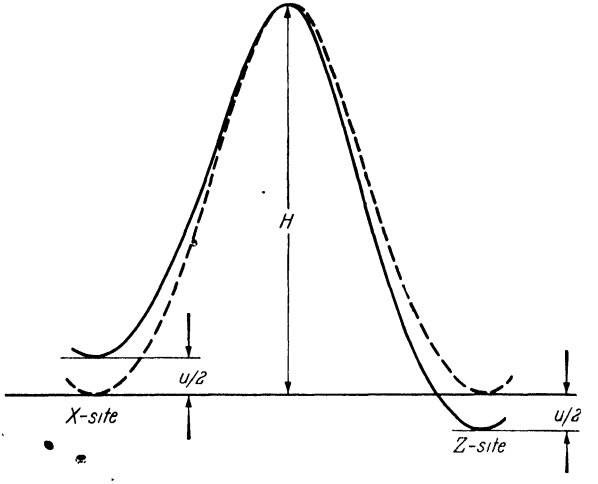


Fig. 6. The effect of a tensile stress in the z direction on the energy barrier for the jump of interstitial atoms between x and z sites

where $n_x + n_y$ has been set equal to $n - n_z$, and the term un_z is approximately equal to $un/3$, since u is already small, of first order. Comparison of equation (58) with equations (35) and (55) shows that

$$\tau^{-1} = \tfrac{3}{2}\Gamma \quad . \quad . \quad . \quad . \quad (59)$$

(or, in terms of equation (52), $\alpha = 3/2$), and

$$\bar{p} = \tfrac{2}{3}\, un/kT \quad . \quad . \quad . \quad . \quad (60)$$

If u is assumed to depend only on σ and not on p (this assumption means that mutual interaction between interstitials is not important, and should certainly be valid for the low concentration of interstitial solutes usually encountered) then, from the reciprocity equation (51), $u = \lambda\sigma$, where λ is defined by equation (34). When this result is substituted into equation (60) and the resulting expression for $\bar{p}$ inserted into (56), we obtain the final expression for Δ_E:

$$\Delta_E = (\tfrac{2}{3}\, E'\lambda^2/k)n/T \quad . \quad . \quad . \quad . \quad (61)$$

The relaxation strength Δ_E is, therefore, proportional to the solute concentration and varies inversely as the absolute temperature. To calculate Δ_E requires only a knowledge of $\lambda = (\partial\varepsilon/\partial p)$ which can be

estimated[62] for the important case of carbon in α-iron from the change in lattice parameters of tetragonal martensite as a function of carbon concentration. The constant that relates Δ_E to n/T can also be determined experimentally from the height of the internal friction peak for a known concentration of solute. The value of this constant is, however, dependent on the crystallographic direction of application of the tensile stress. The above derivation applies to a stress in the $\langle 100 \rangle$ direction; if the stress were in the $\langle 111 \rangle$ direction, SNOEK'S theory predicts that $\Delta_E = 0$, since all three tetragonal axes are then equally inclined to the direction of application of stress. The verification of this predicted result by DIJKSTRA,[63] who found, in single crystals of α-iron, that Δ_E for the $\langle 111 \rangle$ direction is negligible compared to its value in the $\langle 100 \rangle$ direction, is the most striking evidence of the correctness of SNOEK'S picture of the relaxation process. The value of the constant term in the expression (61) for Δ_E has been obtained experimentally, for carbon and nitrogen in polycrystalline α-iron, by DIJKSTRA; it was found that, for internal friction peaks close to room temperature, the maximum damping (which is equal to $\Delta_E/2$) is very closely equal to the weight per cent of interstitial solute. Because of the strong orientation dependence of Δ_E, already discussed, it should be remembered that any texture in the specimens used, will lead to errors in the constant that relates the height of the peak to the solute concentration. On the other hand, relative values of the height of the peak, which may be studied as a function of heat treatment, give a precise measure of the relative residual concentration of interstitial atoms in solution at any time. This is the basis for the study of precipitation which will be described below.

Although the *height* of the internal friction peak is sensitive to orientation, the temperature of maximum internal friction for a given frequency, i.e. the *location* of the peak, is independent of both orientation and the concentration of solute, but is a measure only of the jump rate of interstitials at that temperature (equation (59)).

The generality of the SNOEK mechanism has been demonstrated by the observation of effects, in other b.c.c. lattices containing interstitial solutes, similar to those in α-iron. Thus KÊ investigated relaxation effects due to carbon and oxygen,[64] and also to nitrogen,[65] in tantalum. In the case of oxygen the observed internal friction peak is considerably broader than that corresponding to a single time of relaxation. It was, therefore, suggested that oxygen interstitials in tantalum may be located in two different kinds of positions, so that two types of jumps are possible. Further investigation of carbon in tantalum and also in columbium has been carried out by WERT.[66] In the nonmetallic category, internal friction peaks observed[67] in glass have been attributed to the diffusion of sodium and potassium ions to preferential positions

in the glass structure, in very much the same way as in the case of metals.

The internal friction peak originating in the stress-induced ordering of interstitial solutes has been utilized for extensive investigations of precipitation from supersaturated solutions, especially in the case of α-iron. It has already been shown that Δ_E is proportional to n, the concentration of interstitials *in solution*. This is not the same as the total concentration of solute present. The solute is usually introduced into the lattice by a high temperature treatment in an appropriate chemical atmosphere. At this high temperature, the solubility is usually considerably greater than at lower temperatures, so that if equilibrium is to be maintained, precipitation of a second phase rich in solute must take place at lower temperatures. The solute thus precipitated is usually tightly bound, in the form of an intermetallic compound, and cannot contribute to the internal friction. The height of the peak is, therefore, a measure only of the concentration of interstitial atoms in solid solution, and decreases as precipitation takes place. This suggests a unique method for following the precipitation, since the actual amount of precipitate can be obtained quantitatively as a function of ageing time and temperature. No other physical property that has been used to follow the course of precipitation (electrical resistivity, thermoelectric effect, magnetic properties, density, and microscopic investigation) is determined so completely by the *quantity* of precipitate, and is independent of the condition or size of precipitate particles. This method was first applied by DIJKSTRA[68] to the study of the precipitation of nitrogen and carbon in α-iron. The internal friction peak for nitrogen is at 20°C, and for carbon at 36°C, when the vibration frequency is about 1 sec^{-1}. The rate of precipitation is negligible at these low temperatures, so that a specimen is first quenched from the temperature at which carbon or nitrogen is introduced and then tempered for various periods at the desired precipitation temperature. After each heat treatment the specimen (in wire form) is quenched to prevent further precipitation and the height of the peak determined in a torsion pendulum. In this way, one can obtain the amount of solute left in solution, therefore, the amount precipitated, as a function of tempering time at a given temperature. To obtain a similar tempering curve at a second temperature, it is simply necessary to give the original specimen a re-solution heat treatment and quench. Using this method, DIJKSTRA[68] finds that the precipitation of nitrogen in α-iron occurs in two successive stages. In the first, plates of an unknown precipitate (called "N phase 1") form on (100) planes of the matrix (determined by microscopic examination) and in the second, Fe_4N plates form. It is also found that precipitation is much faster in cold worked α-iron and that the same type of precipitate is no longer formed. This study of precipitation in iron was

continued by WERT[69] and a quantitative interpretation given to the measurements by ZENER.[70] It is shown that the amount of precipitate q would, in the initial stages of the process, follow the law

$$q = b(t - t_0)^r \qquad \ldots \quad (62)$$

where t_0 is the time for nucleation (time at which the nucleus is formed, so that growth occurs thereafter by direct diffusion of solute atoms to the nucleus). ZENER also shows that the value of r depends on the shape of the precipitate particles; for spherical particles $r = 3/2$, for cylindrical particles $r = 2$, and for circular discs $r = 5/2$. The results for precipitation of carbon in α-iron give $r = 3/2$, and for the first stage (N phase 1) of nitrogen precipitation, $r = 5/2$; the shape of particles in these two cases is, therefore, determined. In all cases t_0 is found to be zero within experimental error. Apparently nuclei are present as soon as the material is quenched or, if there is an incubation time, it is very short compared to the time for precipitation.

The experimental curves deviate from the simple power law of equation (62) in later stages of the precipitation process, where competition between adjacent particles for the solute atoms in a given region of solution becomes appreciable. More exact laws are derived to take these interference effects into account.[71] ZENER also shows how the number of precipitate particles per unit volume, and, therefore, the mean spacing of particles, can be obtained from experimental values of the constant b of equation (62).

As a part of the precipitation experiments, it is possible to determine the concentration of solute remaining in solution at any given temperature after precipitation is complete, i.e. the solubility curve for the particular interstitial solute. Smaller solubilities can be determined by the internal friction method than those that one can usually determine by chemical methods. The solubility line for carbon in α-iron was determined[72] down to 150°C, where a solubility of only 10^{-4} weight per cent was observed. This value is lower than all previous estimates and more directly obtained.

The use of the height of the internal friction peak in interstitial solutions to measure the degree of precipitation is extended to the study of precipitation of carbon and nitrogen in cold worked α-iron by HARPER.[73] For a sufficiently large amount of cold work (5 per cent extension is sufficient) normal precipitation, as studied by WERT, does not occur. Instead, an analysis by COTTRELL[74] and BILBY[75] suggests that precipitation of solute atoms into edge dislocations takes place. These authors calculate that such precipitation will follow a law of the form of equation (62) with $r = 2/3$. HARPER verified this power law experimentally, and with the aid of the formula of COTTRELL and BILBY, obtained the density of dislocations in the specimens for

different degrees of cold work. Similar results were obtained for both carbon and nitrogen in α-iron. It is striking that these two cases, for which normal precipitation occurs according to equation (62) with two different r values, follow a 2/3 power law when precipitation occurs in a *cold worked* lattice.

The final illustration of the versatility of SNOEK's internal friction peak is its application to the study of diffusion. The potentialities of internal friction methods for such investigations have already been discussed in connexion with equation (52). In the case of interstitial solutes, we have derived equation (59), which enables us to determine atomic jump rates directly, and quantitatively, from measured relaxation times. The combination of low temperature internal friction measurements with high temperature conventional diffusion measurements makes possible the determination of D and Γ over a very wide range; H and D_0, or Γ_0 (equation (54)), may then be determined with great precision. It would seem that the atomic jump rate Γ is the most fundamental quantity with which we are here concerned, and that both the measured diffusion coefficients and relaxation times would be converted into the appropriate values of Γ by means of equations (53) and (59), respectively. The custom of reporting diffusion coefficients has been so widespread, however, that more often τ-values are converted into diffusion coefficients and compared directly with high temperature diffusion measurements. This focus of attention on D rather than Γ does not lead to any complications in interpretation, because of the simple proportionality between the two quantities. SNOEK[60, 61] reported values of D_0 and H for diffusion of carbon and nitrogen in α-iron, based on his measurements. More precise internal friction data were later reported by WERT and ZENER.[76] The most complete data for carbon in α-iron are obtained[77] by the combination of static relaxation measurements with internal friction and conventional diffusion measurements,[78] which give $D_0 = 0{\cdot}02$ cm^2/sec and $H = 20{,}100$ cal/mole, from a straight line plot of log D versus T^{-1} that covers fourteen powers of ten in D over the temperature range — 35°C to 800°C. It has been shown[76] that values of D_0 may be predicted theoretically if it is assumed that the free energy of activation, or the work required to bring an interstitial atom to the midpoint between two interstitial sites, goes entirely into *elastic straining* of the surrounding lattice. The value of D_0 calculated from this theory for the case of carbon in iron is in excellent agreement with the experimental result reported above. Earlier reported values of D_0 had not fitted the theory so well, because of the difficulty in a precise determination of D_0 when the data for D are only over a limited temperature range. (D_0 is obtained from the intercept at $1/T = 0$ of the straight line plot of log D versus $1/T$. Low temperature data, as obtainable by internal friction methods, are, therefore, very valuable in

making possible a precise extrapolation of the straight line.) In other systems, particularly nitrogen in iron and carbon in tantalum, the agreement of reported values of D_0 with the theory has improved,[66] as the range of measurement was extended. This verification of the strain theory of D_0, wherever precise data for interstitial diffusion is available, suggests that, in other interstitial solutions, the heat of activation can be obtained from only a single measurement of the diffusion coefficient D by the substitution of the calculated D_0 into equation (54).

Examples of Random Substitutional Solutions

A substitutional solid solution that shows no order under zero stress may also show anelastic effects resulting from stress-induced ordering. Such effects have only been observed at relatively high concentrations of solute (when the solute is present to ten or more atomic per cent), in contrast to the interstitial case where minute amounts of solute produce a sizable internal friction peak. The first observation of an internal friction peak due to stress-induced ordering in a substitutional solution was that of ZENER[79] in the case of a single crystal of α-brass (30 per cent Zn). These measurements were only at one frequency and composition and the origin of the peak was not explained. Later KÊ[80] verified the existence of such a peak in (70–30) α-brass at lower frequencies, and obtained a heat of activation in reasonable agreement with the heat of activation for diffusion of zinc in brass. A possible mechanism for this relaxation effect in brass was suggested by ZENER[81] and will be discussed below. The investigation of this type of anelasticity was extended by the writer[82] to α-silver-zinc alloys (silver rich), where similar, but larger, effects were anticipated and found (due to the greater difference in atomic radii of Ag and Zn as compared to Cu and Zn). Extensive work[83] was carried out on the Ag-Zn system and the dependence of internal friction on composition and temperature were obtained. The observed internal friction peak is clearly a volume effect, since it is independent of grain size. However, if the grain size is too fine, the grain boundary internal friction (see the next section) also occurs, thus requiring a resolution of the two effects. The most successful investigations of the ordering internal friction are, therefore, carried out on single crystals or coarse-grained polycrystalline specimens.

Clearly, the explanation for the observations in substitutional alloys cannot be the same as for interstitial alloys. For a lattice that has cubic symmetry, any elastic distortion due to the substitution of a larger or smaller atom at one lattice site must also have cubic symmetry. There are, therefore, no preferential sites for such atoms when stress is applied. On the other hand, the distortion about a *pair* of solute atoms in nearest neighbour positions no longer has cubic symmetry; a pair of

solute atoms will produce tetragonal distortion about the direction of the pair axis. This loss of cubic symmetry in the case of pairs of atoms is the basis for the explanation given by ZENER[81] for the anelastic behaviour of substitutional alloys. He suggests that pairs of substitutional solute atoms behave completely in analogy to single interstitials. For example, if the solute atom is larger than the solvent, then under zero stress, pairs of solutes will be randomly distributed over all crystallographically permissible directions. If, now, a tensile stress is applied in a particular direction, the pair axes will show a preferential alignment in those crystallographic directions which are the nearest to the stress direction. The attempt of the lattice to establish a state of equilibrium at each value of stress will give rise to internal friction in exactly the same way as in the case of interstitial solutions. The time of relaxation will be closely related to the rate of atomic movement in the alloy. A high enough concentration of solute is required, in the substitutional case, for the occurrence of pairs of atoms to be reasonably probable. This stress-induced preferential orientation of pairs of atoms is a form of stress-induced ordering. The type of order involved does not fall into the category of either long range or short range order, for it is assumed that the pairs themselves are formed by random fluctuations. The phenomenon is one of stress-induced ordering nevertheless, since the preference for special directions of pair axes is a deviation from randomness.

Although the description of the relaxation phenomena in substitutional solutions in terms of reorientation of solute atom pairs has the advantage of its direct analogy to the case of interstitial solutions, it cannot be taken too seriously for several reasons. The "pair" concept is probably most nearly satisfactory at low solute concentrations, where most solute atoms are far apart and the number of pairs is quite small. Even in such a dilute solution, where the internal friction effects are probably too small to be observed, it is not correct to think of a fixed number of pairs of solute atoms, since a tendency toward preferential creation or dissolution of nearest neighbour pairs may occur under stress. In actual measurements, solutions containing as high as one-third solute atoms have been used; here clusters of solute atoms considerably more complex than pairs occur with great frequency. The behaviour of such clusters under stress cannot readily be taken into account in terms of a theory of reorientation of solute pairs. In fact, the very concept of *solute* and *solvent* becomes ambiguous as the concentrations approach 50 per cent!

If the pair-reorientation picture is an oversimplification, and, therefore, not an acceptable explanation for the observed behaviour of substitutional solutions, can we still be sure that the phenomenon is one of stress-induced ordering, controlled by volume diffusion within the

alloy? There is excellent evidence that leads to an affirmative answer to this question. In the first place, these effects are observed in well-annealed single crystals and are not structure sensitive. In fact, the observations are quantitatively reproducible in different specimens of the same alloy. The possibility that the observed phenomena are related to the presence of internal interfaces or of dislocations is, therefore, discounted. Secondly, these anelastic effects are only observed in alloys, and are, apparently, larger the greater the atomic size difference of the solute and solvent atoms. Finally, the width of the internal friction peak corresponds very closely to a single time of relaxation, and the heat of activation is of the correct magnitude for volume diffusion. In view of this evidence, it seems almost certain that the relaxation is controlled by the movement of atoms in the lattice, which acts to produce an ordered condition of the alloy under an applied stress. It is, therefore, expected that the measured time of relaxation τ and the appropriate jump frequency Γ, are related by equation (52). The dimensionless constant α cannot be calculated, as it was in the interstitial case, until we know more precisely the exact nature of the ordering process. A reasonable value for α can be estimated,[83] however. In the interpretation of equation (52), it is of great importance to specify what is the appropriate jump frequency Γ. The interpretation of this quantity is not as obvious in a substitutional alloy as in the interstitial case; in a substitutional solution there are two significant rates of atom movement, those corresponding to each of the two types of atoms, say, A and B. If the mechanism of diffusion is not a direct interchange of lattice positions, these two rates, Γ_A and Γ_B, may differ significantly. The well-known experiments of KIRKENDALL[84] and of DA SILVA and MEHL[85] give evidence that the two atomic species actually do not move at the same rate, at least in several f.c.c. alloys. It is, therefore, probable that diffusion occurs by means of the movement of lattice vacancies.[86] In contrast to conventional diffusion measurements where the faster moving atoms make the major contribution to the measured diffusion coefficient,[87] the stress-induced ordering process is controlled by the rate of movement of the *slower* atomic species. This conclusion may be demonstrated as follows: let us suppose that B atoms move much faster than A atoms, or $\Gamma_B \gg \Gamma_A$. At a reasonably high concentration of B, it becomes very improbable that a given B atom does not have at least one nearest neighbour that is also B. The B atoms, therefore, form a continuous network throughout the lattice. If diffusion is by means of vacancies, most of the movement of vacancies will be along this network of B atoms, because the rate of jump of a B into a vacancy is considerably greater than that of an A atom. No redistribution of atoms, or ordering, can be produced by such movement of vacancies along the B network,

since for every B atom that has jumped another simply takes its place. Only on those relatively rare occasions when an A atom makes a jump will a contribution toward re-arrangement of the atom distribution in the lattice occur. Thus Γ_A controls the rate of relaxation, and this quantity is the appropriate jump frequency to be inserted in equation (52). The interpretation of the internal friction effects in terms of the pair reorientation of solute atoms, discussed above, leads one to emphasize the importance of the movement of *solute* atoms,[88] which may not actually be the rate controlling factor at all.

An illustration of internal friction data for Ag-Zn alloys is given in Fig. 7, which shows a series of internal friction peaks at one frequency for different compositions. The figure shows that both the height and location of the peaks are very sensitive to chemical composition. The magnitude of the maximum damping falls off rapidly as the concentration of zinc decreases, due to the smaller degree of ordering that can be produced at lower concentrations. The temperature at which internal friction is a maximum, or at which the time of relaxation $= 1/\omega$ is higher the lower the zinc concentration. This result means that the controlling rate of atomic movement at any given temperature increases with increasing concentration of zinc. In the case of the lowest peak, an additional high temperature damping associated with grain boundaries becomes significant enough to distort the ordering peak. When data similar to that of Fig. 7 is obtained at different frequencies, the temperatures at which the peaks occur, are correspondingly shifted. By carrying out experiments over a wide range of frequencies, the relaxation times for the various silver-zinc alloys may be obtained over a wide range of temperatures, as shown in Fig. 8. Data at the lowest temperatures can only be obtained by the use of frequencies too low to be practicable, and are, therefore, most conveniently obtained by static methods instead of by vibration. The curves of Fig. 8 represent essentially the variation of atomic mobility with temperature.

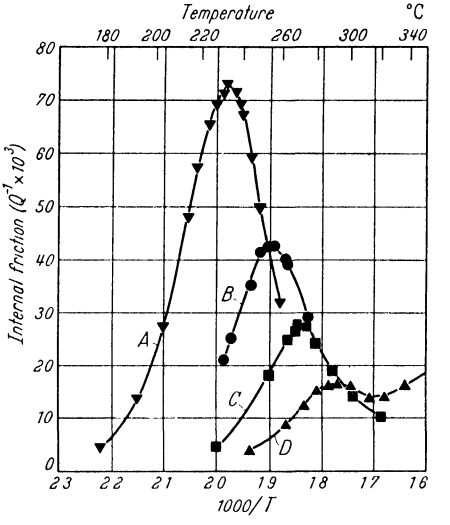


Fig. 7. Internal friction data for a series of silver-zinc alloys (frequency of vibration 0·8 cycle per second at room temperature). The atomic per cent zinc in the various alloys are 30·2, 24·2, 19·3, and 15·8, corresponding to the curves A to D respectively

Unfortunately, no reliable conventional diffusion data is available for extension of these curves to high temperatures. The data obtained by internal friction and static relaxation methods cover a sufficiently large range, however, and is sufficiently precise, that a deviation from the usual straight line semi-logarithmic relation is quite apparent. A straight line plot for log Γ versus T^{-1} is only to be expected if H is a constant (equation (54)). There is no theoretical reason, however, why H may not be dependent on T. Conventional diffusion data have never been reliable enough to report such a deviation, except when channels for diffusion, other than volume diffusion, occur (e.g. diffusion down grain boundaries). The curves of Fig. 8 also show a strong concentration dependence of atomic mobility. An increase of about 10 per cent in zinc concentration means an increase in the controlling rate of atomic movement by a factor of ten. It is interesting to try to see if this strong concentration dependence resides primarily in the term Γ_0 or in H. Extrapolation of the curves of Fig. 8 to $T^{-1} = 0$ shows that the values of Γ_0 for the various concentrations are very nearly the same and, therefore, that the differences in atomic movement as a function of concentration are principally determined by differences in the heat of activation H. The intercepts Γ_0, and the corresponding values of D_0, may be compared with the strain theory of D_0 (discussed earlier in connexion with interstitial diffusion), which has also been applied[89] to substitutional solutions. The agreement with theory is quite good. In the case of conventional

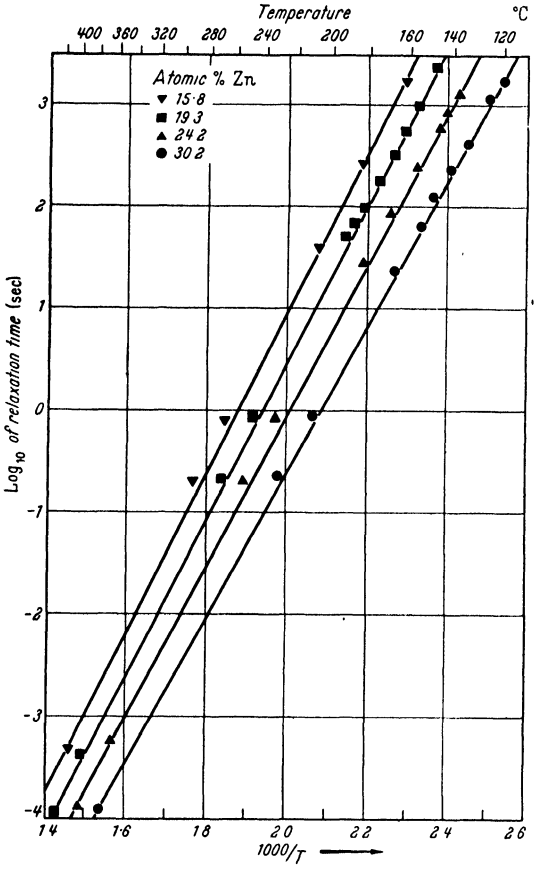


Fig. 8. The variation of the time of relaxation for stress-induced ordering with temperature, for a series of silver-zinc alloys

diffusion data, the agreement between reported values and theoretical values has generally been poor, but this lack of agreement has been related[90] to the presence of mechanisms other than volume diffusion in many of these experiments.

In addition to the above experiments, the stress-induced ordering effect has also been used for experiments fundamental to our understanding of the mechanism of diffusion. In terms of the vacancy mechanism of diffusion, which has been widely accepted, the rate of diffusion or atom movement depends on the product of the concentration of vacancies and the rate of jump of atoms into vacancies. Both of these factors are strongly temperature dependent, but conventional diffusion measurements are incapable of separating the two. On the other hand, by means of low temperature relaxation measurements it has been possible to detect the presence of a nonequilibrium concentration of vacancies in a silver-zinc alloy that had been quenched rapidly from a high temperature.[91] The effect of such quenching is to trap some of the relatively large number of vacancies present at high temperatures, which then produce abnormally rapid atomic movement, and, therefore, abnormally short relaxation times, at low temperatures. For example, the atomic jump rate, at 50°C is increased by a factor of 10^5 over the equilibrium rate in a Ag-Zn wire that had been quenched from 400°C into water at room temperature. Such experiments not only give strong additional evidence for a vacancy mechanism, but also make it possible to study separately the two factors which determine the rate of atomic movement. Furthermore, from observations on the rate at which the "quenched-in" vacancies disappear, it is hoped that the mechanism by which lattice vacancies are created and absorbed will be determined. It has been postulated,[92, 93] for example, that dislocations act as sources and sinks for vacancies.

Examples of Initially Ordered Substitutional Solutions

If a solid solution shows an incomplete state of long range order under zero stress, then, in accordance with the theory presented in the beginning of this section, a change in the state of order will be produced by an applied stress if the ordering process is accompanied by a change of lattice dimensions. This phenomenon of stress-induced change in the degree of order, first predicted by GORSKY,[94] leads to mechanical behaviour corresponding to that of the standard linear solid. Unlike the two cases already considered, for which the solid solution is disordered under zero stress, the characteristic internal friction peak in the present case is observable only as a function of frequency (at constant temperature), and not as a function of *temperature*, at constant frequency. The reason for this difference is that, for the initially ordered solution, the degree of order at zero stress is itself a function of tempera-

ture. Thus, experiments which simply determine internal friction as a function of temperature are difficult to interpret. Such experiments, carried out by KÖSTER[95–98] and collaborators on several different ordering alloys, show that internal friction increases to very high values in the vicinity of the Curie point for long-range order, and continues to rise, as the temperature is increased into the disordered region, until values are obtained which are outside the range that the apparatus can measure (about $\phi = 0{\cdot}1$). In addition to the fact that the degree of order at zero stress depends upon temperature, it has also been shown[5] that in the immediate proximity of the Curie point Δ_E becomes very large, since in this range the ordering energy is a very rapidly varying function of the degree of order. It is, therefore, difficult to draw quantitative conclusions from the type of measurements obtained by KÖSTER, unless supplemented by other experiments.

In the case of CuAu, whose ordered structure is tetragonal ($c/a = 1{\cdot}1$), striking relaxation effects are to be expected, due to stress-induced changes in the degree of long-range order. The existence of such effects was predicted[94] and confirmed[99] by GORSKY, who carried out elastic after-effect measurements at constant temperature. (These experiments are analogous to measurements of internal friction as a function of frequency of vibration.[5])

INTERNAL FRICTION ORIGINATING IN INTERFACES

Stress Relaxation Across Incoherent Interfaces

Internal interfaces in a metal provide important sources of internal friction. A widely studied source of damping is the relaxation of shear stress across incoherent interfaces, such as large angle grain boundaries and interfaces between different phases. The grain boundary case has already been reviewed in this series;[100] the discussion here will, therefore, be brief and will mainly concern itself with the question of the *distribution* of interfaces. This point of view was recently reviewed by the writer.[101]

The concept of stress relaxation across an interface is readily described if we imagine a single interface of limited area located in a metal, a hypothetical case, to be sure. The application of a shear stress to such a specimen produces a state of uniform stress throughout the material, but this state will not remain indefinitely. Because of the inability of an incoherent interface to sustain a shear stress across it, *stress relaxation* will take place by a process of slipping across the interface until the stress there has gone to zero. Shear stress, therefore, becomes redistributed so that it is concentrated along the edges of the interface. This relaxed state is indicated schematically in Fig. 9, where lines of force are used to represent the stress field, such that the number of lines cutting across a unit area is proportional to the shear stress. Because of the

increased stress at the ends, the over-all shear strain in the material is increased as a result of this relaxation process. When stress is released, the existence of a reverse stress across the interface, results in the material being restored to its original state.

Such a material, therefore, behaves like a standard linear solid, where the parameter p, which is now the difference between the mean stress and that across the interface, approaches its equilibrium value according to a relaxation equation (equation (35)). It is readily shown[101] that the time of relaxation τ is proportional to the linear dimension of the interface, but the temperature dependence of τ (i.e. the activation energy H) is independent of the size of the interface if the unit mechanism of the process is everywhere the same. Fig. 9 shows that, in the relaxed condition, the shear stress undergoes a significant decrease in an

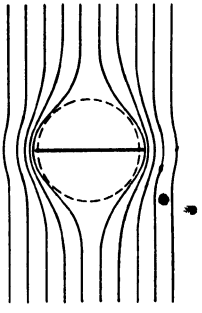


Fig. 9. Schematic representation of an incoherent interface across which relaxation of shear stress has occurred. The shear stress at any point is the number of lines per unit area. The cross section of the sphere of relaxation is represented by the dashed circle

approximately spherical region whose cross-section is outlined by the dashed circle. This region is called the *sphere of relaxation.*

In a well-annealed polycrystalline metal there is a regular collection of interfaces similar to the one just described, in which relaxation, or so-called "viscous slip," is held up at the edges and corners where the various boundaries meet. The spheres of relaxation will not overlap significantly, which means that each interface responds to stress independently of the others. The mechanical behaviour of the material is, therefore, a simple superposition of the contributions due to the individual interfaces. Since the linear dimensions of all interfaces are not the same, the internal friction peak is usually considerably broader (by a factor of 2 or 3) than the peak which corresponds to a single time of relaxation. This is the well-known "grain boundary peak" that has been studied by Kê[33, 102] and reviewed in the article by King and Chalmers.[100] The shift of this peak with frequency gives the heat of activation H for the unit mechanism responsible for the relaxation. A very large ΔM effect accompanies the grain boundary relaxation; the apparent modulus at low frequencies may be 30 per cent lower than the equivalent value for a single crystal.

On the high temperatures side of the grain boundary peak, the

internal friction, falls at first, but then begins to rise again. This effect has been attributed[28, 29] to slipping at the interface corners where the stress is concentrated after relaxation; the same phenomenon results in a permanent set when a static stress is applied. These effects were previously interpreted differently by Kê.[103]

Internal interfaces which are not allowed to grow at high temperatures to form a regular array, like the grain boundaries of a well annealed polycrystalline metal, may show more complicated effects. If the array of interfaces is irregular, or somewhat haphazard, so that large overlapping of the spheres of relaxation occurs, the behaviour of the collection of interfaces can no longer be regarded as a simple superposition of the behaviour of the individual interfaces, but instead we obtain an example of *coupled relaxations* (page 21). In this case, when one interface relaxes it produces a significant increase in the shear stress across a second, and vice versa, so that very large non-elastic strains may be produced. Such a description was used to explain the large internal friction obtained after discontinuous precipitation in an aluminium-zinc alloy.[25] Instead of the usual broadened peak, the internal friction seems to climb to exceedingly high values at high temperatures, or, if temperature is held constant, at low frequencies. Such behaviour could still be described, phenomenologically, in terms of a superposition of relaxation times, but the range of times required would be very large and the values would have little physical significance. The fact that highly cold-worked materials show this type of behaviour, and the interpretation of this fact, will be discussed in the next section.

Although the general descriptive features of stress relaxation across incoherent interfaces may be given without knowledge of the atomic mechanism responsible for this behaviour, the explanation of the kinetics of the process must be closely tied in with the atomic mechanism. Thus, much attention has been given to values for the heat of activation for grain boundary stress relaxation. On the basis of very little evidence available at the time, Kê[104] suggested that the heat of activation for grain boundary relaxation may be the same as that for volume diffusion. This correlation is a surprising one, inasmuch as there seems to be no theoretical reason for it. Subsequent experimental evidence[28, 101, 105] has shown values of the activation energy for grain boundary relaxation both greater and smaller than that for volume diffusion. Apparently, the suspected correlation is not correct, but the mechanism of stress relaxation is still not understood. A mechanism suggested by Mott,[106] and discussed at length in the earlier article,[100] is disputed on the basis of evidence presented by the writer.[101] The effect of minute amounts of impurities on the activation energy for stress relaxation is not clear, as yet, although it is known that impurities may have a large effect on

both the position and height of the grain boundary internal friction peak.[44, 107] Further investigations of the heat of activation for stress relaxation across grain boundaries will most certainly contribute to our understanding of the nature of the grain boundary.

Movement of Twin Interfaces

When transformation from a cubic to a tetragonal structure occurs, heavy twinning is necessary in order to avoid high residual stresses in the material. Such twinned structures have been observed in a Mn-Cu (12 per cent Cu) alloy[108] as well as in In-Tl alloys.[109] An internal friction peak with maximum at about 0°C was observed by WORRELL for the Mn-Cu alloy. Inasmuch as crystallographic coherency exists across a twin interface, the explanation for the observed internal friction cannot be the same as for the viscous-like slipping of incoherent interfaces. The tetragonality of the lattice structure suggests that the application of shear stress will induce movement of a twin interface normal to itself, in such a way as to enlarge the region where the direction of tetragonality is favourable with respect to the applied stress relative to that in which it is unfavourable. The fact that an internal friction peak is observed, leads to the belief[5] that this movement of twin boundaries is a relaxation phenomenon. Further measurements have been made[110] for the Mn-Cu alloy, but internal friction has not yet been studied as a function of frequency. There is, therefore, still some uncertainty as to whether the internal friction actually originates in a relaxation process; if it does, the significance of the time of relaxation must be ascertained.

INTERNAL FRICTION RELATED TO PLASTIC PHENOMENA

The type of internal friction to be discussed here is related to the internal stresses, or dislocations, introduced by cold working and is commonly referred to as "cold work internal friction." The latter designation is not accurate inasmuch as effects are also observed due to the residual dislocations in well annealed materials.

In contrast to our relatively recent awareness of the other sources of internal friction discussed in this article, the study of damping of plastic origin has a long history. As pointed out by GUILLET,[111] even the earliest bell makers were cognizant of the fact that brittle materials gave the lowest damping. In recent years, considerable effort has been directed toward the study of this cold work internal friction. Yet, in spite of the effort expended, our understanding of the nature of this type of damping is in a very elementary stage. The reasons for such slow progress are many. The effects are very complex and a considerable number of variables are involved, among them, the degree of cold work and the type of cold work given the specimen material, the time at room temperature before measurement, the frequency of vibration,

and the impurity content of the specimen. Much of the work in the field is not fundamental in character but is concerned only with the measurement of internal friction, at one temperature and frequency, as a function of some alloying element or a particular heat treatment. Often no attempt is made to see if other sources of internal friction may also be present to complicate the results. As a consequence, it is usually difficult to compare data by different investigators, which often contradict each other.

On the brighter side, there have been a sufficient number of fundamental investigations to give us a collection of important facts. These basic facts will first be reviewed without any attempt to interpret them. Finally, a discussion in terms of dislocation theory will attempt to see to what extent the factual material can be explained.

Effect of Cold Working

The room temperature internal friction of a well annealed material usually has a very low value, of the order 10^{-5}, although this value is apparently very sensitive to the purity of the material. The structure sensitivity of this damping, especially of single crystals, is most striking; various investigators[9, 112, 113] report that slight jarring or almost any handling operation will produce a large increase in damping. In addition, two single crystals prepared of the same material and in exactly the same way differ in their internal friction values, sometimes by as much as a factor of ten.[114, 115] Similarly different parts of one crystal do not show the same internal friction.[12, 116]

The application of a small static stress to a well annealed material will usually raise its internal friction appreciably, even though no observable yielding has occurred. The duration of loading seems to be important since it seems possible[113] to produce a larger damping by dropping or jarring a specimen, than by equivalent static loads. This point bears further investigation, however. Systematic investigation of the effect of cold working on damping has been carried out using various systems of stressing: application of compressive loads, pulling in a tensile machine, drawing through dies. One cannot be certain that the same types of internal stresses are always produced by these various methods and, therefore, that the data obtained are always comparable. Several authors show that the room temperature internal friction as a function of cold work goes through a maximum. ZENER, CLARKE and SMITH[50] report such a maximum in polycrystalline brass at about 5 per cent elongation, for measurements at a frequency of 2000 c.p.s. The observed damping, even at the maximum, is very small, but this may be related, as we shall see, to the fact that about two weeks elapsed between the cold working and the measurements. LAWSON[113] reports very striking maxima for spectroscopically pure polycrystalline copper,

measured at 50 kc/sec. Correspondingly, a minimum is observed in the dynamic Young's modulus of the specimens ($\Delta E/E$ about 6 or 7 per cent) at the same amount of cold work that produces the maximum damping. The height of the maximum in the room temperature damping is greater the lower the temperature at which the deformation is carried out. HASIGUTI and HIRAI[115] obtain similar maxima for copper single crystals, measured at frequencies in the tens of kilocycles. The maxima are not identical for different specimens, but occur for 1–10 per cent elongation and for values $\phi \times 10^3$ in the range 3–10. For large amounts of cold work the internal friction apparently decreases to a constant value greater than the value for the annealed specimen. BOULANGER[117] also observes a maximum in the residual damping (after room temperature annealing) of a nickel-chrome alloy. Other investigators report no maximum, but a monotonic increase of damping with cold working of aluminium[118, 119] and iron,[120] although the rate of rise of the curves is considerably decreased for large amounts of cold work. In cases where simultaneous observations of the dynamic modulus are made, a decrease in the modulus is generally observed to parallel an increase in the internal friction. This decrease in modulus may be as large as several per cent.[121] As we shall see, only part of this decrease may be interpreted as a ΔM effect.

Although the behaviour of pure and impure materials as a function of degree of cold work seems to be about the same, qualitatively, the magnitude of the changes produced by cold working is greater for the purer materials.[118, 119]

Effect of Annealing

The conventional description of the effect of annealing a cold worked metal at successively increasing temperatures is that no significant internal changes occur until recrystallization begins, except for the relief of "macrostresses" responsible for stress-corrosion cracking.[122] This description is based on the fact that resistivity, x-ray line broadening, microstructure, hardness, and tensile strength all show essentially no change as a result of annealing below the recrystallization temperature. (Hardness and tensile strength may even increase slightly for low temperature anneals.) It has, therefore, been supposed that the internal stresses produced by plastic flow are not relieved before recrystallization occurs. Internal friction measurements show surprisingly different phenomena. The room temperature damping (usually at high frequencies) produced by cold working may recover completely to the value it had before cold working, after anneals at temperatures too low to produce any changes in the properties mentioned above, as first shown by FÖRSTER and KÖSTER[95] for brass and low-carbon steel. The prior recovery of internal friction has no influence on the tendency of the

material to recrystallize at sufficiently high temperatures and, conversely, the room temperature damping is little affected by the onset of recrystallization, if this damping has already recovered during a low temperature anneal. More extensive investigations of the annealing of cold worked α-brass are reported by KÖSTER and ROSENTHAL.[121] It is shown that sizeable recovery of internal friction occurs at room temperature in tens of hours. Measurements of the damping as a function of time show that the rate of recovery is greatest at the shortest times. Since measurements in these experiments were started one-half hour after cold working, initial values of the damping may have been considerably higher than any of the reported values. Very rapid measurements that may be taken immediately after plastic deformation, are therefore desirable in annealing studies. The recovery of internal friction at room temperature occurs in about the same time for specimens given prior extensions of from 10 to 60 per cent, except that the starting values are higher for the greater elongation. The work of KÖSTER and ROSENTHAL indicates that the recovery of room-temperature internal friction in cold worked brass takes place in a similar manner at various annealing temperatures, the time of recovery decreasing from tens of hours at room temperature to minutes or seconds at 200°C. Similar measurements on cold worked brass are reported by FUSFELD,[123] who shows that the effective heat of activation for the recovery process is substantially lower than 10,000 calories per mole. Observations of the effect of annealing on internal friction in aluminium[119] and iron[120] are reported by KÖSTER. The results for aluminium seem to stand alone in that they show incomplete recovery below the recrystallization temperature. The case of α-iron is very striking; here the recrystallization temperature is at 500–600°C, yet damping recovers completely at room temperature in ten hours, and at 100°C in minutes. Similar effects are also observed in copper.[115, 42]

The ability of cold worked materials to anneal at room temperature, as far as internal friction measurements are concerned, may, in part, explain some of the discrepancies in the effect of cold work on internal friction as reported by various investigators. The results of BOULANGER[117] show, for example, that whether or not a maximum is obtained in the curve of damping versus degree of cold work, depends on the time that has elapsed between the cold working and the measurements.

The distinct difference between the effect of annealing on internal friction and on most other properties leads to the almost unavoidable conclusion that there are at least two distinct types of internal stress: one responsible for hardness and gross plastic deformation, which is required in order that recrystallization may occur, and the second responsible for internal friction effects.

The dynamic modulus is apparently somewhere between hardness and internal friction in its reaction toward annealing. A small part of the large decrease in modulus, produced by large amounts of cold work, recovers together with the internal friction; the remainder is not affected by anneals below the recrystallization temperature. This residual effect is clearly observed,[121] either after annealing for a long time at room temperature, or after successive anneals at increasing temperatures have completely eliminated the cold work internal friction. The fact that the recoverable part of the decrease in dynamic modulus is eliminated in the same time as the internal friction, is strong evidence that it represents a ΔM effect, i.e. it results from an additional in-phase component of strain produced by the same non-elastic mechanism as that responsible for the damping. The part of the decrease in dynamic modulus which is not recoverable below the recrystallization temperature is apparently a lowering of the modulus M' (equation (9)) that relates the stress to the elastic strain. That the presence of internal strains will produce such a lowering of the elastic modulus, was demonstrated by ZENER.[124] One should also keep in mind the fact that large amounts of cold work may also change the modulus by producing preferential orientation (texture) in a polycrystalline material.

Strongly Amplitude Dependent Damping

Although amplitude dependence of internal friction is observed[125, 126] for vibration at large strain amplitudes, it had been believed[1] that, so long as the maximum strain amplitude in a specimen is kept below 10^{-5} or 10^{-6}, damping is independent of amplitude. That this belief is not always correct is strikingly demonstrated by the experiments of READ,[9, 10] which show that the internal friction of annealed and slightly cold worked single crystals of zinc and copper is strongly amplitude dependent, even at strain amplitudes of 10^{-7}; for the case of zinc the damping may increase by almost one hundred times from low amplitudes up to 10^{-6}. These experiments, which use the piezoelectric method, also show that, associated with the amplitude dependent damping, there is an amplitude dependent ΔM effect, as measured by the decrease in resonant frequency with increasing amplitude of vibration. The resonant frequency in these experiments is in the tens of kilocycles. It is shown that the curve of damping versus strain amplitude for zinc crystals is not reversible, i.e. different values are obtained for decreasing amplitudes than those for increasing amplitudes. These results are interpreted as evidence that the measurements themselves produce internal changes in the crystals. Further experiments on these effects in zinc crystals are reported by several investigators.[127, 114, 49] The experiments on copper crystals show considerably simpler effects,

probably because room temperature represents a much lower "reduced temperature" for copper than for zinc. (Copper has a much higher melting temperature and recrystallization temperature than zinc.) The damping-amplitude curves for copper crystals are not affected by the measurements themselves and, therefore, represent a measure of the mechanical state of the material. Further experiments on copper crystals are those of MARX and KOEHLER[21] and of the writer.[11,12] Fig. 10, taken from reference 12, shows the effect of successive applications of a static compressive stress on the curves of internal friction versus strain amplitude, for a copper crystal. The stresses were sufficiently small so that no observable change of length of the crystal occurred as a result of stressing. Similar data were also given by READ.[10] It should be noted that these curves start in a parabolic fashion, so that the damping below a strain amplitude of 10^{-7} is practically independent of amplitude. One could then ignore the amplitude dependent part of the damping and concentrate only on the internal friction at very low amplitudes. On the other hand, the strong amplitude dependence seems to be the most characteristic feature of this type of damping. Furthermore, it is possible to measure very precisely the *changes* in resonant frequency with strain amplitude, so that the amplitude dependence of both the ΔM effect and the damping may be obtained simultaneously. (Changes in resonant frequency with amplitude can be measured to better than 1 part in 50,000.) For these reasons, the work of the writer concentrates on a quantitative study of the amplitude dependent internal friction and ΔM effect. The experimental results show that the ratio of the amplitude dependent part of damping to that of the ΔM effect is independent of frequency and of the temperature; this fact could not be explained in terms of any mechan-

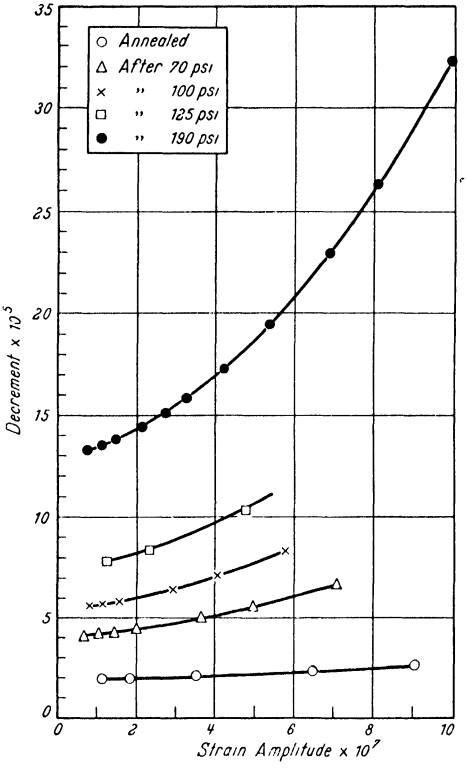


Fig. 10. Effect of successive applications of static compressive stress on the amplitude-dependent internal friction in a single crystal of copper

ism of viscoelastic behaviour. On the other hand, the experimental values of this ratio are in quantitative agreement with the predictions of the theory of static hysteresis as given by equations (49). The exceptional amplitude dependence of this type of internal friction might, in itself, have led us immediately to suspect a mechanism of static hysteresis. This conclusion is suggested by the discussion (in the earlier section on static hysteresis) which shows that such a mechanism inherently leads to strong amplitude dependence because of the non-linearity of the static hysteresis loops, in contrast to viscoelasticity, which is essentially a linear phenomenon at low amplitudes.

The reader may wonder why the phenomenological mechanism might not have been decided with certainty by a simple study of the frequency dependence of internal friction. Such measurements do not provide decisive results in the present case, because of the extreme structure sensitivity of the property we are measuring. As long as one observes that two crystals, grown and treated identically, or that the two halves of one crystal,[116] differ considerably in values of internal friction, it becomes difficult to measure any frequency effects that are not masked by such variations. Even measurements at a series of natural frequencies (overtones) of one given specimen do not represent the properties of the same part of the lattice, since the maximum stress amplitude occurs in different parts for different harmonics. Observed variations with frequency are, therefore, different for different specimens, but the scatter seems to fall about a frequency independence in the case of the copper single crystals.[11] For more highly cold worked polycrystalline materials, greater homogeneity seems to exist, and internal friction is reported[50] to be independent of frequency.

The temperature dependence of the curves of internal friction versus amplitude has also been studied for single crystals of copper in a limited temperature range.[11] The internal friction and ΔM effect increase with temperature in this range of measurement, for which a tentative explanation has been advanced.[128] Other observers report[129, 130] different results for the temperature dependence. Further experimentation along these lines will undoubtedly be fruitful.

Surface conditions seem to have an effect on internal friction, but it is reported[127, 49, 11] that etching the surface results in a lowering of the internal friction curve as a whole without affecting the amplitude dependent part.

The demonstration that amplitude dependent damping in single crystals at high frequencies originates in a mechanism of static hysteresis, and the close relation between this mechanism and strong amplitude dependence, means that it is always desirable to know whether damping is or is not strongly amplitude dependent. Unfortunately, many of the publications give no clue regarding this matter, either because the

apparatus used involved weak electromechanical coupling, so that all strain amplitudes are below 10^{-7} where no amplitude dependence is observed, or because the experimenter did not check on the possibility of amplitude dependence.

Study of a Heavily Cold Worked Material

One investigation by Kê[118] and Zener[131] deserves special attention, since it can be interpreted in terms of a mechanism which differs greatly from static hysteresis. The experiments differ in two ways from the single crystal investigations. First, the measurements are at low frequencies, i.e. obtained by means of a torsion pendulum. Second, the specimens (pure and commercial aluminium) are severely cold worked, by drawing them through dies to a 95 per cent reduction in area. The internal friction of these heavily cold worked specimens is independent of amplitude up to strain amplitudes of 10^{-5}. It originates in anelastic behaviour, since creep produced by static stress is completely recoverable; furthermore, it varies with frequency and temperature as a function of the product $\omega e^{H/RT}$, with H equal to 31,000 cal/mole. The curve of internal friction versus frequency does not show a maximum, but climbs to very high values at low frequencies. This behaviour is completely comparable to that to be expected in the case of coupled relaxations of an irregular network of interfaces, discussed in the previous section. In the present case the interfaces are slip bands, which acquire a viscous behaviour due to the interlocking of the dislocations present in high concentrations in these bands.[131] The irregular distribution of the bands results in a coupling phenomenon, as discussed earlier. The interpretation in terms of a network of slip bands is in agreement with the observations of Wood[132] that a heavily cold worked material is fragmented into crystallites. We may regard the slip bands as the crystallite boundaries.

Discussion in Terms of Dislocation Theory

Most of the important facts concerning the internal friction of plastic origin have been discussed, and the reader, reflecting upon them, will undoubtedly agree that many careful experiments remain to be performed before a detailed understanding of the atomic mechanism of this damping will be possible. An attempt at interpretation of the measurements in terms of dislocation theory must, in the present stage, be somewhat speculative. On the other hand, new experiments are generally the outgrowth of speculation.

In two investigations progress has been made toward an understanding of the mechanism of damping, at least from a phenomenological viewpoint. These are the investigations of amplitude dependent

damping in single crystals of copper at high frequencies, and the study of heavily cold worked aluminium at low frequencies. Let us start with the latter. Here a satisfactory explanation may be given in terms of slip bands or crystallite interfaces. The behaviour of regions that are very dense in dislocations is not conveniently discussed in terms of dislocation theory, for the same reason that it is not profitable to discuss large angle grain boundaries as made up of dislocations: the identity of the individual dislocations is lost. The study of heavily cold worked material seems to fit best into the category of damping resulting from stress relaxation across incoherent interfaces. The fact that the observed heat of activation for the damping in heavily cold worked aluminium is the same as that for grain boundary relaxation in that material, within experimental error, is further evidence that the two effects are basically the same. The grain boundaries, of course, represent a more regular collection of interfaces than the slip bands of a heavily cold worked metal, and, therefore, produce an internal friction peak. Annealing at low temperatures of the heavily cold worked material may result in lowered damping due to the production of greater regularity in the crystallite boundaries without appreciable growth of crystallites (which would lead to recrystallization).

Now let us turn to the behaviour of slightly cold worked materials, where the properties of individual dislocations may be more readily discussed. As already pointed out, we must conclude, from the easy room temperature recovery of cold work internal friction, that the type of internal stress responsible for internal friction must differ from that which produces recrystallization and hardness, in that it is less stable or more mobile. Translated into dislocation language we may say that, where as dislocations tightly bound in internal interfaces are responsible for gross plastic flow, the internal friction in a material that is not too heavily cold worked is associated with relatively free dislocations. We may imagine that during low temperature annealing these free dislocations are not lost from the lattice but simply become bound, either to impurity atoms or to other dislocations (presumably in internal surfaces). The case with which recovery takes place, as indicated by its occurrence at room temperature and by the fact that the heat of activation for the process is small[123] relative to corresponding values for volume diffusion, shows that the recovery process does not involve atomic diffusion. If recovery involves the locking of dislocations by impurity atoms, it is the dislocations that move to the impurity atoms, rather than vice versa. The occurrence of a mechanism of static hysteresis in high frequency measurements is now readily interpreted. The static hysteresis loop is a consequence of a certain number of relatively free dislocations being torn loose from the small barriers that hinder their motion, for each increment of stress $d\sigma$. The stress-strain loop is then really a

step-like affair, each step corresponding to the burst of a single dislocation, which moves almost instantaneously from the original barrier to another position where its motion is again interrupted. Each such dislocation burst produces non-elastic strain. Once a dislocation is pulled loose it will not generally return to its original position unless a reverse stress is applied. The magnitude of an individual step that constitutes the stress-strain loop is a small fraction of an atomic distance, too small to be observed as a step. The stress-strain loop that has been described is analogous to the magnetic hysteresis loop, in which sudden movements of magnetic domain boundaries are induced by an applied magnetic field in the same way as dislocation bursts are induced by an applied stress.

A qualitative estimate of the dependence of internal friction on the degree of cold work may be given in terms of the present interpretation. The damping is proportional to the energy loss per cycle ΔW (equation (27)), and the contribution to this energy loss, in that part of the cycle where the stress is increased by an amount $d\sigma$, may be represented as $d\Delta W$. This latter quantity is proportional to the increase in non-elastic strain corresponding to the stress increment, which in turn is proportional to the product of the number of dislocations moved dN and the mean distance moved L. In the form of an equation,

$$d\Delta W \approx L \,.\, dN$$

It now seems reasonable, that in successive stages of cold working starting from a well-annealed material, the major increase is, at first, in the term dN, while L remains essentially constant; more free dislocations are made available, but the distance that each may move in one burst is fixed by imperfections already present in the lattice. As cold working progresses, the crowding of dislocations soon begins to place a serious limit on L, which begins to decrease rapidly when the density of dislocations is large enough so that they begin to hinder each other's motion. The quantity dN cannot become indefinitely large to counteract this decrease in L, since the tendency for relatively free dislocations to become "bound dislocations" will limit the number of free dislocations available. This argument predicts a maximum in the product $L \,.\, dN$, and, therefore, in the internal friction, as a function of the degree of cold work. The concept of dislocation bursts has, therefore, provided a reasonable explanation for the observed maxima in the curves of damping versus cold work.

Amplitude dependent internal friction has only been studied, and its mechanism established as one of static hysteresis, at high frequencies. No systematic low frequency measurements for well annealed materials and for small amounts of cold work have been obtained. (This lack of data is partly a result of the fact that low frequency measurements of

small values of internal friction are difficult, because the equipment itself has thus far been responsible for a residual damping of about 10^{-4}.) If the dislocation mechanism just described is correct, however, we may anticipate that, given enough time (i.e. at sufficiently low frequencies) or at higher temperatures of measurement, it should be possible for a relatively free dislocation to pull away from its barrier by thermal activation, without having to wait for the applied stress to become sufficiently large to free it. In other words, the mechanical behaviour which is describable in terms of static hysteresis at high frequencies may completely change its character to a viscoelastic type of mechanism at low frequencies. One model that may be presented is that, at low frequencies, instead of dislocations being torn loose from impurity atoms that restrain them, they may drag the impurity atoms along with them. This behaviour would lead to a viscous creep involving a permanent set, since dislocations that have moved have no reason to return to their original positions unless a reverse stress is applied. The low frequency internal friction may be considerably less amplitude dependent than the high frequency damping, and it may show a strong frequency dependence, if this interpretation is correct. In fact, the simplest description of such behaviour would be the Maxwell solid, which represents elasticity plus viscous flow, and leads to an inverse frequency dependence (equation (12)). Unfortunately there are no low frequency damping measurements which can be examined to see if these ideas are correct. There are static measurements, however, which agree very well with the above interpretation. These are the measurements of "microcreep" of single crystals of tin, by CHALMERS,[133] which show that for very small stresses a limited amount of creep occurs. The observed initial creep rate is proportional to the applied stress, and the rate of creep, for any given stress, approaches zero with time. This microcreep is not recoverable in single crystals. It has been attributed[74, 134] to the movement of relatively free dislocations which drag impurity atoms along with them. On the basis of this evidence it therefore seems logical to suppose that internal friction at high frequencies and microcreep may be traced to the same origin: the movement of relatively free dislocations which are torn loose by the stress when the frequency is high, or which surmount the barrier to their movement by an activated process at low frequencies or under static conditions. Although this description provides a semblance of order in the interpretation of phenomena associated with the movement of dislocations at low stress levels and for small degrees of cold work, it should be remembered that there is a large unfilled gap between the high frequency measurements and the static "microcreep" measurements. Furthermore, the former has been studied only in the case of copper crystals and the latter effect only in the case of tin crystals.

Evidently, further experiments are required to see if the predictions are substantiated.

Thermoelastic Internal Friction

The sources of internal friction considered up to this point have all involved some form of atomic re-arrangement, induced by stress. In the present section, damping originating in stress-induced thermal variations will be considered. The basic principles involved in this type of internal friction have been understood for some time and have been adequately reviewed.[1, 5] The present discussion will briefly summarize these principles.

By the thermoelastic effect, we mean the coupling between internal (thermal) and mechanical co-ordinates of a material, as illustrated by the expansion of a solid when heated, and the reciprocal effect, the cooling during an adiabatic expansion. An element of material, to which a small stress is suddenly applied, will instantaneously be strained by an amount ε', and this strain will be accompanied by a change in temperature. If the stress is homogeneous throughout the specimen, the same temperature change will occur at every point. On the other hand, a *non-homogeneous stress* sets up a temperature gradient in the material, and, therefore, allows heat flow to take place. The flow of heat into or out of a given element of the material produces an additional non-elastic strain. In general, a conversion of mechanical energy into heat, or internal friction, will be a result of this heat flow and the associated non-elastic behaviour.

If a periodically varying non-homogeneous stress is applied to the material, a periodically varying temperature gradient is produced. If the variations in the stress are so rapid, i.e. the frequency so high, that insufficient time is allowed for appreciable heat flow to take place during a stress cycle, the process is *adiabatic* and no energy loss, or damping, takes place. Under such conditions the dynamic modulus is the modulus M', the quotient of the stress by the instantaneous strain ε' (equation (9)). The quantity M' is, in this case, known as the adiabatic modulus. In the other extreme, at sufficiently low frequencies, temperature equilibrium in the specimen is constantly maintained, or the process is *isothermal*. The transfer of mechanical energy into heat and vice versa takes place reversibly, so that again there is no mechanical energy lost (or net heat generated) per cycle. The dynamic modulus measured under these conditions is called the isothermal modulus and is lower than the adiabatic modulus. In the intermediate range of frequencies, where the conditions of application of stress are neither adiabatic nor isothermal, the conversion of mechanical energy into heat is not reversible and internal friction effects may be observed. There are two categories of thermoelastic effects, depending on the origin of the non-homogeneity of stress in the specimen.

Macroscopic Thermal Currents

The thermoelastic effect may result from non-homogeneity of stress that is inherent in the type of vibration. This phenomenon is demonstrated most strikingly for the case of the transverse vibration (flexure) of reeds. In this mode of vibration the material that is on the outer or convex side at any instant is expanded and, therefore, cooled, while that on the concave side is compressed and raised in temperature. Heat flow will, therefore, occur across the specimen, if sufficient time is allowed. As already pointed out, the thermoelastic internal friction shows a maximum as a function of frequency, since very small damping occurs at either very low frequencies (the isothermal limit) or at very high frequencies (the adiabatic limit). It can be shown[5] that the case of a reed in transverse vibration is describable, very nearly, in terms of a single time of relaxation (i.e. as a standard linear solid), so that the internal friction peak is described by equation (39). The time of relaxation τ is given by

$$\tau = d^2/\pi^2 D_t \qquad \ldots . \quad (63)$$

where d is the thickness of the reed and D_t the thermal diffusion constant, which is

$$D_t = \frac{\text{thermal conductivity}}{(\text{specific heat})\ (\text{density})}$$

The relaxation strength Δ_E is the fractional difference between the adiabatic and isothermal moduli; it ranges between 10^{-2} and 10^{-4} for various materials. The measurements of BENNEWITZ and RÖTGER[135] and of ZENER, OTIS and NUCKOLLS[136] have verified the theory in all details.

Macroscopic thermal currents may also occur, at least in principle, in a specimen vibrating in a longitudinal mode, where heat flow is to be expected between regions of compression and of dilatation. The distances between these regions, and, therefore, the time of relaxation involved, is too large for the effect to be of any significance when specimens are vibrated in the ordinary range of frequencies. We may, therefore, always regard longitudinal vibrations as adiabatic with respect to macroscopic thermal currents.

No thermoelastic effect results from a non-homogeneous *shear* stress, as in torsional vibration, because shearing is not accompanied by a change of temperature.

Microscopic Thermal Currents

In addition to the possibility that non-homogeneity of stress may result from the type of vibration employed, a polycrystalline metal also offers the possibility of fluctuations in stress from grain to grain, even

when the applied stress is macroscopically homogeneous.[137] This microscopic non-homogeneity is a consequence of the elastic anisotropy of an individual grain, combined with the fact that grains are randomly oriented in a polycrystalline metal. Fluctuations in stress from one grain to another gives rise to heat currents between adjacent grains, and, therefore, to an internal friction peak analogous to that obtained as a result of macroscopic thermal currents. The peak obtained as a result of intergranular thermal currents does not represent a single time of relaxation, however, and the exact form of this curve is not readily evaluated theoretically. We may expect from dimensional reasoning,[137] that internal friction will be a function of the frequency, the mean grain size d_g, and the thermal diffusion constant D_t, only through the product $\omega d_g/D_t$. The maximum in the internal friction may be expected when this product is close to unity. These predictions are verified by the experiments of ZENER and RANDALL.[20, 41]

An awareness of the internal friction that originates in thermal currents is important when one studies the damping due to other sources, if polycrystalline materials are being used, or if the specimen is vibrated transversely. Under either of these circumstances, it is desirable to work in a frequency range where the vibrations are either isothermal or adiabatic with respect to thermal currents, to avoid this type of internal friction.

MAGNETOELASTIC INTERNAL FRICTION

The coupling between magnetic and mechanical properties of a ferromagnetic material provides another important source of internal friction. This "magnetic damping" is closely related to other ferromagnetic phenomena and has, therefore, been reviewed in some detail in the general treatises on ferromagnetism by BECKER and DÖRING[138] and by BOZORTH.[139] The present discussion is a brief résumé of the factors that are responsible for internal friction in ferromagnetic materials.

Magnetoelastic phenomena are best discussed in terms of the domain theory of ferromagnetism, according to which a ferromagnetic substance is composed of small "domains" that are individually magnetized to saturation. In the demagnetized state of a specimen, there is a random distribution of the magnetization vectors of the domains among the various crystallographic directions of "easy magnetization". The effect of a magnetic field is to produce rotation of the directions of magnetization and movements of domain boundaries so as to align the magnetization vectors as closely as possible to the direction of the field.

The coupling between magnetic and mechanical properties is provided by the phenomenon of magnetostriction, which usually manifests itself through the elongation of a material in the direction of magnetization. (The opposite effect, a decrease in length of a magnetized specimen

represents "negative magnetostriction".) From the appropriate reciprocity relation, it follows that application of stress may produce a change in the state of magnetization. The effect of stress is, therefore, somewhat analogous to that of a magnetic field, in that stress can also produce rotation of the magnetization vectors of domains and movement of the domain boundaries. The principal difference between the action of stress and of a magnetic field is that in a field there is a difference in the energy associated with a magnetization parallel and antiparallel to the field, while stress is incapable of distinguishing between two directions of magnetization 180° apart.

As a result of magnetoelastic coupling, the effect of stress is to produce, in addition to the elastic strain, a magnetostrictive strain. The occurrence of this non-elastic strain leads to both internal friction and to what is called, in the present case, the ΔE effect. At high fields, sufficient to align all domains in the direction of the field, the magnetostrictive strain vanishes. The value of Young's modulus measured under these conditions, i.e. when the material is magnetized to saturation, is simply the ratio of stress to elastic strain ε' and may be designated in the usual notation (see equation (9)) as E'. The difference between E' and the value of Young's modulus measured at zero magnetization is the well-known "ΔE effect." The concept of a ΔM effect, which appears throughout this article, is a natural generalization of this ΔE effect in magnetic materials. We will see shortly that Young's modulus of a ferromagnetic specimen depends not only on the field strength but also on the vibration amplitude.

The work of several authors[138, 140, 141] has led to the separation of sources of internal friction in magnetic materials into three types: macro-eddy currents, micro-eddy currents, and static hysteresis. These will now be discussed individually.

Macro-eddy Currents

The sudden application of a tensile stress to a partially magnetized specimen will produce a change in the magnetization. Such a change induces surface eddy currents which produce an additional field inside the specimen such that the total flux inside is unchanged. Associated with these currents are resistive energy losses. The surface eddy currents gradually diffuse inwards thereby allowing the magnetic field strength in the specimen to return to its original value. This diffusion of eddy currents, and, therefore, of magnetic flux, may be regarded[5] as analogous to the thermal diffusion which produces thermoelastic internal friction. It results in a peak in the curve of internal friction versus frequency; this peak is not expressible as a single time of relaxation, however. The damping depends on the thickness of the specimen d, the resistivity ρ, and the angular frequency ω only through the quantity

$d^2\omega/\rho$. In the low frequency limit, damping is proportional to this quantity, and in the high frequency limit, the damping varies inversely as the square root of this quantity. This type of internal friction is a typical example of anelasticity, and shows the usual independence of vibration amplitude. There is no damping due to macro-eddy currents in a demagnetized specimen, due to the fact the application of stress to such a specimen produces no gross magnetization. (This fact is a consequence of the inability of stress to distinguish between two directions 180° apart; if no preferential direction of magnetization exists at zero stress, a finite stress cannot create one.) These characteristics (the peak obtained as a function of frequency, the amplitude independence, and the vanishing of the effect in a demagnetized specimen) distinguish the internal friction due to macro-eddy currents and make it possible to separate this type of damping from the others.

Micro-eddy Currents

Although the application of stress to a demagnetized specimen produces no gross magnetization, the fact that there are local variations of magnetization within the material, due to the domain structure, means that stress will produce local changes of magnetization. These local magnetization changes will add vectorially to zero, but the "micro-eddy currents" that accompany them will be associated with finite total energy losses. Like the macro-eddy current case, the damping that results from micro currents is independent of vibration amplitude and strongly frequency dependent. The two cases differ, however, in that micro-eddy currents produce losses even at zero magnetization, and that these losses are proportional to the frequency over a wide range of frequency, i.e. internal friction does not go through a peak. Furthermore, this source of damping is highly structure sensitive; according to a calculation of BECKER and DÖRING,[138] internal friction varies inversely as the square of the magnitude of internal stresses. In qualitative agreement with this theory, various authors show[142, 143] that this type of damping is reduced by cold working and increased by annealing.

Magnetoelastic Static Hysteresis

When measurements are taken at sufficiently low frequencies, eddy current losses are effectively eliminated, because the rate of application of stress and, therefore, the rate of change of magnetization is very small. Under these conditions we obtain a stress-strain hysteresis loop analogous to the hysteresis loop of magnetization versus field strength; in both cases the occurrence of hysteresis is associated with sudden movements of domain boundaries which produce, irreversibly, a change in the size of one domain at the expense of another. This type of stress-strain loop is another example of the behaviour that we have termed

static hysteresis. Characteristic of such behaviour is frequency independence, and a strong dependence of damping on strain amplitude even at low amplitudes of vibration. This type of magnetic damping is discussed in detail by BOULANGER.[32]

The similarity in the origin of the magnetization-field strength hysteresis loop to that of the stress-strain loop means that the Rayleigh law, which states that magnetic hysteresis losses are proportional to the cube of the maximum field strength, applies equally well to the mechanical hysteresis losses. The energy dissipated per cycle should, then, be

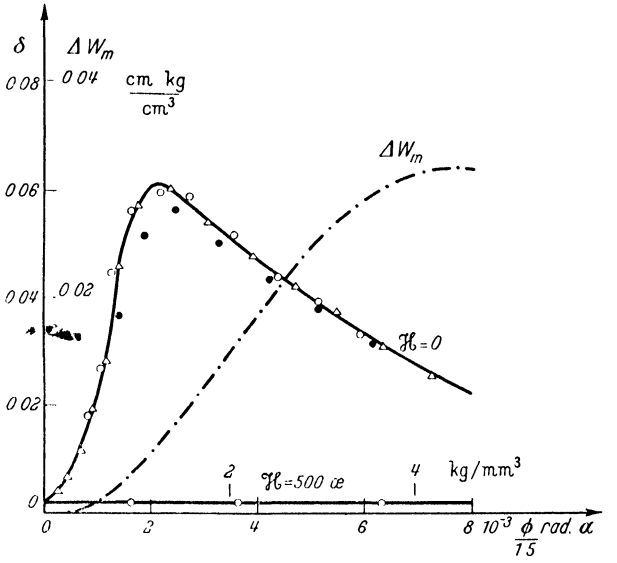


Fig. 11. Internal friction of Armco iron as a function of amplitude of vibration, at zero field and at saturation. Points indicated by different symbols represent different vibration frequencies

proportional to σ_1^3, where σ_1 is the stress amplitude of vibration. Since the vibrational energy is proportional to σ_1^2, we find (from equation (27)) that the internal friction is proportional to σ_1. This linear amplitude dependence of damping at low amplitudes has been verified by FÖRSTER and KÖSTER[144] and by SNOEK.[145] Correspondingly, it is shown[144a] that Young's modulus decreases linearly with σ_1. This result again shows that for a mechanism of static hysteresis, both the damping and ΔM effect show the same form of amplitude dependence, as expected from equations (49). As the vibration amplitude continues to increase, the energy dissipated per cycle must deviate from Rayleigh's law and, finally, approach a constant value for stress amplitudes sufficiently large to produce a condition of saturation. Correspondingly, the internal friction, which initially obeys a first power law must go

through a maximum as a function of amplitude and, finally, obey an inverse square law. The measurements of BOULANGER, shown in Fig. 11, demonstrate this behaviour. From the difference between the internal friction curves for zero field and for saturation, the energy loss ΔW_m, due to ferromagnetic effects, can be calculated (the dashed curve of Fig. 11). This curve is initially a cubic law and eventually approaches a constant value at large amplitudes. The fact that essentially the same internal friction curve is obtained at different frequencies guarantees that eddy current losses do not contribute to the observed damping.

OTHER SOURCES OF INTERNAL FRICTION

The sources of damping already discussed by no means exhaust all possibilities, but are simply the ones that have been most thoroughly investigated up to the present. The ability of stress to induce atomic re-arrangements of various kinds seems to offer wide opportunities for further application of internal friction methods, particularly in the study of phase transformations. Such applications have, thus far, been only of a preliminary nature.

An important type of phase transformation is the precipitation of a second phase, which occurs by a process of nucleation and growth. We have already discussed the effect of precipitation on internal friction when precipitation is discontinuous and results in the formation of internal interfaces. Studies of the precipitation of interstitial solutes in b.c.c. lattices, in terms of the rate of disappearance of solute atoms from solid solution, have also been discussed. No mention has been made, however, of any internal friction produced by the precipitate particles themselves, in the early stages of continuous precipitation. Such effects have been studied, for the most part, on age hardenable engineering materials; internal friction, as measured at one frequency and at room temperature, may then be regarded as another mechanical property (like hardness) which varies with the degree of precipitation. On the other hand, fundamental studies of such damping effects, in relatively simple alloys, may reveal valuable information about the nature of the early stages of precipitation, that cannot be obtained by other methods. An investigation by KÊ[146] reveals an "anomalous" internal friction in aluminium containing $\frac{1}{2}$ per cent copper, which he claims is related to the precipitation process and to the presence of a state of cold work. Further experiments by the writer on this alloy indicate that cold working may not be necessary to this effect, and also show that the "anomalous" damping is difficult to reproduce. This phenomenon seems to be a sensitive measure of one stage in the precipitation process.

Internal friction methods may also be of value in the case of phase transformations that do not occur by nucleation and growth. If the

transformation is accompanied by a change of dimensions, the application of stress to the material will induce or retard the transformation at temperatures in the immediate vicinity of the transformation temperature. Stated differently, the lattice is relatively unstable near the transformation temperature and may be greatly influenced by an external stress. Large non-elastic deformation, or plasticity, and large internal friction effects are, therefore, expected in the vicinity of the transformation temperature. Such behaviour has been observed for the β to ζ transition in Ag-Zn alloys,[97] for the λ-point transition in ammonium chloride,[147] and for the α to γ transformation in ferronickels.[148, 149] The mechanism of this type of internal friction is, of course, closely related to the mechanism of the particular phase transformation, and, therefore, deserves further detailed investigation.

References

[1] Zener, C. *Proc. Phys. Soc.* 52 (1940) 152
[2] Seitz, F. *The Physics of Metals* New York, 1943 Chap. X
[3] —— and Read, T. A. *J. Appl. Phys.* 12 (1941) 538
[4] Zener, C. *Trans. Amer. Inst. Min. Met. Engrs.* 167 (1946) 155
[5] —— *Elasticity and Anelasticity of Metals* Chicago, 1948
[6] —— *ibid.* p. 138
[7] —— *Phys. Rev.* 60 (1941) 455
[8] Eshelby, J. D. *Proc. Roy. Soc. A* 197 (1949) 396
[9] Read, T. A. *Phys. Rev.* 58 (1940) 371
[10] —— *Trans. Amer. Inst. Min. Met. Engrs* 143 (1941) 30
[11] Nowick, A. S. *Phys. Rev.* 80 (1950) 249
[12] —— *Symposium on the Plastic Deformation of Crystalline Solids* Pittsburgh, 1950 p. 155
[13] Minorsky, N. *Non-Linear Mechanics* Edward Bros., Ann Arbor, Michigan 1947 Chap. XII
[14] Voigt, W. *Ann. Phys.* 47 (1892) 671
[15] Stokes, G. G. *Trans. Camb. Phil. Soc.* 7 (1845) 287; *Math. and Phys. Papers I* p. 75
[16] Joos, G. *Theoretical Physics* New York 1934, Chap. V
[17] Maxwell, J. C. *Phil. Mag.* (IV) 35 (1868) 134; *Scientific Papers II* p. 30
[18] Alfrey, T. *Mechanical Behavior of High Polymers* New York, 1948 Chap. B IV
[19] —— and Doty, P. M. *J. appl. Phys.* 16 (1945) 700
[20] Randall, R. H., Rose, F. C. and Zener, C. *Phys. Rev.* 56 (1939) 343
[21] Marx, J. and Koehler, J. S. *Symposium on the Plastic Deformation of Crystalline Solids* Pittsburgh 1950 p. 171
[22] Zener, C. *J. appl. Phys.* 18 (1947) 1022
[23] Rayleigh, Lord *The Theory of Sound* London, 1894
[24] Alfrey, T. *Quar. appl. Math.* 2 (1944) 113
[25] Nowick, A. S. *J. appl. Phys.* 22 (1951) 925
[26] Eyring, H. and Halsey, G. *High Polymer Physics, A Symposium* New York, 1948 p. 61
[27] Alfrey, T. *Mechanical Behavior of High Polymers* New York, 1948 Chap. B III

[28] PEARSON, S. *Royal Aircraft Establishment* Report No. Met. 67, Farnborough, 1951
[29] ——, GREENOUGH, G. B. and SMITH, A. D. N. *Nature, Lond.* 167 (1951) 1023
[30] BURGERS, J. M. *First Report on Viscosity and Plasticity* Amsterdam, 1935 p. 19
[31] CHEVENARD, P. *Rev. Met.* 39 (1942) 65
[32] BOULANGER, Ch. *ibid.* 46 (1949) 321
[33] KÊ, T. S. *Phys. Rev.* 71 (1947) 533
[34] FINE, M. E. *J. Metals* 2 (1950) 1322
[35] NORTON, J. *Rev. Sci. Inst.* 10 (1939) 77
[36] SIEGEL, S. *J. Acoust. Soc. Amer.* 16 (1944) 26
[37] JACOBS, R. B. and BANCROFT, D. *Rev. Sci. Inst.* 9 (1938) 279
[38] CABARAT, R. *Rev. Met.* 46 (1949) 617
[39] BORDONI, P. G. *Nuovo Cim.* 4 (1947) 177
[40] SACK, H. S., MOTZ, J., RAUB, H. L. and WORK, R. W. *J. appl. Phys.* 18 (1947) 450
[41] ZENER, C. and RANDALL, R. H. *Trans. Amer. Inst. Min. Met. Engrs.* 137 (1940) 41
[42] WEGEL, R. L. and WALTHER, H. *Physics* 6 (1935) 141
[43] ZENER, C., VAN WINKLE, D. and NIELSEN, H. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 98
[44] KÊ, T. S. *J. appl. Phys.* 20 (1949) 1226
[45] FÖRSTER, F. *Z. Metallk.* 29 (1937) 109
[46] QUIMBY, S. L. *Phys. Rev.* 25 (1925) 558
[47] SIEGEL, S. and QUIMBY, S. L. *ibid.* 49 (1936) 663
[48] COOKE, W. T. *ibid.* 50 (1936) 1158
[49] WERT, C. A. *J. appl. Phys.* 20 (1949) 29
[50] ZENER, C., CLARKE, H. and SMITH, C. S. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 90
[51] QUIMBY, S. L. *Phys. Rev.* 39 (1932) 345
[52] FÖRSTER, F. and BREITFELD, H. *Z. Metallk.* 30 (1938) 343
[53] FUSFELD, H. I. *Rev. Sci. Instr.* 21 (1950) 612
[54] SMITH, A. D. N. *J. Sci. Instr.* 28 (1951) 106
[55] HUNTINGTON, H. B. *Phys. Rev.* 72 (1947) 321
[56] LAZARUS, D. *ibid.* 76 (1949) 545
[57] LIPSON, H. *Progress in Metal Physics II* Chap. I, London, 1950
[58] CADY, W. D. *Piezoelectricity* New York, 1946
[59] WOODRUFF, E. *Phys. Rev.* 16 (1903) 321
[60] SNOEK, J. *Physica* 6 (1939) 591
[61] —— *ibid.* 8 (1941) 711
[62] Reference 5, p. 119
[63] DIJKSTRA, L. J. *Philips Research Reports* 2 (1947) 357
[64] KÊ, T. S. *Phys. Rev.* 74 (1948) 9
[65] —— *ibid.* 74 (1948) 914
[66] WERT, C. *J. appl. Phys.* 21 (1950) 1196
[67] FITZGERALD, J. V. *J. Amer. Ceramic Soc.* Oct. 1951
[68] DIJKSTRA, L. J. *J. Metals* 1 (1949) 252
[69] WERT, C. A. *J. appl. Phys.* 20 (1949) 943
[70] ZENER, C. *ibid.* 20 (1949) 950
[71] WERT, C. and ZENER, C. *ibid.* 21 (1950) 5
[72] WERT, C. *J. Metals* 2 (1950) 1242
[73] HARPER, S. *Phys. Rev.* 83 (1951) 709
[74] COTTRELL, A. H. *Rep. Conf. on Strength of Solids* Phys. Soc. London, 1948 p. 30
[75] —— and BILBY, B. A. *Proc. Phys. Soc.* 62 (1949) 19

[76] WERT, C. and ZENER, C. *Phys. Rev.* 76 (1949) 1169
[77] WERT, C. *ibid.* 79 (1950) 601
[78] STANLEY, J. K. *J. Metals* 1 (1949) 752
[79] ZENER, C. *Trans. Amer. Inst. Min. Met. Engrs.* 152 (1943) 122
[80] KÊ, T. S. *J. appl. Phys.* 19 (1948) 285
[81] ZENER, C. *Phys. Rev.* 71 (1947) 34
[82] NOWICK, A. S. *Phys. Rev.* 82 (1951) 340
[83] —— *Phys. Rev.*, in press
[84] SMIGELSKAS, A. D. and KIRKENDALL, E. O. *Trans. Amer. Inst. Min. Met. Engrs.* 171 (1947) 130
[85] DA SILVA, L. C. C. and MEHL, R. F. *J. Metals* 3 (1951) 155
[86] HUNTINGTON, H. B. *Atom Movements* Amer. Soc. for Metals, 1951 p. 69
[87] DARKEN, L. S. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 184
[88] LE CLAIRE, A. D. *Phil. Mag.* 42 (1951) 673
[89] ZENER, C. *J. appl. Phys.* 22 (1951) 372
[90] NOWICK, A. S. *ibid.* 22 (1951) 1182
[91] —— *Phys. Rev.* 82 (1951) 551
[92] NABARRO, F. R. N. *Rep. Conf. on Strength of Solids* Physical Soc. London, 1948 p. 75
[93] SEITZ, F. *Phys. Rev.* 79 (1950) 890, 1002
[94] GORSKY, W. S. *Phys. Z. Sow.* 6 (1934) 77
[95] FÖRSTER, F. and KÖSTER, W. *Z. Metallk.* 29 (1937) 116
[96] KÖSTER, W. *ibid.* 32 (1940) 145
[97] —— *ibid.* 32 (1940) 151
[98] —— and SCHNEIDER, A. *ibid.* 32 (1940) 156
[99] GORSKY, W. S. *Phys. Z. Sow.* 8 (1936) 562
[100] KING, R. and CHALMERS, B. *Progress in Metal Physics I* Chap. 3, London, 1949
[101] NOWICK, A. S. *Metal Interfaces* Amer. Soc. for Metals, 1952
[102] KÊ, T. S. *Phys. Rev.* 72 (1947) 41
[103] —— *J. appl. Phys.* 21 (1950) 414
[104] —— *Phys. Rev.* 73 (1948) 267
[105] ROTHERHAM, L., SMITH, A. D. N. and GREENOUGH, G. B. *J. Inst. Metals* 79 (1951) 439
[106] MOTT, N. F. *Proc. Phys. Soc.* 60 (1948) 391
[107] KÊ, T. S. *J. appl. Phys.* 20 (1949) 274
[108] WORRELL, F. T. *ibid.* 19 (1948) 929
[109] GUTTMAN, L. *J. Metals* 2 (1950) 1472
[110] SIEFERT, A. V. and WORRELL, F. T. *J. appl. Phys.* 22 (1951) 1257
[111] GUILLET, L. *Rep. Conf. on Strength of Solids* Phys. Soc. London, 1948 p. 116
[112] FOUND, G. H. *Trans. Amer. Inst. Min. Met. Engrs.* 161 (1945) 120
[113] LAWSON, A. W. *Phys. Rev.* 60 (1941) 330
[114] SWIFT, I. H. and RICHARDSON, J. E. *J. appl. Phys.* 18 (1947) 417
[115] HASIGUTI, R. R. and HIRAI, T. *ibid.* 22 (1951) 1084
[116] SACK, H. S. Discussion to reference 12
[117] BOULANGER, Ch. *C. R. Acad. Sci. Paris* 226 (1948) 1170
[118] KÊ, T. S. *J. Metals* 2 (1950) 575
[119] KÖSTER, W. *Z. Metallk.* 32 (1940) 282
[120] —— *Arch. Eisenhüttenw.* 14 (1940–41) 271
[121] —— and ROSENTHAL, K. *Z. Metallk.* 30 (1938) 345
[122] BRICK, R. M. and PHILLIPS, A. *Structure and Properties of Alloys* Chap. 4, New York, 1949
[123] FUSFELD, H. I. To be published
[124] ZENER, C. *Acta Cryst.* 2 (1949) 163

[125] NORTON, J. *Trans. Amer. Inst. Min. Met. Engrs.* 137 (1940) 49
[126] FÖPPL, O, *J. Iron Steel Inst.* 134 (1936) 393
[127] READ, T. A. and TYDNALL, E. P. T. *J. appl. Phys.* 17 (1946) 713
[128] FRANK, F. C. Discussion to reference 12.
[129] BORDONI, P. G. *Ric. Sci.* 19 (1949) 851
[130] PITTENGER, J. *Phys. Rev.* 83 (1951) 872
[131] KÊ, T. S. and ZENER, C. *Symposium on the Plastic Deformation of Crystalline Solids* Pittsburgh, 1950 p. 185
[132] WOOD. W. A. *Proc. Roy. Soc. A* 172 (1939) 231
[133] CHALMERS, B. *ibid.* 156 (1936) 427
[134] COTTRELL, A. H. and JASWON, M. A. *ibid.* 199 (1949) 104
[135] BENNEWITZ, K. and RÖTGER, H. *Phys. Z.* 37 (1936) 578; *Z. f. Tech. Phys.* 19 (1938) 521
[136] ZENER, C., OTIS, W., and NUCKOLLS, R. *Phys. Rev.* 53 (1938) 100
[137] ZENER, C. *ibid.* 53 (1938) 90
[138] BECKER, R. and DÖRING, W. *Ferromagnetismus* Berlin, 1939
[139] BOZORTH, R. M. *Ferromagnetism* New York, 1951
[140] BECKER, R. and KORNETSKI, M. *Z. Phys.* 88 (1934) 634
[141] KERSTEN, M. *Z. tech. Phys.* 15 (1934) 463
[142] WILLIAMS, H. J., BOZORTH, R. M. and CHRISTENSEN, H. *Phys. Rev.* 59 (1941) 1005
[143] KÖSTER, W. *Z. Metallk.* 35 (1943) 246
[144] FÖRSTER, F. and KÖSTER, W. *Naturwiss.* 25 (1937) 436
[145] SNOEK, J. *Physica* 8 (1941) 745
[146] KÊ, T. S. *Phys. Rev.* 78 (1950) 420
[147] LAWSON, A. W. *Phys. Rev.* 57 (1940) 417
[148] SCHEIL, E. and THIELE, W. *Arch. Eisenhüttenw.* 12 (1938) 103
[149] —— and REINACHER, G. *Z. Metallk.* 36 (1944) 63

2

THE MECHANISM OF OXIDATION OF METALS AND ALLOYS AT HIGH TEMPERATURES

*Karl Hauffe**

Introduction

THE reaction between a solid metal and oxygen or an oxidizing gas (e.g. a halogen or sulphur) seems to be one of the simplest chemical reactions. Experiments show, however, that the mechanism is rather complex, because in many cases the reaction product separates the reactants and forms a compact layer of oxide or halide on the metal surface. The oxidation can advance only as fast as the reactants diffuse across the oxide layer. It is therefore of interest to investigate the structure of these layers and the mechanism whereby atoms, ions and electrons diffuse in them. Assuming that there is no diffusion through pores or along grain boundaries, WAGNER[79] made the working hypothesis that only cations and anions migrate across the compact layer, and that the migration of electrically neutral atoms or molecules may be neglected. In the next chapter we shall consider the laws of motion of ions and electrons within the crystal. We shall see that WAGNER'S theory of tarnishing reactions[79–81, 86] and its applications to alloys by WAGNER[82, 83] and HAUFFE and associates[30, 31, 20, 21] has some success in explaining experimental observations.

The interconnexion of diffusion processes and phase boundary reactions is pictured in Fig. 1. Sometimes, the reactions at the phase boundaries may be rate-determining factors. In most cases, however, the migration of ions controls the rate of oxidation. On this basis it may be assumed either that metal ions and electrons migrate from the interface between metal and oxide across the oxide to the outer surface or that oxygen ions migrate in the opposite direction.

Reactions of metals and alloys with oxygen, sulphur and halogens at high temperatures are called tarnishing reactions. We must distinguish between oxidation processes leading to thick oxide layers and those leading to thin (or very thin) films. In the first case it is to be expected that the rate of oxidation is inversely proportional to the thickness of the oxide layer (or proportional to the square root of the time) and that electroneutrality holds. Under these conditions WAGNER'S theory is

* At present Professor of Physical Chemistry at the University of Berlin, Germany.

valid. At the other extreme, in very thin films, there is a diffuse electric double layer, where the electrical field strength is so high that the rate of migration of ions is no longer proportional to the field strength, i.e. it is no longer inversely proportional to the thickness of the oxide film, but depends quasi-exponentially upon the thickness. MOTT[65] was the first to show that under these conditions the rate of oxidation will decrease more rapidly with increasing thickness of the oxide film than is accounted for by the parabolic law. This is one of the various ways of interpreting the so-called logarithmic rate law. In Chapter 3.1 the formation of thin

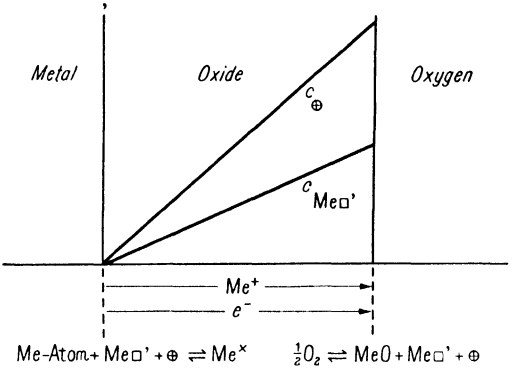


Fig. 1. Migration processes and phase boundary reactions during the oxidation of a metal forming a defect semiconducting oxide layer. (Here $\oplus$ and $Me\square'$ represent an electron hole and a metal ion vacancy, respectively. The expression "electron hole" indicates that somewhere in the lattice of the compound an electron is missing. Fig. 1 shows that electrons migrate by filling electron holes and ions migrate by filling ion vacancies)

films will be discussed in general, using the theories of CABRERA and MOTT[5] and ENGELL and HAUFFE.[12]

For thick layers, where the concentrations or activities of the reactants at the two boundaries of the oxide layer are independent of time, and thermodynamic equilibrium at the respective interfaces is practically established, the reaction rate obeys Fick's law, giving the increase of thickness Δx per unit time t,

$$\frac{\mathrm{d}(\Delta x)}{\mathrm{d}t} = \frac{k'}{\Delta x} \qquad \ldots . \quad (1)$$

where k' is a constant. Upon integration,

$$(\Delta x)^2 = 2k't \qquad \ldots . \quad (2)$$

Instead of measuring the thickness it is often easier to determine the increase of mass in grammes per cm^2 $\Delta m/A$. Equation (2) then leads to

$$\left(\frac{\Delta m}{A}\right)^2 = 2k't \qquad \ldots . \quad (3)$$

This is the parabolic rate law, which was first derived by TAMMANN[73] in 1920 and independently by PILLING and BEDWORTH[67] in 1923.

The Formation of Thick Compact Layers of Oxide on a Metal Surface

A compact layer is formed, if the molar volume of the oxide is equal to or greater than that of the metal. This general rule is only a rough one and requires modifying in view of experimental results. The theoretical basis for understanding the departure from this rule has been given by Frank and van der Merwe.[17] According to these authors it is reasonable to assume that, if the degree of misfit is less than about 15 per cent, the layer will take up the lattice parameter of the metal. If it is, however, greater than about 15 per cent, it will take up very nearly its own unstrained lattice parameter. In the latter case the layer can break away from the surface or become porous. In the first part of this report we shall confine ourselves to thick compact layers. In 4·3 thin layers will be dealt with.

Lattice Defects in Ionic Crystals

Motion of ions in a crystal lattice is understandable, if one assumes, as Frenkel has done, that not all ions are in their normal lattice positions.[18] In pure silver halides some cations have left their normal lattice points to occupy interstitial positions; in this way one cation vacancy appears per interstitial ion, as is shown in Fig. 2. The anion lattice is practically undisturbed. In a concentration gradient or a weak electrical field the silver ions can move via interstitial positions or via vacancies. Frenkel's assumptions well explain experimental results. Tubandt showed that, in silver halides, practically only silver ions are mobile.[75, 76] Furthermore, energy calculations led Jost[51, 52] to the conclusion that for the silver halides less energy is required to move a silver ion from a normal to an interstitial site than to form a pair of vacancies (proposed by Schottky for alkali halides and discussed below) by moving a cation and an anion to the surface. More detailed calculations by Mott and Littleton[63] somewhat modified this. Experiments involving the addition of the halides of divalent metals to silver halides, first carried out by Koch and Wagner,[55] have confirmed Jost's viewpoint. Teltow[74] has repeated these experiments with a considerable increase in experimental accuracy and also in the number of determinations with additions of divalent metal salts such as the halides of lead and cadmium.

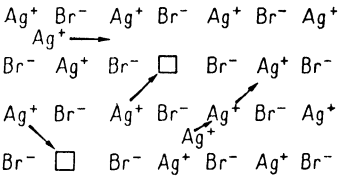


Fig. 2. Lattice defects in silver bromide. According to Frenkel. (vacant silver sites indicated by a square)

Fig. 3 shows another type of defect, first discussed by Schottky.[71]

Under the influence of thermal motion cations and anions near the surface of the alkali halides leave their normal lattice points diffusing to the surface and leaving vacancies behind. These vacancies can move in the lattice and their existence thus makes the normal ions more mobile. Experiments to determine transport numbers show that the conductivity is entirely ionic in character and due to the migration of both cations and anions. The temperature coefficient of mobility will usually be different for the two ions. The energy required to form a vacancy was first computed by JOST and NEHLEP[53] and later investigated by more refined methods by MOTT and LITTLETON.[63] SCHOTTKY defects seem to be the rule also in some other crystals, e.g. in lead halides and some sulphides and oxides.

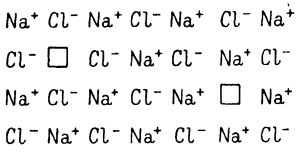


Fig. 3. Lattice defects in sodium chloride. According to SCHOTTKY

These types of lattice defects are only to be found in stoichiometric crystals. The WAGNER-SCHOTTKY disorder theory of crystals explains experimental observations and, moreover, its application led to further fruitful experiments. Thus WAGNER and co-workers[81] have developed the basic principles of ionic crystals possessing deviations from the ideal, integral, stoichiometric composition. In solids like Cu_2O there are cation vacancies $Cu\square'$ and electron holes $\oplus$. From the corpuscular viewpoint, therefore, an electron hole is represented by a Cu^{2+} ion (see Fig. 4). In the presence of oxygen, cuprous ions diffuse from the interior to the surface, there combining with chemisorbed oxygen which, on account of its electron affinity, has captured electrons, forming more Cu_2O molecules at the surface. It is, therefore, evident that the number of Cu^+-vacancies and electron holes increases with increasing oxygen pressure. Since in most cases the concentration of lattice defects is relatively low, one may tentatively apply the laws of ideal solutions. Under these assumptions we can write

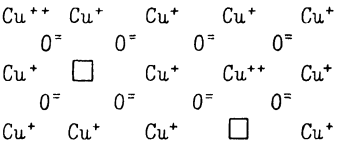


Fig. 4. Lattice defects in cuprous oxide. According to WAGNER. Squares indicate vacant copper ion lattice points. A divalent copper ion representing an electron hole

$$\tfrac{1}{2}O_2^{(g)} \rightleftharpoons Cu_2O + 2Cu\square' + 2\oplus \quad . \ . \ . \ . \quad (4)$$

Applying the law of mass action we have

$$c^2_{Cu\square'} \cdot c^2_{\oplus} = K \cdot p_{O_2}{}^{1/2} \quad . \ . \ . \ . \quad (5)$$

where K is the equilibrium constant and $c_\oplus$ is the concentration of electron holes, which is equal to the concentration of cuprous ion vacancies $c_{Cu\square'}$. Since the electrical conductivity κ is proportional to $c_\oplus$, we have

$$\kappa = \text{const} \cdot p_{O_2}{}^{1/8} \quad . \ . \ . \ . \quad (6)$$

WAGNER and associates[9, 27] find $\kappa \propto p_{O_2}^{1/7}$ in good agreement with theory. GUNDERMANN and WAGNER[28] showed that the conductivity due to cuprous ions is also proportional to $p_{O_2}^{1/7}$. NiO behaves in a similar manner. Semi-conductors in which the current is carried by electron holes are called electron-defect semi-conductors (or p-semi-conductors). Such conductors are Cu_2O, NiO,[3] FeO,[85] CoO,[85] Bi_2O_3,[57] Cr_2O_3.[33]

Solids containing an excess of metal are in distinct contrast to the above group of oxides. A representative of this group is zinc oxide. An excess of metal may be present in the form of zinc ions in interstitial positions and an equivalent number of quasi-free electrons ⊖ as shown in Fig. 5. At high temperatures ZnO dissociates, as follows

$$ZnO \rightleftharpoons Zn\bigcirc^{\cdot} + \ominus + \tfrac{1}{2}O_2^{(g)} \quad . \ . \ . \ . \quad (7a)$$

or

$$ZnO \rightleftharpoons Zn\bigcirc^{\cdot\cdot} + 2\ominus + \tfrac{1}{2}O_2^{(g)} \quad . \ . \ . \ . \quad (7b)$$

In these equations $Zn\bigcirc^{\cdot}$ and $Zn\bigcirc^{\cdot\cdot}$, respectively, represent an interstitial zinc ion, the point to the right of the circle indicating the positive net-charge. According to MILLER'S self-diffusion measurements at 1000°C[61] the assumption of the existence of interstitial zinc is better justified than is the existence of oxygen ion vacancies. Again applying the law of mass action we obtain for the electrical conductivity

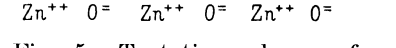


Fig. 5. Tentative scheme of lattice defects in zinc oxide as proposed by WAGNER

$$\kappa = \text{const. } p_{O_2}^{-1/4} \quad . \ . \ . \ . \quad (8a)$$

and

$$\kappa = \text{const. } p_{O_2}^{-1/6} \text{ respectively} \quad . \ . \ . \ . \quad (8b)$$

In agreement with (8) it is found experimentally that the electrical conductivity of zinc oxide decreases with increasing partial pressure of oxygen at 600°C[2]; $\kappa \propto p_{O_2}^{-1/4}$ to $p_{O_2}^{-1/5}$. CdO,[2] TiO_2,[10, 34] Al_2O_3[29] behave similarly. The special problems encountered in interpreting the observed properties of intrinsic semi-conductors such as CuO[35] will not be discussed, because it is not yet known what type of lattice defects prevail in this case.

Wagner's Theory of Oxidation of Metals

Following WAGNER we shall now derive the general formula of high-temperature oxidation of metals, which expresses the rate of oxidation in terms of the free energy of oxidation, the electrical conductivity, and the transport numbers of ions and electrons. Assuming that the ions and electrons migrate independently of each other, one obtains

for u_i, the average drift velocity in cm . s^{-1} of particles of type i in the direction of the x-axis

$$u_i = -B_i \left(\frac{1}{N} \frac{d\mu_i}{dx} + z_i e \frac{d\phi}{dx} \right) \quad . . . . \quad (9)$$

where B_i represents the mobility of particles of type i, i.e. the steady-state velocity due to a unit force per particle, corresponding to a change in the chemical potential per particle of one unit of energy (= 1 erg) for a displacement of one unit of length (= 1 cm). N is Avogadro's number, e is the electronic charge, ϕ is the local electrical potential, z_i is the electrical valency and μ_i is the chemical potential of a mole of particles of type i in ergs per mole.

If $\dot{n}_i/A$ is the number of moles migrating per unit cross section per unit time, then there follows from equation (9)

$$\frac{\dot{n}_i}{A} = c_i u_i = -B_i c_i \left(\frac{1}{N} \frac{d\mu_i}{dx} + z_i e \frac{d\phi}{dx} \right) \quad . . . . \quad (10)$$

where c_i is the concentration of particles of type i in moles per unit volume (= 1 cm^3).

As it is often desirable to eliminate B we introduce the electric current I (in amperes) and the electric conductivity κ (ohm^{-1} . cm^{-1}) in combination with the electric field strength ($d\phi/dx$) in absolute units and the factor 300 to convert to practical units by means of the relation

$$I = A\kappa 300 \left| \frac{d\phi}{dx} \right|. \quad . . . . \quad (11)$$

In the case of electrolysis $d\mu/dx$ is zero, and we obtain for the transport of ions

$$\left(\frac{n_i}{A} \right)_{\text{electrolysis}} = \frac{300}{96500} \cdot t_i \kappa \cdot \left| \frac{d\phi}{dx} \right| \quad . . . . \quad (12)$$

where t_i is the transport number of particles of type i. Comparing equations (10) and (12) we get

$$\frac{\dot{n}_i}{A} = \frac{300}{96500} \cdot \frac{t_i \kappa}{|z_i| eN} \left(-\frac{d\mu_i}{dx} - z_i eN \frac{d\phi}{dx} \right). \quad . . \quad (13)$$

In view of the chemical equilibrium between ions, electrons and electrically neutral atoms, we have

$$\mu_{\text{Me}} = \mu_1 + z_1 \mu_3, \quad . . . . \quad (14a)$$

$$\mu_{\text{X}} = \mu_2 + z_2 \mu_3, \quad . . . . \quad (14b)$$

where μ_{Me} and μ_{X} are the chemical potentials of the electrically neutral reactants, the metal Me and the non-metal X, respectively. These chemical potentials μ_{Me} and μ_{X} are not independent of each

other. According to the special form of the Gibbs-Duhem equation,

$$n_{\mathrm{Me}}\mathrm{d}\mu_{\mathrm{Me}} + n_{\mathrm{x}}\mathrm{d}\mu_{\mathrm{x}} = 0,$$

we have in the case of a small deviation from a stoichiometrically composed solid

$$\mathrm{d}\mu_{\mathrm{Me}} \approx -\left|\frac{z_1}{z_2}\right| \mathrm{d}\mu_{\mathrm{x}}. \qquad \ldots . \quad (15$$

On the basis of these equations WAGNER derives the following equation for the formation of an oxide, halide or sulphide

$$\frac{\dot{n}_{equ}}{A} = \left\{\frac{300}{96500} \cdot \frac{1}{Ne}\int_{\mu_x'}^{\mu_x''} \frac{1}{|z_2|}(t_1 + t_2)t_3\kappa \mathrm{d}\mu_x\right\} \frac{1}{\Delta x} \qquad \ldots . \quad (16$$

$$\frac{\dot{n}_{equ}}{A} = k_r \cdot \frac{1}{\Delta x}$$

where $\dot{n}_{equ}/A$ is the rate of both outward migration of metal ions and inward migration of non-metal ions expressed in equivalents per centimetre per second, Δx is the instantaneous thickness of the reaction product in centimetres, e is the electronic charge in absolute electrostatic units, z_2 is the valence of negatively charged anions, μ_x is the chemical potential of the non-metal in ergs per gram atom, where a single dash and the double dash following the μ_x indicate the equilibrium between the reaction product and the metal and between the reaction product and the non-metal, respectively; t_1, t_2 and t_3 are the transport numbers of metal ions, non-metal ions and electrons, respectively, and the constant k_r is called "rational rate constant," i.e. the reaction rate in equivalents per centimetre per second for a layer of the reaction product of 1 cm thickness.

Recently WAGNER[81] has derived a formula, similar to equation (16) which contains only the self-diffusion coefficients D_1^* and D_2^* of metal and non-metal ions, respectively. Using EINSTEIN'S equation,

$$D_i^* = B_i kT \qquad \ldots . \quad (17$$

WAGNER finally obtains the following expression for the rational rate constant k_r

$$k_r = c\int_{a_x'}^{a_x''}\left(\frac{z_1}{|z_2|} D_1^* + D_2^*\right)\mathrm{d}\ln a_x$$

$$\text{if } t_3 \approx 1 \qquad \ldots . \quad (18$$

$$= c\int_{a_{Me'}}^{a_{Me''}}\left(D_1^* + \frac{|z_2|}{z_1} D_2^*\right)\mathrm{d}\ln a_{Me}$$

where $c = c_1/z_1 = c_2/z_2$ is the concentration of metal or non-metal ions in equivalents per cubic centimetre and a_{Me} and a_x are the corresponding activities. This equation, however, is fulfilled only if the transport number of electrons t_3 is practically equal to unity.

Both equations (16) and (18) are useful for the calculation of the rate constant of oxidation reactions. To use equation (16) the electrical conductivity and the transport numbers of ions and electrons as functions of the partial pressure of the non-metal are required. Use of equation (18) is recommended if the determination of self-diffusion coefficients is easier than the determination of transport numbers.

As already pointed out by HAUFFE,[30, 36] on the basis of WAGNER'S theory of oxidation, the chemical viewpoint, i.e. the distinction between oxidation, halogenation and sulphuration reactions, is not of paramount importance. We must, however, clearly distinguish between oxidation processes in which the intermediate layer is mainly an electronic conductor and those in which it is mainly an ionic conductor. In the former case we must also distinguish between conductors in which the current is carried by quasi-free electrons and conductors in which the current is carried by electron holes. This distinction is of decisive importance for tarnishing reactions.

According to the general equation (16) the electronic conductivity is the rate determining factor if the intermediate layer is an ionic conductor. An example is the formation of AgBr on a silver surface exposed to bromine vapour. Here the rate of bromination will decrease with decreasing electronic conductivity and decreasing concentration of electron holes. In contrast to these systems with an ionic conducting layer the rate of oxidation can be increased or decreased in systems with an electronic conductor as intermediate layer if the concentration of the lattice defects is altered, for in this case the ionic conductivity is the rate determining factor.

Oxidation Processes Involving Ionic Conducting Layers

Consider the reaction between silver and gaseous bromine. According to the general scheme silver ions and electrons migrate from the interface between Ag and AgBr across the bromide to the outer surface and here they react with adsorbed bromine atoms. Whereas the concentration of silver vacancies and silver ions in interstitial positions is independent of the vapour pressure of the bromine, the concentration of electron holes is a function of the partial vapour pressure of the bromine. According to WAGNER we have[80]

$$\tfrac{1}{2}\mathrm{Br}_2^{(g)} \rightleftharpoons \mathrm{AgBr} + \mathrm{Ag}\square' + \oplus. \qquad \ldots . \quad (19)$$

Applying the law of mass action we obtain for the hole conductivity

$$\kappa_{\oplus} = \kappa_{\oplus(p_{\mathrm{Br}_2}=1)}\, p_{\mathrm{Br}_2}^{1/2}, \qquad \ldots . \quad (20)$$

and finally by combining with equation (16) for the rate of bromination

$$\frac{\dot{n}_{\mathrm{AgBr}}}{A} = \left\{\frac{300}{96500}\,\kappa_{\oplus(p_{\mathrm{Br}_2}=1)}\,\frac{RT}{Ne}\,p_{\mathrm{Br}_2}^{1/2}\right\}\frac{1}{\Delta x} \qquad \ldots . \quad (21)$$

if $p_{Br_2} = p_{Br_2}'' \gg p_{Br_2}'$. In agreement with this, WAGNER was able to show that between 200 and 400°C the rate of bromination is proportional to the square root of the bromine vapour pressure. Furthermore the transport number t_3 of electron holes at 200°C and $p_{Br_2} = 0{\cdot}23$ atm is expected from bromination experiments to be about 17 per cent, whereas from conductivity measurements it was found to be 12 per cent.[80] This result is in satisfactory agreement with WAGNER'S assumptions. The same mechanism may be assumed to apply to the reaction of silver with chlorine.

Oxidation Processes Involving Excess Semi-conducting Layers

Next, we consider the reaction between zinc and oxygen. Experiments showed that the electric conductivity of zinc oxide is mainly due to quasi-free electrons $\ominus$ and, to a lesser extent, to interstitial zinc ions. Consequently, in contrast to the halogenation of silver, the rate determining factor in the oxidation of zinc is the migration of zinc ions via interstitial positions, i.e. the ionic conductivity. Applying the law of mass action to equation (7)

$$c_{Zn\bigcirc\cdot} \cdot c_{\ominus} = K p_{O_2}^{-1/2}; \qquad \ldots\ldots \quad (22)$$

and putting $c_{Zn\bigcirc\cdot} = c_{\ominus}$, we obtain for the conductivity due to zinc ions

$$\kappa_{Zn} = \kappa_{Zn(p_{O_2}=1)}\, p_{O_2}^{-1/4}. \qquad \ldots\ldots \quad (23)$$

Proceeding in the same way as in the derivation of equation (21) we finally obtain for the rate of oxidation

$$\frac{\dot{n}_{ZnO}}{A} = \left\{\frac{300}{96500}\,\kappa_{Zn(p_{O_2}=1)}\,\frac{RT}{Ne}\left(\sqrt[4]{\frac{1}{p_{O_2}'}} - \sqrt[4]{\frac{1}{p_{O_2}''}}\right)\right\}\frac{1}{\Delta x}. \qquad \ldots\ldots \quad (24)$$

It follows from equation (24) that the rate of oxidation of zinc is independent of the partial pressure of oxygen. This conclusion was experimentally confirmed by WAGNER and GRUENEWALD[86] and can easily be understood. Fig. 6 shows the distribution of interstitial zinc ions and of quasi-free electrons in the lattice. Near the oxide/metal interface the oxide is rich in interstitial zinc ions and quasi-free electrons, whereas their concentration at the outer surface is very small. It is, therefore, apparent that the concentration gradient of $Zn\bigcirc\cdot$ in zinc oxide, and thus the rate of oxidation of zinc, should be almost independent of the oxygen pressure.

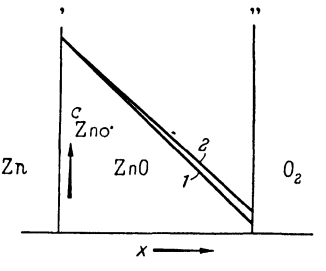


Fig. 6. Concentration gradients in zinc oxide at oxygen pressures of 0·1 atm (Curve 1) and 0·001 atm (Curve 2)

The rate constant of zinc oxidation, as found by WAGNER and GRUENEWALD, is $k_r = 7{\cdot}5 \,.\, 10^{-11}$ g^2 cm^{-4} h^{-1} at 400°C. In contrast to

this, the rate constant of oxidation of spectroscopically pure zinc under exactly similar experimental conditions was found by GENSCH and HAUFFE[20] to be $k_r = 9 \cdot 10^{-10}$ at 390°C. It must be assumed that the high value of k_r is a sign of great purity of the zinc used. The oxidation of zinc alloys, mentioned below, will show that this conclusion is justified.

It is to be expected that the oxidation of cadmium proceeds according to the same mechanism, for FAIVRE and MICHEL[14] found that an increasing excess of metal leads to an increase of the lattice parameter. This as well as the dependence of the electric conductivity on the partial pressure of oxygen makes it probable that there are metal ions in interstitial positions, and quasi-free electrons. Conductivity measurements carried out by EARLE,[10] and HAUFFE and associates[34] on TiO_2 show that this, too, is an excess semi-conductor; however, it contains no interstitial cations, but oxygen ion vacancies as shown by EHRLICH.[11] GULBRANSEN and associates[22] have investigated the dependence of the rate of oxidation of titanium on oxygen pressure.

Oxidation Processes Involving Deficit Semi-conducting Layers

Good examples of this type of process are the reactions of iron and copper with oxygen, the wüstite-phase and cuprous oxide having been investigated in particular by JETTE and FOOTE,[49] DARKEN and GURRY,[8] and by WAGNER and associates.[85] Unfortunately the rate of oxidation of iron to FeO at 950 and 1000°C is determined by surface reactions.[39] We shall, therefore, rather consider the oxidation of copper investigated by WAGNER and GRUENEWALD.[86]

Since Cu_2O is an electronic conductor, the velocity of diffusion of copper ions and electrons is determined by the more slowly moving copper ions, moving via vacancies. According to equation (16) the copper ion conductivity $\kappa_{Cu} = \kappa \cdot t_1$ is the rate determining factor. Using the experimental result

$$\kappa_{Cu} = \kappa_{Cu(p_{O_2}=1)}\, p_{O_2}^{1/7} \quad \ldots\ (25)$$

and the relation

$$d\mu_x = \tfrac{1}{2} d\mu_{O_2} = \tfrac{1}{2} RT d\ln p_{O_2} \quad \ldots\ (26)$$

we obtain for the rate of oxidation of copper

$$\frac{\dot{n}_{Cu_2O}}{A} = \frac{300}{96500}\, \kappa_{Cu(p_{O_2}=1)}\, \frac{7}{4}\, \frac{RT}{Ne} \left(\sqrt[7]{p_{O_2}''} - \sqrt[7]{p_{O_2}'}\right) \frac{1}{\Delta x} \quad \ldots\ (27a)$$

or

$$= (1 + z_{Cu^+})\, B_{Cu\square}\, kT \left(\sqrt[7]{p_{O_2}''} - \sqrt[7]{p_{O_2}'}\right) \frac{K^{1/7}}{\Delta x} \quad \ldots\ (27b)$$

The mobility of the cation vacancies can be calculated from the electric conductivity, the transport number of the cations, and the copper

deficit in cuprous oxide. These data have been used by WAGNER and GRUENEWALD to calculate the values of the rational rate constant k_r as indicated in Table 1. Observed and calculated values of k_r are in good agreement. Furthermore, the experimental observations led to $k_r \propto p_{O_2}^{1/7}$, in fairly good agreement with the theoretically expected relation $k_r \propto p_{O_2}^{1/8}$ (see Fig. 7).

TABLE 1

Calculated and Measured Rate Constants of the Formation of Cuprous Oxide on Copper at 1000°C as measured by WAGNER and GRUENEWALD

Oxygen Pressure	*Rational Rate Constant* k_r • (*equivalents* . cm^{-1} . s^{-1})	
in atm	*Calculated*	*Observed*
$8{\cdot}3 \cdot 10^{-2}$	$6{\cdot}6 \cdot 10^{-9}$	$6{\cdot}2 \cdot 10^{-9}$
$1{\cdot}5 \cdot 10^{-2}$	$4{\cdot}8 \cdot 10^{-9}$	$4{\cdot}5 \cdot 10^{-9}$
$2{\cdot}3 \cdot 10^{-3}$	$3{\cdot}4 \cdot 10^{-9}$	$3{\cdot}1 \cdot 10^{-9}$
$3{\cdot}0 \cdot 10^{-4}$	$2{\cdot}1 \cdot 10^{-9}$	$2{\cdot}2 \cdot 10^{-9}$

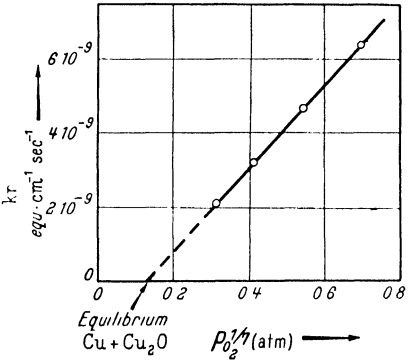


Fig. 7. Dependence of the rate of oxidation of copper on the partial pressure of oxygen at 1000°C according to WAGNER and GRUENEWALD

SHOCKLEY and associates,[1] and CASTELLAN and MOORE,[7] have investigated the self-diffusion of copper in cuprous oxide which was growing on copper but unfortunately the cuprous oxide was not homogeneous. Therefore their self-diffusion coefficients cannot be used in WAGNER'S third formula (27*c*) for the rate of oxidation of copper (cf. equation (18)), namely

$$k_r = (1 + z_{Cu^+}) c D^{*''}_{Cu^+} \{1 - (c'_{Cu\square'}/c''_{Cu[\square']})\} \quad . \; . \; . \; . \; (27c)$$

where $D^{*''}_{Cu^+}$ is the self-diffusion coefficient of copper ions in cuprous oxide in equilibrium with oxygen.

In the same manner WAGNER and GRUENEWALD have investigated the oxidation of nickel and found $k_r \propto p_{O_2}^{1/4}$, whereas theoretically $k_r \propto p_{O_2}^{1/6}$ should be obtained (see equations (48) and (49)). The discrepancy between experimental results and theoretical expectation is not yet understood. Recently HAUFFE and RAHMEL[37] have investigated the reaction between iron and sulphur at 670°C. According to the equation

$$\tfrac{1}{2}S_2^{(g)} \rightleftharpoons FeS + Fe\square'' + 2\oplus \quad . \quad . \quad . \quad . \quad (28)$$

the rate of sulphuration of iron is a linear function of the 1/4 and 1/7 power of the sulphur vapour pressure (see Fig. 8). The theoretical relation is $k_r \propto p_{S_2}^{1/6}$. The agreement is fairly good.

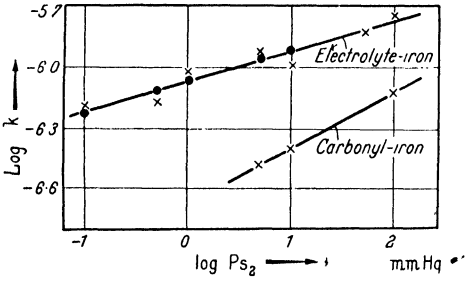


Fig. 8. Dependence of the rate of sulphuration of iron on the partial pressure of sulphur at 670°C according to HAUFFE and RAHMEL (k in $g^2 \cdot cm^{-4} \cdot s^{-1}$)

APPLICATION OF WAGNER'S THEORY OF OXIDATION TO THE OXIDATION OF ALLOYS

In the above chapters we considered the oxidation of pure metals and developed quantitative relations between transport of both ions and electrons and the rate of oxidation of metals. In this chapter we shall consider the influence of the additions of small amounts of a foreign metal on the rate of oxidation. Accepting WAGNER'S theory of oxidation and assuming that the mobilities of the disordered ions and electrons will not be markedly changed by the presence of foreign ions, we must diminish the concentration of ionic or electronic defects in order to obtain a decrease in the rate of oxidation. Conversely the velocity of oxidation will increase on increasing the concentration of ionic defects, electron holes or quasi-free electrons. To understand this we must consider the application of the law of mass action to ionic and electronic lattice defects in mixed crystals.

Oxidation Processes Involving Ionic Conducting Layers

WAGNER,[87] HAUFFE and associates[32–35, 38, 40] have shown that the concentration of vacancies and ions in interstitial positions, or of

electron holes and quasi-free electrons, and likewise the electrical conductivity can be greatly varied by the addition of ions of higher or lower valencies. This will be shown for the case of silver bromide. For a FRENKEL type of defect we have the following equilibrium

$$AgBr_{undisturbed} \rightleftharpoons Ag\square' + Ag\bigcirc^{\cdot}, \quad \ldots\ldots (29)$$

and with the law of mass action we obtain

$$c_{Ag\square'} \cdot c_{Ag\bigcirc^{\cdot}} = K. \quad \ldots\ldots (30)$$

This equation is valid for pure silver bromide and for silver bromide containing small quantities of another metal bromide. In the case of pure silver bromide we have the special condition

$$c^{\circ}_{Ag\square'} = c^{\circ}_{Ag\bigcirc^{\cdot}} \quad \ldots\ldots (31)$$

If cadmium bromide is dissolved in silver bromide, divalent cadmium ions replace silver ions and one cation vacancy $Ag\square'$ appears for each divalent cadmium ion replacing a silver ion in the lattice $Cd\bullet^{\cdot}(Ag)$ (see Fig. 9). The equation for the solution of cadmium bromide in silver bromide then reads

Ag^+	Br^-	Ag^+	Br^-	Cd^{++}	Br^-
Br^-	$\square$	Br^-	Ag^+	Br^-	Ag^+
Ag^+	Br^-	Ag^+	Br^-	$\square$	Br^-
Br^-	Ag^+	Br^-	Cd^{++}	Br^-	Ag^+

Fig. 9. Lattice defects in a mixed crystal of AgBr-$CdBr_2$ as proposed by WAGNER

$$CdBr_2 \rightleftharpoons Cd\bullet^{\cdot}(Ag) + Ag\square' + 2AgBr. \quad \ldots\ldots (32)$$

Consequently, because of electro-neutrality, in a binary solution the number of cation vacancies is equal to the number of cadmium ions plus the number of silver ions in interstitial positions

$$c_{Ag\square'} = c_{Ag\bigcirc^{\cdot}} + c_{Cd\bullet^{\cdot}(Ag)}. \quad \ldots\ldots (32)$$

Using (30), (31) and (32) we have

$$\frac{c_{Ag\square'}}{c^{\circ}_{Ag\square'}} = \frac{c_{Cd}}{2c^{\circ}_{Ag\square'}} + \left[\left(\frac{c_{Cd}}{2c^{\circ}_{Ag\square'}}\right)^2 + 1\right]^{1/2} \quad \ldots\ldots (33)$$

where $c_{Cd} = c_{Cd\bullet^{\cdot}(Ag)}$.

It follows from (33) that on increasing the concentration of cadmium bromide in silver bromide the concentration of cation vacancies and, therefore, the electric conductivity will increase. Experiments carried out by KOCH and WAGNER[54] and TELTOW[74] confirm the above conclusion. This result is of importance for the interpretation of the rate of bromination of silver alloys as determined by HAUFFE and GENSCH.[31, 21]

For the rate of bromination of silver and silver alloys the electron hole conductivity is the rate determining factor and we must, therefore, find a relation between the hole conductivity and the cadmium concentration in silver and the $CdBr_2$ concentration in AgBr, respectively.

Applying the law of mass action to (19) at constant bromine vapour pressure

$$c_{Ag\square'} \cdot c_{\oplus} = K \text{ (for } p_{Br_2} = \text{const)}. \quad \ldots \quad (34)$$

Observing the relation

$$\frac{k}{k^\circ} = \frac{\kappa_{\oplus}}{\kappa^\circ_{\oplus}} = \frac{c_{\oplus}}{c^\circ_{\oplus}} = \frac{c^\circ_{Ag\square'}}{c_{Ag\square'}} \quad \ldots \quad (35)$$

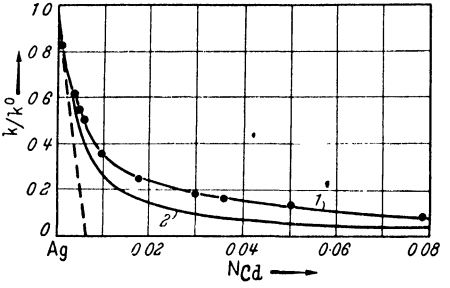


Fig. 10. The dependence of the ratio of the rate constant of bromination of the silver-cadmium alloy, k, to that of the pure silver, k°, on the amount of cadmium added N_{Cd} (atom fraction) as measured by HAUFFE and GENSCH (T = 330°C and p_{Br_2} = 170 mm Hg), Curve 1 observed and Curve 2 calculated

and using equation (33) we finally obtain for the ratio of the oxidation rate of silver-cadmium alloys k to that of pure silver k°

$$\frac{k}{k^\circ} = \frac{1}{\frac{c_{Cd}}{2c^\circ_{Ag\square'}} + \left[\left(\frac{c_{Cd}}{2c^\circ_{Ag\square'}}\right)^2 + 1\right]^{1/2}} \quad \ldots \quad (36)$$

Using TELTOW'S value for $c^\circ_{Ag\square'} = 2{\cdot}9 \cdot 10^{-3}$ at 330°C the calculated values of k agree satisfactorily with the values of k obtained by the experiments shown in Fig. 10. Silver-lead and silver-zinc alloys behave similarly under identical experimental conditions (see Fig. 11).

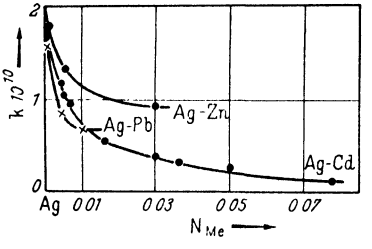


Fig. 11. Dependence of the rate constant of bromination of the silver alloy, k_r (equ. $cm^{-1}\ s^{-1}$), on the increasing amount of cadmium, lead or zinc added as measured by GENSCH and HAUFFE (T = 330°C and p_{Br_2} = 170 mm Hg)

According to the order of increasing solubility, $ZnBr_2 < PbBr_2 < CdBr_2$, we have

$$\left(\frac{k}{k^\circ}\right)_{\min}^{Ag/Cd} < \left(\frac{k}{k^\circ}\right)_{\min}^{Ag/Pb} < \left(\frac{k}{k^\circ}\right)_{\min}^{Ag/Zn}.$$

Similar results were obtained by HIMMLER,[47] who found that the rate of chlorination of silver is decreased by the presence of lead chloride and cadmium chloride. The chance of developing silver alloys more resistent to attack by halogens at high temperatures than silver-cadmium alloys does not seem to be very great at present, if it is remembered that only metals which are soluble in silver, whose ionic radius is 1·1 ± 0·2 Å and whose valency is 2 or 3 can be used. Only if the radius falls within the above limits can a solubility in AgBr large enough to produce a high concentration of silver ion vacancies be expected. Employing GOLDSCHMIDT'S values for the ionic radii only the following cations need be considered: Cd^{2+} (1·03), Pb^{2+} (1·32), Cu^{2+} (0·96), Tl^{3+} (1·05), Sr^{2+} (1·27) and La^{3+} (1·22). The use of silver alloys containing more than one alloying metal is not desirable, because the more soluble metal ion diminishes the solubility of the other metal ions in silver halides.[89] The resistance of ternary silver alloys is thereby often reduced.

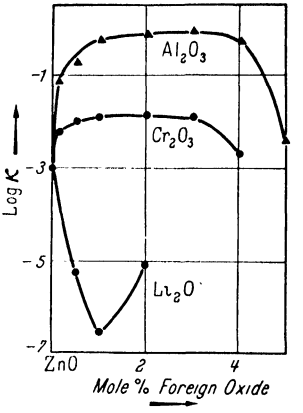


Fig. 12. Variation of the electrical conductivity of zinc oxide with increasing amounts of trivalent and monovalent cations at 390 °C as measured by HAUFFE and VIERK (p_{air} = 1 atm)

Oxidation Processes Involving Quasi-free Electron Conducting Layers

We shall now consider the influence of small amounts of aluminium and lithium in zinc on the rate of oxidation of zinc. If we oxidize a zinc-aluminium alloy at 390°C in an atmosphere of oxygen, a layer of zinc oxide containing Al_2O_3 is formed. Measurements of the electric conductivity of zinc oxide carried out by HAUFFE and VIERK[40] (see Fig. 12) showed that with increasing amounts of Al_2O_3 the concentration of quasi-free electrons increases while that of zinc ions in interstitial positions decreases. The equation corresponding to the solution of Al_2O_3 in zinc oxide is

$$Al_2O_3 \rightleftharpoons 2Al\bullet^{\cdot}(Zn) + 2\ominus + 2ZnO + \tfrac{1}{2}O_2^{(g)} \quad . \; . \; . \; (37a)$$

or

$$\tfrac{1}{2}O_2^{(g)} + 2Zn\circ^{\cdot} + Al_2O_3 \rightleftharpoons 2Al\bullet^{\cdot}(Zn) + 4ZnO. \quad . \; . \; (37b)$$

Fig. 13 shows aluminium ions in zinc ion positions. Because of

electro-neutrality quasi-free electrons $\ominus$ will appear while interstitial zinc ions will disappear; for the same reason we have

$$c_{\ominus} = c_{\mathrm{Zn}\bigcirc\cdot} + c_{\mathrm{Al}\bullet\cdot(\mathrm{Zn})}. \quad \ldots \quad (38)$$

Thus, using equation (38) and substituting in equation (22) at constant oxygen pressure, we obtain for the ratio of the rate of oxidation constants of the Zn/Al alloy k to that of the pure zinc k°

$$\frac{k}{k^{\circ}} = \frac{1}{\dfrac{c_{\mathrm{Al}}}{2c^{\circ}_{\mathrm{Zn}\bigcirc\cdot}} + \left[\left(\dfrac{c_{\mathrm{Al}}}{2c^{\circ}_{\mathrm{Zn}\bigcirc\cdot}}\right)^2 + 1\right]^{1/2}} \quad \ldots \quad (39)$$

where $c_{\mathrm{Al}} = c_{\mathrm{Al}\bullet\cdot(\mathrm{Zn})}$ is the concentration of aluminium ions at zinc ion lattice points.

Measurements of the electric conductivity of mixtures of zinc oxide

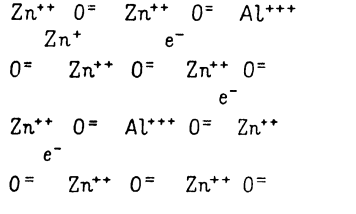


Fig. 13. Lattice defects in a mixed crystal of ZnO-Al_2O_3

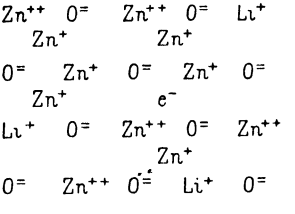


Fig. 14. Lattice defects in a mixed crystal of ZnO-Li_2O

and aluminium oxide at 400°C showed that the concentration of interstitial zinc ions in pure ZnO can be estimated as $c_{\mathrm{Zn}\bigcirc\cdot} \approx 4 \cdot 10^{-5}$.[40] It is easy to show, therefore, that if the addition amounts to more than 0·1 atom per cent, equation (39) becomes

$$\frac{k}{k^{\circ}} \approx \frac{c^{\circ}_{\mathrm{Zn}\bigcirc\cdot}}{c_{\mathrm{Al}}}. \quad \ldots \quad (40)$$

We can treat the oxidation of a zinc-lithium alloy in the same way. For zinc oxide containing lithium oxide we have

$$\tfrac{1}{2}\mathrm{O}_2{}^{(g)} + 2\ominus + \mathrm{Li}_2\mathrm{O} \rightleftharpoons 2\mathrm{Li}\bullet'(\mathrm{Zn}) + 2\mathrm{ZnO}. \quad \ldots \quad (41)$$

Assuming the validity of equation (41) we see that adding lithium oxide to zinc oxide causes an increase in the concentration of interstitial zinc ions and a decrease in the concentration of quasi-free electrons (see Fig. 14).

Furthermore, the condition of electro-neutrality requires

$$c_{\mathrm{Zn}\bigcirc\cdot} = c_{\ominus} + c_{\mathrm{Li}\bullet'(\mathrm{Zn})}. \quad \ldots \quad (42)$$

Thus, combining equation (42) and the equilibrium condition at constant oxygen pressure, equation (22), we finally obtain for the ratio of

the oxidation rate constant of the Zn/Li alloy k to that of pure zinc k°

$$\frac{k}{k^\circ} = \frac{c_{Li}}{2c^\circ_{Zn\bigcirc\cdot}} + \left[\left(\frac{c_{Li}}{2c^\circ_{Zn\bigcirc\cdot}}\right)^2 + 1\right]^{1/2}; \quad \ldots . \quad (43)$$

and if $c_{Li} = c_{Li\bullet(Zn)} \gg c^\circ_{Zn\bigcirc\cdot}$

$$\frac{k}{k^\circ} \approx \frac{c_{Li}}{c^\circ_{Zn\bigcirc\cdot}}. \quad \ldots . \quad (44)$$

The experimental results obtained by GENSCH and HAUFFE[20] confirm the validity of the assumption made (see Fig. 15). Table 2 shows the observed and the calculated values for the oxidation rate constants of spectroscopically pure zinc and for zinc alloys. The curious dependence

TABLE 2

Calculated and Measured Rate Constants of the Formation of Oxide on Zinc and Zinc Alloys at 390°C, as measured by GENSCH and HAUFFE ($p_{O_2} = 1$ atm)

Zinc alloys in atom per cent	*Rate Constant of Oxidation* k ($g^2\ cm^{-4}\ h^{-1}$)	
	Observed	*Calculated*
zinc, pure	$0{\cdot}8 \cdot 10^{-9}$	—
zinc + 0·1 Al	$1{\cdot}0 \cdot 10^{-11}$	$4 \cdot 10^{-11}$
+ 1·0 Al	$< 1{\cdot}0 \cdot 10^{-11}$	$4 \cdot 10^{-12}$
zinc + 0·4 Li	$2{\cdot}0 \cdot 10^{-7}$	$1 \cdot 10^{-7}$
zinc + 0·1 Tl	$2{\cdot}0 \cdot 10^{-9}$	—
+ 0·5 Tl	$2{\cdot}0 \cdot 10^{-10}$	- -
+ 1·0 Tl	$1{\cdot}3 \cdot 10^{-10}$	- -

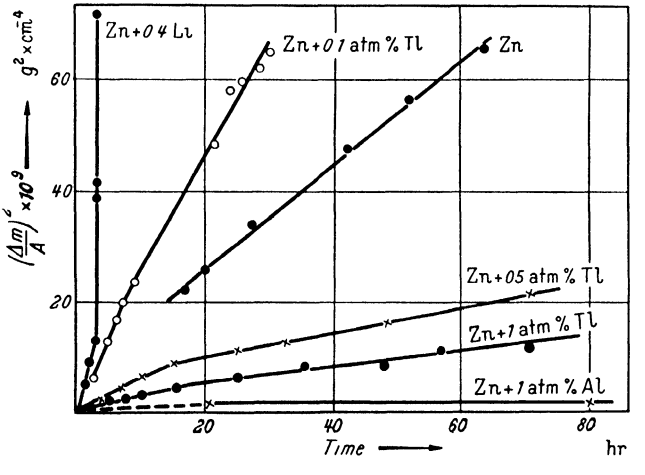


Fig. 15. Influence of foreign metals on the rate of oxidation of zinc, as measured by GENSCH and HAUFFE ($p_{O_2} = 1$ atm and $T = 390°C$)

of the rate of oxidation of the zinc-thallium alloys on the thallium content of the alloy is remarkable. As shown by Fig. 15 the rate of oxidation of an alloy containing 0·1 atom per cent Tl is greater than that of pure zinc, whereas a greater thallium content (0·5 and 1·0 atom per cent) leads to a rate of oxidation which is lower than that of pure zinc. This behaviour can only be understood if it is tentatively assumed that small additions of thallium enter the zinc oxide lattice as univalent cations whereas greater quantities enter as trivalent cations.

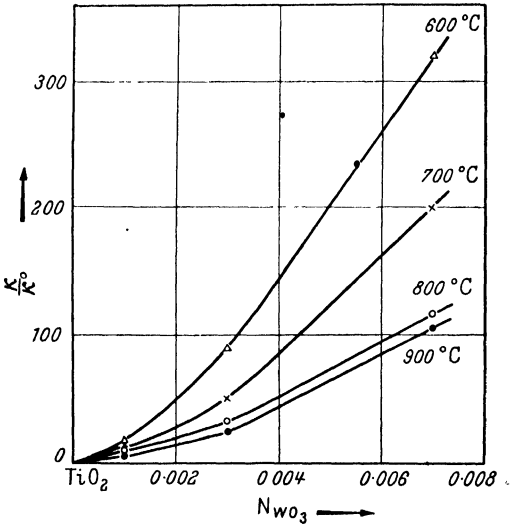


Fig. 16. Dependence of the electrical conductivity of TiO_2 on the increasing amounts of WO_3 (N_{WO_3} = mole fraction) added (p_{air} = 1 atm). According to HAUFFE, GRUNEWALD and GREESE

Since the common impurities of zinc are aluminium, iron and gallium, but not alkali metals, oxidation produces zinc oxide layers containing trivalent cations and giving a lower rate of oxidation than observed for pure zinc. It is, therefore, permissible to assume that the high value of the rate of oxidation of zinc mentioned on page 79 is a direct indication of the purity of the zinc used.

From the theoretical viewpoint, and according to conductivity measurements on mixed crystals of TiO_2 and WO_3 recently carried out by HAUFFE, GRUNEWALD and GREESE[34] (cf. Fig. 16), the rate of oxidation of titanium may be expected to diminish in the presence of tungsten, which decreases the concentration of oxygen ion vacancies. An investigation of the oxidation of titanium alloys from this special point of view has not yet been made. The dissolution of tungsten oxide in titanium oxide can be formulated as follows:

$$WO_3 \rightleftharpoons W\bullet\bullet(Ti) + 2\ominus + TiO_2 + \tfrac{1}{2}O_2 \quad . \; . \; . \; . \quad (45)$$

$$O\square^{\cdot} + WO_3 \rightleftharpoons W\bullet\bullet(Ti) + \ominus + TiO_2 + \tfrac{1}{2}O_2^{(g)}. \quad . \; . \; (45a)$$

According to the tentatively proposed equilibrium,

$$TiO_2 \rightleftharpoons O\square^{\cdot} + \ominus + \tfrac{1}{2}O_2^{(g)}, \qquad \ldots \quad (46)$$

the rate of oxidation of titanium and titanium alloys, in contrast to the oxidation of zinc and its alloys, seems to be determined by the diffusion of oxygen ions. HAUFFE and PFEIFFER[39] show that the rate of

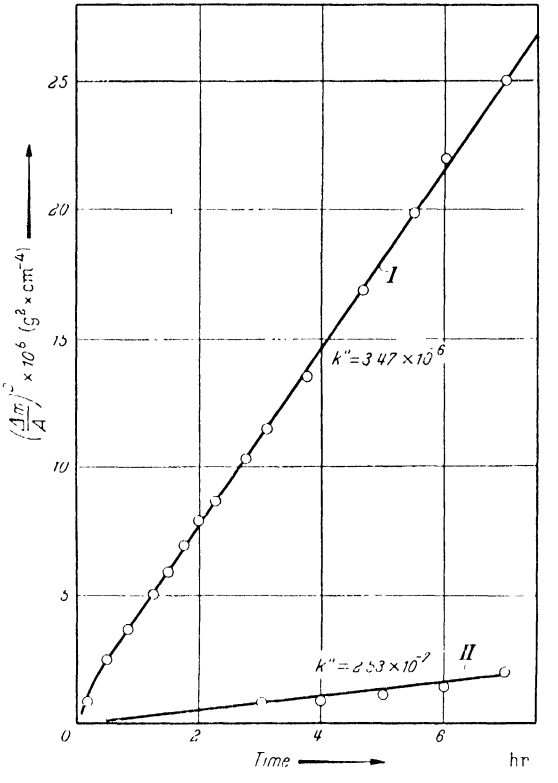


Fig. 17. Decrease of the rate of oxidation of titanium due to the presence of WO_3 (curve II) in oxygen of 1 atm at 800°C as measured by HAUFFE and PFEIFFER

oxidation of titanium decreases in the presence of WO_3 vapour in oxygen (Fig. 17).

Oxidation Processes Involving Electron-hole-conducting Layers

The rate of oxidation of a metal covered by an electron-hole-conducting oxide layer can be increased or decreased by addition of a foreign metal whose oxide dissolves in the mother oxide and which gives monovalent or trivalent cations respectively. The cause of this behaviour can be understood by applying the theory of lattice defects

to mixed oxide crystals. Fig. 18 shows schematically what happens when monovalent ions, i.e. Li^+, enter the crystal lattice of nickel oxide which is of the hole conducting type. According to the equation

$$\tfrac{1}{2}O_2^{(g)} + Li_2O \rightleftharpoons 2Li\bullet'(Ni) + 2\oplus + 2NiO \quad . \; . \; . \quad (47)$$

Ni^{++}	$O^=$	Ni^{+++}	$O^=$	Ni^{++} $O^=$
$O^=$	Li^+	$O^=$	Ni^{++}	$O^=$ Ni^{++}
Ni^{++}	$O^=$	Ni^{+++}	$O^=$	Li^+ $O^=$
$O^=$	□	$O^=$	Ni^{++}	$O^=$ Ni^{+++}
Ni^{+++}	$O^=$	Ni^{++}	$O^=$	Ni^{++} $O^=$

Fig. 18. Lattice defects in a mixed crystal of NiO-Li_2O

where $Li\bullet'(Ni)$ represents a lithium ion in a nickel position and the $'$ the negative charge of this "impurity centre," we obtain an increase of the electron hole concentration with an increasing concentration of lithium oxide, and, therefore, an increase in the electrical conductivity. The conductivity measurement of VERWEY and associates[78] have confirmed the presence of this type of defect (see Fig. 19).

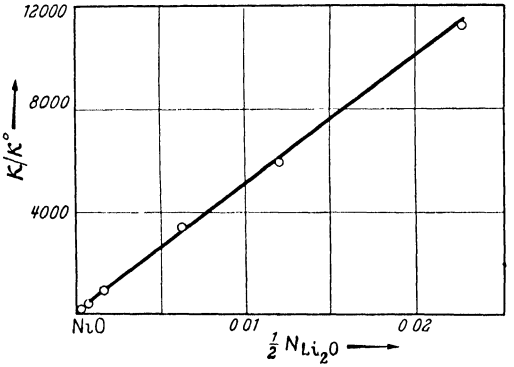


Fig. 19. Dependence of the electrical conductivity of NiO on the increasing amount of Li_2O added, as measured by VERWEY, HAAYMAN and ROMEYN at 20°C

In analogy to equation (4) the equilibrium condition for lattice defects in nickel oxide in an atmosphere of oxygen reads

$$\tfrac{1}{2}O_2^{(g)} \rightleftharpoons NiO + Ni\square'' + 2\ominus. \quad . \; . \; . \; . \quad (48)$$

Applying the ideal law of mass action, we get

$$c_{Ni\square''} \cdot c_{\oplus}^{2} = K \text{ (if } p_{O_2} = \text{const)} \quad . \; . \; . \; . \quad (49)$$

where $c_{Ni\square''}$ represents the concentration of nickel ion vacancies.

Furthermore, electroneutrality leads to

$$c_{\oplus} = 2c_{Ni\square''} + c_{Li\bullet'(Ni)}. \quad . \; . \; . \; . \quad (50)$$

If $c_{Li} = c_{Li\bullet'(Ni)} \gg c^{\circ}_{Ni\square''}$ and if we substitute silver ions for lithium ions, then the ratio of the oxidation rate constant of a nickel-silver

alloy to that of pure nickel k°, calculated from equations (49) and (50), equals

$$\frac{k}{k^\circ} \approx \left(\frac{2c^\circ_{\mathrm{Ni}\square''}}{c_{\mathrm{Ag}}}\right)^2 \quad \ldots . \quad (51)$$

The decrease in the rate of oxidation due to the presence of silver in nickel is far less than expected from equation (51). The cause seems to be the small solubility of silver oxide in nickel oxide, as observed even at 1000°C by HAUFFE and PFEIFFER.[41]

Making use of the remarkably great solubility of lithium oxide in nickel oxide and the fact that at 1100°C the vapour pressure of lithium

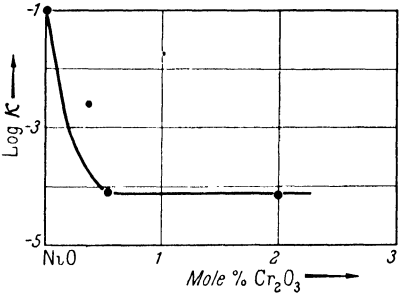


Fig. 20. Dependence of the electrical conductivity of NiO on the increasing amount of Cr_2O_3 added, according to HAUFFE and BLOCK

oxide is satisfactorily high, HAUFFE and PFEIFFER[39] investigated the attack of oxygen on nickel at 1000° and 1100°C in the presence of Li_2O vapour. According to equations (47), (50) and (51*a*)

$$\frac{k}{k^\circ} \approx \left(\frac{2c^\circ_{\mathrm{Ni}\square''}}{c_{\mathrm{Li_2O}}}\right)^2 \quad \ldots . \quad (51a)$$

we obtain a decrease in the rate of oxidation of nickel (see Fig. 19). At present a quantitative evaluation is not yet possible.

On the other hand, if chromium oxide is dissolved in nickel oxide the electric conductivity is lowered considerably (see Fig. 20).[38] According to WAGNER's scheme shown in Fig. 21, the presence of trivalent chromium ions lead to an increase in the concentration of nickel ion vacancies and to a decrease in the concentration of electrons. The equation for the solution of Cr_2O_3 in nickel oxide then reads

$$\begin{array}{cccccc} Ni^{++} & O^= & Cr^{+++} & O^= & Ni^{++} & O^= \\ O^= & \square & O^= & Ni^{++} & O^= & Cr^{+++} \\ Ni^{++} & O^= & Ni^{++} & O^= & \square & O^= \\ O^= & Cr^{+++} & O^= & Cr^{+++} & O^= & Ni^{++} \end{array}$$

Fig. 21. Lattice defects in a mixed NiO-Cr_2O_3 crystal as proposed by WAGNER

$$2\oplus + Cr_2O_3 \rightleftharpoons 2Cr\bullet^{\cdot}(Ni) + 2NiO + \tfrac{1}{2}O_2^{(g)} \quad \ldots \quad (52a)$$

or

$$Cr_2O_3 \rightleftharpoons 2Cr\bullet^{\cdot}(Ni) + Ni\square'' + 3NiO. \quad \ldots . \quad (52b)$$

Here the symbol $Cr\bullet^{\cdot}(Ni)$ represents a chromium ion in a nickel ion site.

Remembering that the condition of electro-neutrality holds in the $NiO\text{-}Cr_2O_3$ phase we obtain

$$c_{Ni\square''} = \tfrac{1}{2}c_{\oplus} + \tfrac{1}{2}c_{Cr\bullet(Ni)} \quad . \quad . \quad . \quad . \quad (53)$$

and, approximating,

$$c_{Ni\square''} \approx \tfrac{1}{2}c_{Cr} \text{ (if } c_{Cr} = c_{Cr\bullet(Ni)} \gg c^{\circ}_{Ni\square''}). \quad . \quad . \quad . \quad . \quad (54)$$

The ratio of the oxidation rate constant of a nickel-chromium alloy k to that of pure nickel k° is

$$\frac{k}{k^{\circ}} \approx \frac{c_{Cr}}{2c^{\circ}_{Ni\square''}} \text{ (if } c_{Cr} \gg c^{\circ}_{Ni\square''}). \quad . \quad . \quad . \quad . \quad (55)$$

Although there had been many empirical investigations on the oxidation of nickel alloys, the first experiments with the theory of lattice defects in view were carried out by WAGNER and ZIMENS.[83] In agreement with the requirements of equation (55), Table 3 shows that the rate of oxidation of nickel increase with increasing amounts of chromium. The relatively rapid decrease in the rate of oxidation at higher concentrations of chromium (> 6 atom per cent Cr) is probably due to the formation of a compact layer of the spinel phase $NiCr_2O_4$, in which ionic diffusion processes occur extremely slowly.

TABLE 3

The rate of oxidation of nickel-chromium alloys at 1000 °C as measured by WAGNER and ZIMENS

Weight per cent Cr	t sec	$\Delta m/A$ g . cm^{-2}	$k = \frac{1}{t}\left(\frac{\Delta m}{A}\right)^2$
—	57600	$4{\cdot}1 \cdot 10^{-3}$	$3{\cdot}1 \cdot 10^{-10}$
0·3	57600	$8{\cdot}8 \cdot 10^{-3}$	$14{\cdot}0 \cdot 10^{-10}$
1·0	57600	$12{\cdot}3 \cdot 10^{-3}$	$26{\cdot}0 \cdot 10^{-10}$
3·0	57600	$13{\cdot}4 \cdot 10^{-3}$	$31{\cdot}0 \cdot 10^{-10}$
10·0	57600	$2{\cdot}9 \cdot 10^{-3}$	$1{\cdot}5 \cdot 10^{-10}$

As was pointed out by WAGNER and ZIMENS the quantitative discussion of the experimental results is rather difficult, because we must assume that nickel and chromium ions do not possess the same mobilities, so that we are here dealing with a solid solution whose composition is a function of the distance from the oxide/metal interface. If the mobility of the chromium ion is small, the region near the oxide/metal interface will be rich in chromium oxide. Conversely, if the mobility of the chromium ion is the larger of the two, the concentration of chromium oxide near the outer surface will be above the average concentration. At temperatures above 1000°C no accumulation of Cr_2O_3 near the

outer surface is in fact observed, but this may be because Cr_2O_3 is rather volatile. Concerning this problem an electron diffraction and electron microscope study carried out by GULBRANSEN, PHELPS and HICKMAN[25] is of importance. The quantitative treatment of the different mobilities of the divalent nickel ions and trivalent chromium ions opens up new possibilities of understanding many of the rather complicated oxidation processes of alloys.

The oxide layer formed on nickel-chromium alloys with 10–20 per cent Cr is composed either of a spinel ($> 1000°C$)[7a, 60] or of Cr_2O_3 ($< 1000°C$)[25, 56]. In addition the greater mobility of chromium ions in spinel and NiO-Cr_2O_3-layers observed by HAUFFE and PSCHERA[44] seems to be of importance for the formation of oxide layers. In the case of the oxidation of iron-aluminium alloys an increase of Al_2O_3 in the oxide layer was found by KORNILOW and SCHIPIKELMANN.[55] An extensive discussion of this matter was given by HAUFFE.[36]

THE FORMATION OF MULTI-LAYERED SCALE CONTAINING VARIOUS METAL OXIDE PHASES

One of the important factors determining the resistance of a metal or alloy to oxidation at high temperatures is the structure of the oxide layer formed at the surface. An extensive investigation of structure has been carried out by GULBRANSEN and associates.[22–26] Applying the electron diffraction reflection technique they could observe the structure of the oxide layer *in situ* at temperatures at which it is actually protective. The results of these measurements are of particular importance for the class of oxidation processes in which, from a thermodynamic point of view, two or three oxide layers can be expected.

One of the most thoroughly investigated processes is the oxidation of iron at different temperatures and pressures of oxygen. The first relatively complete studies of the oxidation of iron at elevated and low temperatures was carried out by JACKSON and QUARRELL[48] and FISCHBECK[16] and NELSON.[66] They recognized that the formation of different oxides and their structural transformations are functions of time as well as temperature. They found that below 400°C the stable oxide on iron was Fe_3O_4, and that FeO was formed between 515 and 900°C. Between 400 and 515°C, FeO, Fe_3O_4 or a mixture of the two oxides were found. JIMORI[50] could distinguish between the cubic Fe_2O_3 and Fe_3O_4 by means of x-ray analysis; by electron diffraction this would have been impossible. From x-ray evidence JIMORI postulated a transition of γ-Fe_2O_3 with cubic spinel structure to α-Fe_2O_3 with rhombohedral structure at 340°C. GULBRANSEN and HICKMAN found a lower transition point at about 225°C. For long periods of oxidation at up to 400°C they observed the formation of α-Fe_2O_3. Alongside this evidence we should note that for thick films formed at 500°C the

inverse spinel Fe_3O_4 is the only oxide observed. Finally, we have studies by BÉNARD and associates[4] of the influence of temperature on the kinetics of formation of different iron oxide layers (see Fig. 22). Thus

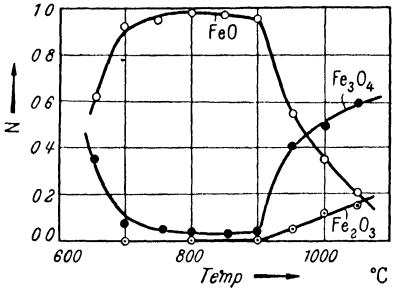


Fig. 22. Influence of temperature on the kinetics of formation of different iron oxide layers, as measured by BÉNARD and COQUELLE. (N = mole fraction of FeO, Fe_3O_4 and Fe_2O_3, respectively)

among all these results there seem to be many disagreements; they are due to difficulties which may be listed as difficulties of

(*a*) reproducibility,
(*b*) accurate measurement of the surface temperature,
(*c*) accurate determination of the time of oxidation,
(*d*) controlling the film thickness, and
(*e*) controlling the pressure of oxygen.

Recently BIRCHENALL and associates[8a] have investigated the mechanism and the kinetics of the scaling of iron under the pressure of 1 atm of oxygen between 400 and 1200°C. The results indicate that in no case, above a temperature of about 625°C, does the layer contain more than 10 per cent by volume of the higher oxides Fe_3O_4 and Fe_2O_3

TABLE 4

Approximate percentages of FeO, Fe_3O_4 and Fe_2O_3 in scale formed on iron oxidized in oxygen; according to DAVIES, SIMNAD and BIRCHENALL

T°C	t min.	FeO *pct*	Fe_3O_4 *pct*	Fe_2O_3 *pct*
600	100	50	50	—
625	110	70	30	—
700	225	98	2	—
750	100	96	4	—
800	100	96	2	2
900	100	90	9	< 1
1000	25	95	4	< 1
1090	25	94	5	< 1

(see Table 4). To measure the rate of growth of Fe_3O_4 and Fe_2O_3, samples of FeO and Fe_3O_4 were prepared and oxidized in oxygen. The values for the parabolic rate constants are given in Table 5. Below

TABLE 5

Values for the Parabolic Rate Constants k (g cm^{-2} $min^{-1/2}$) of the Oxidation of Iron in Oxygen; from DAVIES, SIMNAD and BIRCHENALL

T°C	k	*Remarks*
400	$3{\cdot}4 \cdot 10^{-6}$	Oxidation of Fe to form multilayer scales
500	$5{\cdot}0 \cdot 10^{-5}$	
600	$9{\cdot}0 \cdot 10^{-5}$	
700	$6{\cdot}0 \cdot 10^{-4}$	
800	$1{\cdot}85 \cdot 10^{-3}$	
900	$3{\cdot}9 \cdot 10^{-3}$	
1035	$9{\cdot}0 \cdot 10^{-3}$	
1130	$1{\cdot}24 \cdot 10^{-2}$	
1200	$1{\cdot}87 \cdot 10^{-2}$	
850	$2{\cdot}3 \cdot 10^{-4}$	Oxidation of FeO
910	$3{\cdot}3 \cdot 10^{-4}$	
930	$3{\cdot}9 \cdot 10^{-4}$	
950	$4{\cdot}9 \cdot 10^{-4}$	
800	$4{\cdot}5 \cdot 10^{-5}$	Oxidation of Fe_3O_4
900	$1{\cdot}5 \cdot 10^{-4}$	
988	$3{\cdot}0 \cdot 10^{-4}$	

570°C the oxide layer formed is almost pure Fe_3O_4. The reactions involved, neglecting the formation of Fe_2O_3, are:

$$3Fe + 2O_2 \rightleftharpoons Fe_3O_4 \quad \ldots\ldots (56a)$$

$$3FeO + \tfrac{1}{2}O_2 \rightleftharpoons Fe_3O_4. \quad \ldots\ldots (56b)$$

BIRCHENALL showed that, for a given thickness of Fe_3O_4, the increase in weight for process (56*a*) is four times as great as for process (56*b*). Now BIRCHENALL and associates tentatively suggest that the growth of each layer is a diffusion controlled process and that the interface reaction rates may be neglected. Applying a modified Pfeil technique using radioactive silver Ag^{110} as inert markers they find a very interesting diagram of the oxidation process of iron (see Fig. 23). Whereas for the formation of FeO the rate determining step is the migration of Fe^{2+}-ions, in hematite it is not the iron ions that wander but the oxygen ions. In magnetite, however, both ions migrate. These results are consistent with the semi-conducting properties of the oxides of iron.

On oxidizing copper at 1000°C at an oxygen pressure of about 100 mm Hg, an outer and an inner layer of CuO and Cu_2O, respectively, are formed. Just as the molar volume of Cu_2O is greater than that of

metallic copper, so the molar volume of CuO is greater than that of Cu_2O (1 g — atom of Cu = 7·1 cm³; ½ mole of Cu_2O = 11·7 cm³; 1 mole of CuO = 12·3 cm³). The cupric oxide layer on the cuprous oxide layer is, therefore, compact. Under these conditions PILLING and BEDWORTH,[68] and also FEITKNECHT[15] and FROEHLICH[19] found that the rate of oxidation of copper is independent of the oxygen pressure. This experimental observation is easily understood. The excess of

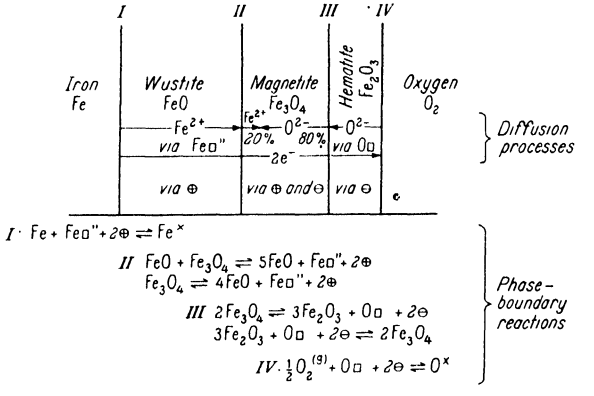


Fig. 23. Schematic diagram of the diffusion of ions across the different oxide layers during the oxidation of iron on the basis of the experimental results of DAVIES, SIMNAD and BIRCHENALL

oxygen in the Cu_2O phase at the Cu_2O/CuO interface is determined by the two-phase equilibrium Cu_2O + CuO, as is also, therefore, the concentration gradient of copper ion vacancies in the cuprous oxide layer. On the basis of the rate of oxidation of copper to Cu_2O, and of Cu_2O to CuO, WAGNER[88] suggested that the ratio by weight of the two oxides formed (at 1000°C and an oxygen pressure of 1 atm) is given by:

$$\frac{\Delta x_{CuO}}{\Delta x_{Cu_2O}} = \frac{k_{rCuO}}{k_{rCu_2O}} = \frac{2{\cdot}5 \,.\, 10^{-12}}{6{\cdot}5 \,.\, 10^{-9}} = 0{\cdot}4 \,.\, 10^{-3}$$

i.e. under the above experimental conditions practically pure cuprous oxide is formed, only the outer surface, representing 0·1 per cent of the total thickness of the film, consisting of cupric oxide. It can be assumed that the sulphuration of iron at 670°C and a sulphur vapour pressure of 1 atm gives to an iron sulphide layer of analogous composition. Here the rate of sulphuration of iron to FeS is a thousand times faster than that of FeS to FeS_2, as was recently shown by HAUFFE and RAHMEL[37] ($k_{FeS} = 3{\cdot}3 \,.\, 10^{-8}$ and $k_{FeS_2} = 1{\cdot}4 \,.\, 10^{-11}$ (equ. $cm^{-1}\ s^{-1}$)). A great number of metals and alloys have been investigated from the same point of view, in particular by GULBRANSEN and associates.[22–26]

The Formation of Thin Compact Oxide Films on Metals

Besides Wagner's theory Mott and Cabrera[5] have recently put forward another theory of the oxidation of metals. Wagner's formula is valid only if the oxide layer is so thick that the concentrations of interstitial ions (or ion vacancies and free electrons, or electron holes) are equal to one another throughout most of its thickness. Mott and Cabrera derive a formula especially for the rate of oxidation in thin and very thin oxide films with a thickness between 10 and 1000 Å. Applying the fundamental assumptions of the Schottky-Mott boundary layer theory[72, 64] (Randschicht-Theorie) of crystal rectifiers, Mott[65] was the first to show that in the vicinity of the metal/oxide interface there is a diffuse electrical double layer because the electrons move more quickly to the adsorbed oxygen atoms than the ions. This process plays an important role if the oxide film is thinner than 100 Å. In such a thin film the electrical field strength is so high that the rate of migration of ions is no longer proportional to the field strength (i.e. it is no longer inversely proportional to the thickness of the oxide film) but depends quasi-exponentially upon the thickness.

Mott and Cabrera theoretically gave different conditions requiring the logarithmic law, the parabolic law ($x^2 = 2kt$), the cubic law ($x^3 = 3kt$, experimentally confirmed by Campbell and Thomas[6] for the oxidation of copper), and lastly the linear law ($x = kt$, experimentally confirmed by Gulbransen and Wysong[26] for the oxidation of aluminium). Wagner[90] recently pointed out that the possibilities are not yet exhausted. Whereas Mott's theory of oxidation processes assumes one double layer at the oxide/metal interface, Wagner expects during the growth of an oxide film an overlapping of two double layers, the one at the metal/oxide interface and the other at the oxide/oxygen interface. Thus another interpretation of the logarithmic law seems possible, especially for films with a thickness greater than 100 Å.

Moore and Lee's experimental results[62] for the oxidation of zinc, which cannot be explained either by the theory of Mott-Cabrera or by Wagner's theory, led Engell and Hauffe[12] to give an explanation which might be of general interest and will, therefore, be described in some detail. For layers of up to 1000 Å Moore and Lee found that the parabolic law of growth (and at the same time the rate of growth) was dependent on the pressure of oxygen, whereas Wagner's theory predicts a rate of growth independent of the oxygen pressure. Wagner and Gruenewald had, indeed, found for thick layers ($>$ 1000 Å) a parabolic growth with the rate of growth independent of oxygen pressure. Using the Schottky-Spenke boundary-layer theory of crystal rectifiers, Engell and Hauffe advance an explanation of the above observations.

During the formation of the oxide layer two regions must be distinguished. As shown by Fig. 24 there is a small region A in which an electric field has been set up because of a depletion of quasi-free electrons which, owing to their relatively great mobility, have been captured by the chemisorbed oxygen atoms on the surface. In this region the migration of the interstitial zinc ions is mainly due to this electric field, while in the region B the motion of the interstitial zinc ions is mainly caused by the concentration gradient. The thickness l they call the critical thickness.

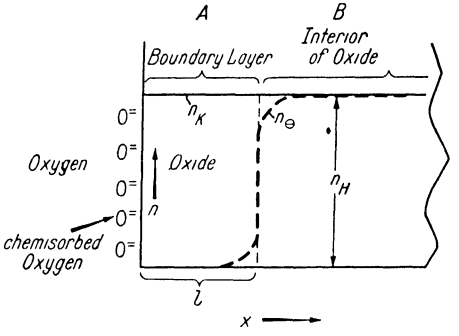


Fig. 24. Scheme of the distribution of quasi-free electrons and interstitial zinc ions in the boundary-layer (region A) and in the interior of the semi-conductor (region B). (l is the critical thickness of the boundary layer)

For both regions of the oxide layer the current of quasi-free electrons $I_\ominus$ and of interstitial zinc ions I_K is given by the usual equations.

$$I_\ominus = D_\ominus \frac{\partial n_\ominus}{\partial x} - E(x) \,.\, n_\ominus \,.\, B_\ominus \qquad \ldots\ldots (57a)$$

$$I_K = D_K \frac{\partial n_K}{\partial x} + v_K \,.\, E(x) n_K \,.\, B_K \qquad \ldots\ldots (57b)$$

Here B is the mobility in $\mathrm{cm^2 \,.\, V^{-1}\, s^{-1}}$, n_K and $n_\ominus$ the concentration of interstitial metal ions and quasi-free electrons respectively, v_K the charge of the metal ions in interstitial sites and $E(x)$ the electric field strength in $\mathrm{V \,.\, cm^{-1}}$ at the point x. Equations (57a), (57b) and the Boltzmann equation

$$\frac{n_R}{n_H} = \exp\left(-\frac{V_D e}{kT}\right) \qquad \ldots\ldots (58)$$

exactly determine the distribution of positive and negative charges in the boundary layer. Here V_D is the diffusion potential, n_R and n_H the concentration of quasi-free electrons in the boundary layer and in the interior of the oxide layer.

The system is stationary, if the concentration of interstitial cations in the boundary layer is such that the divergence of the current of cations moving across the whole layer (A and B) is zero. Since the electric field may be neglected in the interior of the oxide ($E = 0$ for $x \geq l$) we obtain, neglecting the transport of cations in the boundary layer by diffusion,

$$I_K = \left(D_K \frac{\partial n_K}{\partial x}\right)_{x \geq l} = (v_K n_K(x) B_K \,.\, E(x))_{x < l} \quad . \quad . \quad . \quad . \quad (59)$$

With $D_K = B_K \boldsymbol{k}T/\boldsymbol{e}$ and Poisson's equation $\partial E/\partial x = \frac{4\pi \boldsymbol{e} v_K}{\varepsilon} n_K(x)$, ε being the permittivity, we obtain on integrating between the limits $x = 0$ and $x = l$ for the concentration of interstitial cations in the boundary layer

$$n_K(x) = \left\{\frac{\varepsilon}{8\pi \boldsymbol{e}} \cdot \frac{kT}{v_K{}^2 \boldsymbol{e}} \cdot \left(\frac{\partial n_K}{\partial x}\right)_{x \geq l}\right\}^{1/2} (l - x)^{1/2} \quad . \quad . \quad . \quad . \quad (60)$$

By means of a second integration from 0 to l we finally obtain for the thickness of the boundary layer

$$l = \text{const} \left\{\frac{kT}{\boldsymbol{e}} \left(\frac{\partial n_K}{\partial x}\right)_{x \geq l}\right\}^{-1/3} V_D{}^{2/3} \quad . \quad . \quad . \quad . \quad (61)$$

For a critical discussion of these equations see ENGELL, HAUFFE and SCHOTTKY.[13]

Using the equation for the chemisorption of oxygen on zinc oxide

$$\tfrac{1}{2}O_2{}^{(g)} + 2\ominus^{(II)} \rightleftharpoons O^{2-(A)} \quad . \quad . \quad . \quad . \quad (62)$$

HAUFFE and ENGELL[42] finally obtain for $n_{O^{2-}}{}^{(A)}$

$$n_{O^{2-}}^{(A)} = \frac{1}{4}\left\{\frac{\varepsilon}{\pi e} \cdot n_H \cdot \ln \frac{p_{O_2}{}^{1/2} n_H{}^2 K}{n_{O^{2-}}^{(A)}}\right\}^{1/2} \quad . \quad . \quad . \quad . \quad (63)$$

Therefore $n_{O^{2-}}^{(A)}$, the number of chemisorbed oxygen atoms, increases logarithmically with increasing partial pressure of oxygen. All these formulae are only valid, if there is an "exhaustion layer" (the layer A of Fig. 24). They seem to be approximately true in zinc oxide up to 300°C.

If the entire thickness of the oxide layer is less than l the rate of oxidation is dependent on the partial pressure of oxygen even though the diffusion of metal ions to the surface is rate-determining. The reasoning is as follows.

At the beginning of the oxidation and for a constant partial pressure of oxygen, which leads to a constant diffusion potential V_D the electric field strength $E(x_1)$ will decrease with increasing thickness x and we have $E(x_1) > E(l)$, if $x_1 < l$. Now it may be assumed that the majority of the electrons arrive at the surface of the oxide, initially, by means of

the tunnel effect and later, for a thicker boundary layer, by means of thermionic emission (cf. SCHOTTKY and MOTT-CABRERA). If their concentration is n we can put $n \gg n_H$, and n is practically independent on the thickness of the layer x_1, if $x_1 < l$. Electrons reaching the surface must do so against the electric field.

Let N be the maximum number of surface sites per square centimetre available to chemisorbed oxygen atoms, and K be a constant. We obtain:

$$\frac{n_{O^{2-}}^{(A)}}{N - n_{O^{2-}}^{(A)}} = K \cdot p_{O_2}^{1/2} \cdot n^2 \exp\left(-2E(x_1)\frac{x_1 e}{kT}\right) \quad . \quad . \quad . \quad . \quad (65)$$

For small surface concentrations ($n_{O_2^-}^{(A)} \ll N$) we have

$$E(x_1) = \frac{4\pi e}{\varepsilon} \cdot 2n_{O^{2-}}^{(A)} = K' \cdot p_{O_2}^{1/2} \cdot n^2 \cdot \exp\left(-2E(x_1)\frac{x_1 e}{kT}\right) \quad . \quad (66)$$

where

$$K' = K \cdot \frac{8\pi e}{\varepsilon} N.$$

We finally obtain

$$E(x_1) = \frac{kT}{4x_1 e} \ln(p_{O_2} \cdot n^4 \cdot K'^2) - \ln E(x_1). \quad . \quad . \quad . \quad . \quad (67)$$

In this equation $\ln E(x_1) \ll E(x_1)$, since $E(x_1)$ is about 10^5 V/cm. If there is a strong electric field the current due to diffusion may be neglected compared with the current due to the electric field and we obtain from (57*b*)

$$I_K = n_K \cdot B_K \cdot E(x_1) \nu_K. \quad . \quad . \quad . \quad . \quad (68)$$

Combining (57*b*), (67) and (68) we finally have for the current of cations (i.e. the number of the cations entering the surface per square centimetre per second),

$$I_K = n_K \cdot B_K \cdot \nu_K \cdot \frac{kT}{4x_1 e} \ln(p_{O_2} \cdot n^4 \cdot K^{12}) \quad . \quad . \quad . \quad . \quad (69)$$

Comparison of (57*b*) and (66) shows that the rate of formation of thin films is proportional to the surface concentration of chemisorbed oxygen ($I_K \propto E(x) \propto n_{O^{2-}}^{(A)}$).

If the transport of cations to the surface determines the rate of oxidation of the metal, equation (69) gives the following rate law

$$\frac{dx}{dt} = \frac{k}{x} \quad . \quad . \quad . \quad . \quad (70)$$

where

$$k = \frac{Vm}{N} \cdot B_K \cdot \nu_K \cdot \frac{kT}{4e}(\ln p_{O_2} + \ln(n^4 K^{12}))$$

and V_m and N are the molar volume of the oxide and Avogadro's number, respectively. Equation (70) shows that the rate constant of oxidation k is a logarithmic function of the partial pressure of oxygen

$$k = \text{const} \cdot \ln p_{O_2} + \text{const}' \qquad \ldots . \quad (71)$$

The validity of equations (70) and (71) was experimentally confirmed for the oxidation of zinc by MOORE and LEE. As shown in Figs. 25 and

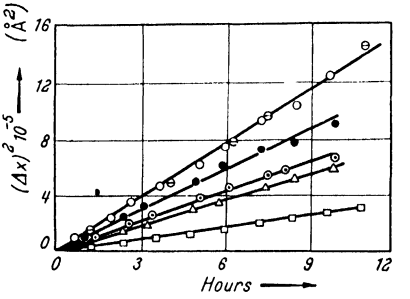


Fig. 25. Oxidation of zinc at 390°C, in parabolic plot, according to MOORE and LEE

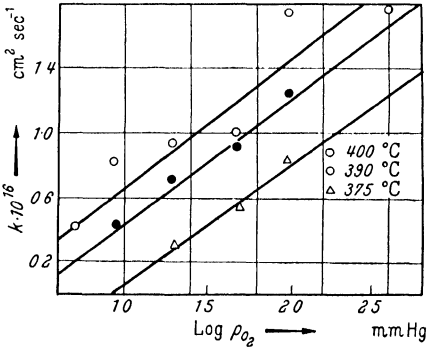


Fig. 26. Logarithmic dependence of the parabolic rate constant k of the oxidation of zinc on the partial pressure of oxygen, according to ENGELL and HAUFFE, on the basis of measurements carried out by MOORE and LEE

26 the rate of oxidation follows the parabolic rate law and k is a logarithmic function of the partial pressure of oxygen, as is to be expected according to equation (71).

WAGNER and GRUENEWALD also investigated the influence of the partial pressure of oxygen on the rate of oxidation of zinc and found k to be independent of p_{O_2}. This result is in good agreement with WAGNER'S theory of oxidation as well as with the views advanced in this article, for in WAGNER'S experiments the condition $x \gg l$ was fulfilled.

This article does not aim at giving a complete review of all important experimental facts concerning the oxidation of metals and alloys. For example, there has been no reference to the interesting experiments on the internal oxidation of copper and silver alloys carried out by RHINES and associates[69, 70] and MEIJERING and RATHENAU.[59] The main purpose has been to stimulate the metal expert by giving him some of the new theoretical points of view on this matter. I would like to express my gratitude to DR. R. LANDSBERG for numerous stimulating discussions, and for the linguistic help he has given me.

REFERENCES

[1] BARDEEN, J., BRATTAIN, W. H. and SHOCKLEY, W. *J. chem. Phys.* 14 (1946) 714

[2] VON BAUMBACH, H. H. and WAGNER, C. *Z. phys. Chem.* 22 (1933) 199

[3] —— —— *Z. phys. Chem.* (B) 24 (1934) 59

[4] BÉNARD, J. and COQUELLE, O. *Compt. rend.* 222 (1946) 796, 884; BÉNARD, J. *Bull. Soc. Chim.* France (1949) 73

[5] CABRERA, N. and MOTT, N. F. *Rep. Progress Phys.* 12 (1949) 163

[6] CAMPBELL, W. E. and THOMAS, U. B. *Trans. Elektrochem. Soc.* 91 (1947) 345

[7] CASTELLAN, G. W. and MOORE, W. J. *J. chem. Phys.* 17 (1949) 41

[7a] CHALMERS, B. and QUARRELL, A. G. *Physical Examination of Metals*, p. 225, London 1941

[8] DARKEN, L. S. and GURRY, R. W. *J. Amer. Chem. Soc.* 67 (1945) 1398

[8a] DAVIES, M. H., SIMNAD, M. T. and BIRCHENALL, C. E. *Trans. Amer. Inst. Min. Met. Engrs.* (1951) 889

[9] DÜNWALD, H. and WAGNER, C. *Z. phys. Chem.* (B) 17 (1932) 467

[10] EARLE, M. D. *Phys. Rev.* 61 (1942) 61

[11] EHRLICH, P. *Z. Elektrochem* 45 (1939) 362

[11a] EVANS, U. R. *Metallic Corrosion*, London (1948)

[12] ENGELL, H. J. and HAUFFE, K. *Metall*, 6 (1952) 285

[13] —— —— and SCHOTTKY, W. *Z. phys. Chem.*, in press

[14] FAIVRE R. and MICHEL, A. *Compt. rend.* 207 (1938) 159

[15] FEITKNECHT, W. *Z. Elektrochem* 35 (1929) 142

[16] FISCHBECK, K. and SALZER, F. *Metallwirtschaft* 14 (1935) 733 753

[17] FRANK, F. C. and MERWE, J. VAN DER *Proc. Roy. Soc.* (A) 198 (1949) 203 and 216; (A) 200 (1950) 125

[18] FRENKEL, J. *Z. Physik* 35 (1926) 652

[19] FROEHLICH, K. W. *Z. Metallk.* 28 (1936) 368

[20] GENSCH, CH. and HAUFFE, K. *Z. phys. Chem.* 196 (1950) 427

[21] —— —— *Z. phys. Chem.* 195 (1950) 386

[22] GULBRANSEN, E. A. and ANDREW, K. F. *Trans. Amer. Inst. Min. Met. Engrs., Metals Trans.* 185 (1949) 741

[23] —— and HICKMAN, J. W. *Trans. Amer. Inst. Min. Met. Engrs., Techn. Publ.* No. 2068 (1946)

[24] —— and RUKA, R. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 1500

[25] ——, PHELPS, R. T. and HICKMAN, J. W. *Ind. and Eng. Chem.* (Anal. Edit.) 18 (1946) 640

[26] —— and WYSONG, W. S. *J. Phys.* and *Colloid. Chem.* 51 (1947) 1087

[27] GUNDERMANN, J. HAUFFE, K. and WAGNER, C. *Z. phys. Chem.* (B) 37 (1937) 148

[28] —— and WAGNER, C. *Z. phys. Chem.* (B) 37 (1937) 155

[29] HARTMANN, W. *Z. Physik* 102 (1936) 709

[30] HAUFFE, K. *Metalloberfläche* 3A (1949) 213; 5A (1951) 1
[31] —— and GENSCH, CH. *Z. phys. Chem.* 195 (1950) 116, 386
[32] See, e.g., HAUFFE, K. Lattice defects and conductivity of ionic and electronic solid conductors, in *Ergeb. exakt. Naturwiss.* 25 (1951) 193
[33] HAUFFE, K. and BLOCK, J. *Z. phys. Chem.* (Schottky-Festband) 198 (1951) 232
[34] ——, GRUNEWALD, H. and GREESE, R. *Z. phys. Chem.*, in press
[35] —— —— *Z. phys. Chem.* (Schottky-Festband) 198 (1951) 248
[36] —— *Z. Metallk.* 42 (1951) 34
[37] —— and RAHMEL, A. *Z. phys. Chem.* 199 (1952) 152
[38] —— *Ann. Phys.* (6) 8 (1950) 201
[39] —— and PFEIFFER, H. *Z. Elektrochem.*, 56 (1952) 390
[40] —— and VIERK, A. L. *Z. phys. Chem.* 196 (1950) 160
[41] —— and PFEIFFER, H. *Z. Metallk.*, in press
[42] —— and ENGELL, H. J. *Z. Elektrochem.*, 56 (1952) 366
[43] —— and FLINT, G. *Z. phys. Chem.*, 200 (1952) 199
[44] —— and PSCHERA, K. *Z. anorg. allgem. Chem.* 262 (1950) 147
[45] HICKMAN, J. W. and GULBRANSEN, E. A. *Trans. Amer. Inst. Min. Met. Engrs., Techn. Publ.* No. 2144 (1947)
[46] —— —— *Analyt. Chem.* 20 (1948) 158; *J. Phys. and Colloid Chem.* 52 (1948) 1186; *Trans. Amer. Inst. Min. Met. Engrs., Techn. Publ.* No. 2144 (1947); HICKMAN, J. W. *Trans. Amer. Inst. Min. Met. Engrs., Techn. Publ.* No. 2483
[47] HIMMLER, W. *Z. phys. Chem.* 195 (1950) 129
[48] JACKSON, G. and QUARRELL, A. G. *Iron and Steel Inst.*, Special Report No. 24 (1939) 65
[49] JETTE, E. R. and FOOTE, F. *J. Chem. Phys.* 1 (1933) 29
[50] JIMORI, J. *Sci Papers Inst, Phys. Chem. Res.* (Tokyo) 34 (1937) 60
[51] JOST, W. *J. Chem. Phys.* 1, 466 (1933), *Z. phys. Chem.* (A) 169 (1934) 129
[52] —— *Trans. Faraday Soc.* 34 (1938) 860
[53] —— and NEHLEP, G. *Z. phys. Chem.* (B) 32 (1936) 1; 34 (1936) 348
[54] KOCH, E and WAGNER, C. *Z. phys. Chem.* (B) 38 (1937) 295
[55] KORNILOW, J. J. and SCHPIKELMANN, A. J. *Ber. Akad. Wiss. U.S.S.R.* 53 (1946) 813; see also, ARCHAROW, W. J. *Bull. Acad. Sci. U.S.S.R.* (1946) 127
[56] LUSTMAN, B. *Iron Coal Trades Rev.* 154 (1947) 889
[57] MANSFIELD, R. *Proc. Phys. Soc.* (B) 62 (1949) 476
[58] MERWE, J. VAN DER *Discuss. Faraday Soc.* 5 (1949) 201
[59] MEIJERING, J. L. and RATHENAU, G. W. *Philips, Res. Rep.* 2 (1947) 81
[60] MIYAKI, S. *Sci. Papers Inst. Phys. Chem. Res.* Tohoku 31 (1937) 161
[61] MILLER, P. *Phys. Rev.* 60 (1941) 890
[62] MOORE, W. J. and LEE, J. K. *Trans. Faraday Soc.* 47 (1951) 501
[63] MOTT, N. F. and LITTLETON, M. J. *Trans. Faraday Soc.* 34 (1938) 485
[64] —— *Proc. Roy. Soc.* A 171 (1939) 944
[65] —— *Trans. Faraday Soc.* 35 (1939) 1175; 36 (1940) 472; 43 (1947) 429
[66] NELSON, B. J. *J. chem. Phys.* 6 (1938) 606
[67] PILLING, N. B. and BEDWORTH, R. E. *J. Inst. Metals* 29 (1923) 529
[68] —— —— *Ind. and Eng. Chem.* 17 (1925) 372
[69] RHINES, F. N. *Trans. Amer. Inst. Min. Met. Engrs.* 137 (1940) 246; RHINES, F. N. and GROBE, A. H. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 318; RHINES, F. N., JOHNSON, W. A. and ANDERSON, W. A. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 205
[70] —— *J. Corrosion and Material Protection* 5 (1947) 15

[71] SCHOTTKY, W. *Z. phys. Chem.* (B) 29 (1935) 335
[72] —— *Z. Physik* 113 (1939) 367; 118 (1942) 539; SCHOTTKY, W. and SPENKE, E. *Wiss. Veröffentl. Siemens-Werke* 18 (1939) 225
[73] TAMMANN, G. *Z. anorg. allgem. Chem.* 111 (1920) 78
[74] TEITOW, J. *Ann. Physik* 6 (1949) 63, 71
[75] TUBANDT, C. and EGGERT, S. *Z. anorg. allgem. Chem.* 110 (1920) 196
[76] —— *Z. anorg. allgem. Chem.* 115 (1921) 105
[77] VERNON, W. H. J., AKEROYD, E. J. and STROUD, E. G. *Ind. Metals* 65 (1939) 301
[78] VERWEY, E. J. W., HAAYMAN, P. W. and ROMEYN, F. C. *Chem. Weekblad* 44 (1948) 705
[79] WAGNER, C. *Z. phys. Chem.* (B) 21 (1933) 25
[80] —— *Z. phys. Chem.* (B) 32 (1936) 447
[81] —— Diffusion and High Temperature Oxidation of Metals in *Atom Movements*, p. 153 Ohio (1951)
[82] —— *J. Corrosion and Material Protection* 5 No. 5 (1947)
[83] —— and ZIMENS, K. E. *Acta Chem. Scand.* 1 (1947) 547
[84] —— and HAMMEN, H. *Z. phys. Chem.* (B) 40 (1938) 197
[85] —— and KOCH, E. *Z. phys. Chem.* (B) 32 (1936) 439
[86] —— and GRUENEWALD, K. *Z. phys. Chem.* (B) 40 (1938) 455
[87] —— *J. chem. Phys.* 18 (1950) 62
[88] —— Chemische Reaktionen der Metalle, in *Handb. der Metallphysik*, Vol. 1, Part 2, p. 143, Leipzig 1940
[89] —— and ZIMENS, K. E. *Acta Chem. Scand.* 1 (1947) 539
[90] —— *Methods of High Temperature Oxidation Testing and Evalution of Observations*, p. 110 Cleveland 1951

3

GASES IN METALS

C. R. Cupp

THE important relationships between metals and the elements normally thought of as gases has been under consideration for many years. In earlier times studies were confined mainly to solubility, diffusion, and embrittlement problems and casting defects caused by gases; now more subtle effects on the physical and mechanical properties of the metals are being considered. There are several books[1–5] and reviews[6–11] dealing with the theme of this paper. Other reviews of a more specialized nature will be referred to throughout. The present paper is an attempt to summarize the more important experimental facts and to mention the significance of these as interpreted by various observers. In keeping with the purpose of this series, particular reference will be made to more recent observations, but the results of earlier work must, at the same time, be given due recognition.

As in an earlier review,[9] it should be pointed out that only those elements will be considered which exist as gases in their uncombined states, at normal temperature and pressure, namely, hydrogen, oxygen, nitrogen, the inert gases and the halogen gases. As a supplement to the list of references, a number of papers dealing with such combined gases as sulphur dioxide, carbon monoxide, carbon dioxide, hydrogen sulphide, etc., will be given, but no consideration of these will be made here. It is felt that purely adsorption phenomena, and corrosion effects due to gases (including surface oxidation) are not part of the subject at hand. Commercial degassing practice will not be reviewed, although some mention will be made of its necessity in various operations. Experimental methods will not be described in any detail; the reader is referred to the original literature for this information. This paper will deal with the solution and diffusion of gases in solid metals; with the absorption, retention, and evolution of gases by metals; and with effects upon the properties of the solid metals. The consideration of the marked effect of gases on many of the physical properties of metals will be the main theme of this work.

GAS-METAL RELATIONSHIPS

FAST has argued[6] that the permeability of metals to gases is related to the ability of a given gas to combine with a given metal, diffusion rates being high where compound formation is possible. RHINES[9] has amended this statement to say that diffusion is appreciable in those

gas-metal couples in which the gas is soluble in the metal, to a significant degree. SMITH[2] has made a more pointed generalization, in the case of hydrogen at least, by emphasizing that increased permeability of the metal, however produced, is always accompanied by an increase in the "occlusive capacity" of the metal for hydrogen. At any rate if one considers that the halogen gases, although they react actively with practically all metals, have been found not to diffuse to any marked degree through any metals, or to dissolve in appreciable amounts in solid metals, then it must be assumed that FAST'S[6] statement is not correct.

Inert Gases

Turning to the role of inert gases, there can be no compound formation and the possibility of solubility is questionable. It has been reported recently[12] that the amount of argon absorbed by uranium at room temperature is less than 1 part in 3×10^5 by weight and at 600°C, less than 1 part in 3×10^4, which is the magnitude of the uncertainties of measurements of the gas volume. Elsewhere[13] it is indicated that inert gases are not dissolved to any measurable extent in solid metals. Similarly the evidence of diffusion of inert gases through metals is not well confirmed. The earliest report found in the literature[14] showed that there is no permeability of metals to inert gases. From preliminary investigations[15] of the removal of inert gases in discharge tubes at low pressures it was erroneously concluded that the gas penetrates into the cathode as ions, while the cathode metal is disintegrated and redeposited by diffusion. Later work of the same nature showed quite conclusively[16] that the apparent loss of neon from lamp tubes was due not to the passage of the gas through iron either at room temperature or at a bright red heat, but to an apparent adsorption upon the iron under the influence of an electric discharge. As a very brief summary of the work by LUMPE and SEELIGER[16] one would say that they found that, in the discharge tube, the inert gas particles were "shot into" ("hineingeschossen") the cathode and there "remained stuck" ("stecken bleiben") in some sort of mechanical trap. A curve showing the distribution of the particles indicated a greater concentration at the surface. On heating to a few hundred degrees Celsius the gas particles are released and "diffuse" ("diffundieren") out of the cathode. No claim was made that there was any true equilibrium solubility of the inert gas in the iron cathode, and it was not suggested that diffusion proceeded to cause a uniform distribution of the gas throughout the cathode.

One other experiment, of doubtful value,[17] indicated that at 505°C and a gas pressure of 72 cm Hg the diffusion of helium through aluminium was as rapid as that of hydrogen. The measured rate was $0{\cdot}026 \times 10^{-7}$ ml N.T.P. × mm thickness per sq cm heated area per

second. Since this anomaly was not investigated further it must remain subject to doubt.

Nitrogen

In his recent review RHINES[9] points out that the compound rule proposed by FAST[6] seems to hold in a general way for oxygen, nitrogen and hydrogen, but the correlation between solubility and diffusion is better. He states further:

"Nitrogen forms compounds with, is soluble in, and diffuses through, most of the metals of the last four groups of the periodic table of the elements. It is generally nonreactive with, insoluble in, and fails to diffuse significantly through, the first-group metals and only sparingly, when at all, through the metals of the second and third groups."

At an earlier time SMITHELLS[1] made the following generalization:

"Nitrogen is only soluble in those metals which are capable of forming nitrides. We have previously noted the same distinction both in activated adsorption and diffusion, and the close connexion between the three kinds of gas-metal equilibria is again emphasized. Nitrogen has been shown to be insoluble within the limits of the experimental method in cobalt,[18*] copper,[19*] silver,[20*] and gold.[21*] Those metals which, like iron and molybdenum, absorb nitrogen, show the same kind of changes in lattice structure, mechanical and electrical properties as are shown when the hydride-forming metals absorb hydrogen. The nitrogen-metal systems have not been investigated so completely as those of hydrogen, but the close parallel between them is evident from the data."

Later SMITHELLS[7] stated that "nitrogen is soluble in a limited number of metals and although there is evidence of simple solubility in some cases, all metals which dissolve this gas are known to form definite nitrides under suitable conditions."

An examination of the literature following this brief statement will indicate that SMITHELLS' generalization is probably closer to the truth than RHINES' statement. It is recognized that some of the errors in RHINES' statement may be due to the type of periodic table being used. Nitrogen is generally soluble in groups VA, VIA and VIIA but (except for iron) not VIII, if those are to be taken as the "last four groups." It is also soluble in group IVA. RHINES is quite correct about nitrogen's relation with groups IA and IB. There is some ambiguity in his statement regarding the second and third group elements, where nitrogen reacts with several elements to form nitrides but has not been reported to dissolve in the same elements. SMITHELLS' statement seems to be more generally acceptable.

On the basis of the two statements above, it seems worthwhile to

* This number refers to the list of references at the end of the present paper.

TABLE 1

		Transition Heavy Metals								*Heavy Metals*		*Non Metals*			
IA	IIA											IIIB	IVB	V	VI
Li^n	Be^n											B	C	N	O
Na	Mg^n											Al^{n*}	Si	P	S
		IIIA	IVA	VA	VIA	VIIA	VIII			IE	IIB	IIIB	IVB	VB	VIB
K	Ca^n														
Rb	Sr	Sc^n	Ti^{nx}	V^{nx}	Cr^{nx}	Mn^{nx}	Fe^{nx}	Co*	Ni*	Cu*	Zn^{n*}	Ga^n	Ge^n	As	Se
Cs	Ba^n	Y	Zr^{nx}	Nb^{nx}	Mo^{nx}	Tc	Ru	Rh	Pd	Ag	Cd^{n*}	In^n	Sn*	Sb*	Te
	Ra	Rare earth	Hf^x	Ta^x	W^{nx*}	Re	Os	Ir	Pt	Au	Hg^n	Tl*	Pb*	Bi*	Po
		Ac	Th^{nx}	Pa	U^{nx}										

Rare earth La^n, Ce^n, Pr^n, Nd^n, etc.

n = nitride formation reported

x = nitrogen solubility reported

* = no nitrogen solubility found

investigate the literature to determine which metals do form nitrides and to find whether these do dissolve nitrogen under other conditions. Table 1 shows the elements of interest and their positions in the periodic table. WELLS[22] reports that most elements combine with nitrogen, and the nitrides include some of the most stable compounds known. He reports that the ionic nitrides include those of lithium, beryllium, magnesium, the metals of IIA and IIB, germanium and thorium, but not those of the IVA metals of the type TiN. A nitride of mercury, Hg_3N_2, also occurs but may be different in structure from the other Group II nitrides. The nitrides of beryllium, magnesium and calcium in Group IIA are described as being anti-isomorphous with oxides M_2O_3, i.e. the metal atoms occupy the positions of the oxygen atoms and the nitrogen those of the metal atoms in the M_2O_3 structure. Thus the structure becomes a compound of the type M_3N_2.

Covalent nitrides include the nitrides of Group III. These have great chemical stability and are described by WELLS[22] as being covalent nitrides. These compounds are formed of equal numbers of atoms of elements of Group III and V, and the nitrides of aluminium, gallium and indium crystallize with the wurtzite structure.

WELLS[22] reports numerous nitrides of the transition elements of varied structures. The characteristic feature of these nitrides is that the small non-metal atoms occupy some or all of the interstices of the metal lattice. The relative positions of the metal and nitrogen atoms is controlled by the number and nature of the interstices occupied by the nitrogen atoms, and they may be the same as in certain simple structures such as the sodium chloride and calcium fluoride structures. The following compounds have been reported to crystallize with the rock-salt structure: ScN, CeN, LaN, TiN, ZrN, VN, NbN, PrN.

We have now determined which metals will form nitrides at higher temperatures; these same metals may dissolve nitrogen at lower temperatures. Quantitative measurements are not important here but a review of the literature, indicated in Table 2, will show that there is a definite solubility of nitrogen in the metals listed.

TABLE 2

Metal	*References*	*Metal*	*References*
Iron	3, 10, 23, 24, 25, 26, 27	Hafnium	27
Zirconium	3, 10, 27, 28, 29	Thorium	27
Tantalum	3, 27	Vanadium	27, 33
Molybdenum	3, 27, 30, 31, 32, 10, 25	Niobium	27
Tungsten	3, 10, 27, 32	Chromium	27, 33
Manganese	3, 27, 33	Uranium	27
Titanium	27, 29, 33		

HÄGG[34] and SMITHELLS and ROOKSBY[35] have found that, although tungsten reacts very slowly to form the nitride W_2N at high temperatures, no solid solutions of nitrogen in tungsten are formed.

It has been shown[36] that in the ranges investigated in the experimental work, there is no solubility of nitrogen in cadmium, thallium, zinc, lead, bismuth, tin, and antimony. More recently it was found[33] that there is no solubility of nitrogen below 1200°C in nickel, cobalt, copper, and tungsten. In the references above, which indicated that tungsten would dissolve nitrogen, the measurements were made above 1200°C, the solubility being 0·304 $cm^3/100$ g at 2400°C, so that it is rather small if it exists at all.

RÖNTGEN and BRAUN[37] found that no measurable amount of nitrogen is absorbed by aluminium between 700°C and 1000°C because a surface film of aluminium nitride is formed which prevents diffusion of the gas into the metal. At most there may be 1 c.c./100 g in the molten metal at the melting point, but the solubility in the solid metal would be much smaller than that. On the other hand, CHAUDRON[38] and MOREAU[39, 40] found relatively large amounts of nitrogen (in one case 7·4 c.c./100 g in the solid metal) dissolved in aluminium which had been electrolytically polished so that no adsorbed layer would affect the results. Their method of degassing was to make the specimen the cathode in a vacuum discharge tube and apply a potential to the terminals of from 5 to 150 kV with a current of about 5 mA. Degassing took place in the cold, the surface molecules (of aluminium nitride presumably) being ionized and drawn into the electric field, replacement being effected by diffusion in the metal. In hydrogen-metal systems the experimental results presented by these workers all appear to indicate a much higher solubility than one would expect. It is felt that their findings in the aluminium-nitrogen system may possibly result from the decomposition of the aluminium nitride compound and do not, in fact, represent the true solid solubility of nitrogen in aluminium.

As was stated earlier, it seems that SMITHELLS' statement that nitrogen is only soluble in those metals which are capable of forming nitrides, is the most applicable generality that can be made about the nitrogen-metal systems.

Oxygen

The measurement of the solubility of oxygen in metals is quite difficult, since, except in the case of the noble metals, a separate oxide phase is formed when the limit of the solid solubility is exceeded. Thus the true solubility of the oxygen may be small in comparison with the total amount of oxygen present in the metal. In many molten metals the large content of dissolved oxygen is sharply decreased on solidification

when the gas may be evolved or may form an oxide of the metal. Generally, the solubility of oxygen in metals is of the same order of magnitude as that of hydrogen, but it is usually more difficult to determine. Unless the metal being used in an investigation is very pure, then the usual method of establishing equilibrium by diffusion inwards from an external oxide phase is liable to lead to erroneous results. As will be shown later, the gas used for preferential internal oxidation of impurity elements to form more stable oxides than the metal being examined, may be interpreted as gas going into solution. Thus a large error may be introduced where the true solubility is, in fact, quite small.

The oxygen-copper system seems to have been investigated more thoroughly than any others. HANSON, MARRYAT and FORD[41] placed the solubility limit at less than 0·009 per cent oxygen at 1000°C and ALLEN and STREET[42] suggested it to be not over 0·005 per cent oxygen at 500°C. The earliest data on spectroscopically pure copper were given by RHINES and MATHEWSON,[43] who reported that the solid solubility increases from 0·007 per cent by weight at 600°C to about 0·015 per cent at 1050°C. A later paper by PHILLIPS and SKINNER[44] showed that previous determinations had given results which were too high. They found the solubility to range from 0·0022 per cent at 700°C to 0·0077 per cent at 1050°C. The probable reason for the apparent difference of solubility is the experimental difficulties in the analytical determination of such small percentages of dissolved gas.

Usually the solubility of oxygen in metals is complicated by the formation of oxide phases, so that complex phase diagrams result. A diagram given by BARRER[3] illustrates the great variety of solid solutions possible. Earlier HERTY and GAINES[45] showed that oxygen must dissolve in liquid iron as FeO, and probably dissolves in the same manner in the solid metal. The maximum solubility in alpha-iron is probably less than 0·05 per cent and may be slightly higher in gamma-iron.

In the cobalt-oxygen system, SEYBOLT and MATHEWSON[46] found from analysis and micrographic observation that the solid solubility of oxygen in cobalt increases from about 0·004 per cent by weight at 550°C to 0·020 per cent at 875°C, at which temperature the hexagonal form of cobalt changes into the face-centred cubic form. In the gas-free metal this transformation occurs at 850°C. In the cubic cobalt, the oxygen content increases almost linearly at constant pressure from 0·004 per cent at 875°C to 0·022 per cent at 1450°C.

A similar solubility of oxygen in the form of nickel oxide (NiO) has been reported by MERICA and WALTENBERG.[47]

The solubility of oxygen in silver has been studied by STEACIE and JOHNSON,[20] who showed that it is proportional to the square root of

the pressure, is not in any simple relation to the temperature, and is a minimum at 400°C. They reported that, at a pressure of 80 cm of mercury, the solubility varies from 0·142 c.c. of oxygen *per c.c.* of silver at 200°C to 0·087 c.c. at 400°C, 0·193 c.c. at 700°C, and 0·354 c.c. at 800°C. They suggest that below 400°C, the dissolved oxygen exists in the form of Ag_2O, and above this temperature as oxygen atoms, since Ag_2O is unstable at the higher temperature. SIMONS[48] argues that the minimum is due to the rate of increase of the dissociation pressure of Ag_2O with temperature being greater than the rate of increase in solubility of Ag_2O in silver.

SMITHELLS[27] indicates that oxygen is probably slightly soluble in palladium and rhodium, but there are no published data. Oxygen is insoluble in mercury, gold, osmium, iridium, and platinum. If there is any solid solubility of oxygen in tin and lead it would be very small. GROSHEIM-KRISKO, *et al*[49] report that the oxygen content of molten lead at 350°C to be $48 - 60 \times 10^{-5}$ weight per cent. Solid solutions containing up to 40 atomic per cent of oxygen without the separation of an oxide phase are possible in zirconium[28, 50, 51] and titanium.[52] Oxygen is also soluble in vanadium, niobium and tantalum.

Hydrogen

In the metal-hydrogen systems there are four different types of interaction between the gas and the metal, as follows:

1. Hydrogen will react to form hydrides of a salt-like nature with the alkali and alkaline-earth metals. These primary salt-like hydrides are ionic, and generally crystalline, non-volatile compounds. These compounds, with the negative hydrogen or hydride ion H^- are analogous to the corresponding halides. GIBB[53] and RUNDLE[54] have fairly recent papers dealing in some detail with the formation of these hydrides, and their references will lead to a broader view than need be given here.
2. With the metals of Group IVB, VB and VIB covalent hydrides are formed yielding compounds, gaseous at room temperature, analogous to the hydrocarbons and typified by the familiar arsine. Some of the covalent hydrides formed are AsH_3 and SiH_4.
3. A number of metals, such as chromium, iron, cobalt, nickel, platinum, copper, silver, molybdenum, magnesium, and aluminium, dissolve hydrogen to form true solutions as indicated by the fact that (*a*) the hydrogen solubility varies as $\sqrt{P}$, and (*b*) solubility increases with increasing temperature.
4. With the rare-earth metals and titanium, zirconium, thorium, hafnium, vanadium, niobium, tantalum, uranium, and

manganese, hydrogen forms what DUSHMAN[10] calls "pseudo-hydrides." SMITH[2] describes these metals as being "exothermic occluders," while those in section (3) above are "endothermic occluders." Although the solubility varies as $\sqrt{P}$ for certain ranges of values of pressure (P), it decreases inversely as the temperature. Palladium behaves similarly with respect to hydrogen.

As far as the writer has been able to determine, hydrogen is not absorbed by tungsten (may have small solubility), gold, zinc (?), cadmium (?), indium, thallium, or mercury. Uranium has been reported to have a small solubility of hydrogen in one reference[27] and to be an "exothermic occluder" in another reference.[2]

Solid tin and lead have a very small capacity to absorb hydrogen.

GAS-METAL EQUILIBRIUM, DIFFUSION AND SOLUTION

It has been stated in the introduction to this paper that the effects of gases on the physical and mechanical properties of metals will be the main subject of study. However, the chemical considerations of the overall problem must not be entirely overlooked. In the preceding section, the various possible gas-metal combinations were reviewed in an attempt to show where one might expect to find unusual physical effects. In comparatively recent years there has been an increasing awareness of the profound effects which gases may have on solid metals. Much more study is needed, however, before a reasonable theory can be developed to cover the entire gas-metal field.

The formidable volume of work which has been devoted to the study of gas-metal equilibria is indeed impressive. It is felt that it is better not to consider this field at all, rather than to create a false impression by a too brief review; therefore, references will be made to a few fairly recent reviews devoted largely to this phase of the work. The reader is warned, not unjustly it is felt, to be extremely critical of solubility measurements and diffusion rates particularly for oxygen. As an example of this, EVANS[55] has listed the reported solubilities of oxygen in iron as found by fourteen different groups. The results vary from 6 to 0·006 per cent oxygen (at 1000°C), a marked variation. EVANS considers the serious discrepancy to be due to: (*a*) effects of traces of impurities which are often not reported in the literature, (*b*) the use of powder samples which adsorb large amounts of gas, which is reported as dissolved gas, (*c*) contamination of the iron by refractories or volatile compounds, (*d*) insufficient time allowed for equilibrium conditions to be reached, and (*e*) variation in the methods used, and consequent variation in the inherent errors. Similar types of errors are probably the cause of variations in the reported solubilities in other systems.

Most papers on gases and metals refer to SMITHELLS' monograph[1] (and occasionally to a paper[7]) which is now, justifiably, a classic in this field. In these treatments, SMITHELLS deals with the three processes involved between gases and metals: adsorption, diffusion and solution, and he shows how the laws governing these phenomena have been evolved.

In the few years after the publication of SMITHELLS' book, the subject of gases and metals advanced rapidly, so that by 1941 there was a need for a comprehensive review of the subject. BARRER[3] fulfilled at least part of the need by devoting a long section of his valuable monograph to the subject of diffusion of gases through solid metals. In this survey, BARRER dealt with the work which had been done up to that time, and developed a reasonable explanation for the phenomenon of gaseous diffusion through solid metals. The present author feels that this monograph is still to be highly recommended, although it is in need of revision.

Shortly before BARRER'S book was published, SIMONS and HAM[56] recorded their belief that the solution and diffusion of gases through metals could be considered as purely chemical phenomena. It was claimed that the specific nature of the solubility of gases in metals indicates that the gases are present mainly as compounds with the metal system. Since the rate of diffusion of a gas through a metal is inversely proportional to the thickness of the metal, it was felt that the adsorbed film of gas, usually postulated in diffusion processes, could not be related to the empirical results. For this and other reasons, diffusion was, therefore, treated as a case of chemical equilibrium between the metal and the gas, and on this basis the usual diffusion laws were derived without reference to the mechanism of the reaction at the gas-metal interface.

STEARN and EYRING[57] briefly considered the diffusion of gases in metals from the point of view of a theory of the absolute rates of reactions in which the slow process of diffusion entails the surmounting of an energy barrier. In this paper some doubt was thrown on the possibility of "rift" diffusion which had been suggested as a mechanism by SMITH and his co-workers.[2]

FAST[58] has discussed the mechanism of diffusion of gases in metals, a subject about which, he says, ". . . very little is yet known with certainty." In general the gas, in atomic (or ionic) form, penetrates between the metal ions of the lattice. Unlike SIMONS and HAM,[56] FAST thinks that the surface effect and not the diffusion itself is frequently the deciding factor in the diffusion velocity. In a later paper FAST[59] considered further the permeability of gases through metal walls. As a result of the diatomic character of the gas phase, there are five stages involved in the penetration of gases through metal walls:

(1) dissociation of molecules into atoms or ions at the surface of entry, (2) penetration of the atoms or ions into the metal, (3) diffusion in the metal, (4) a transition from the dissolved to the adsorbed state at the exit surface, and (5) reassociation of the components into diatomic molecules. The speed of the whole permeation process is determined by the slowest of these five successive steps. Examples were given of experimental cases where each of these five stages is the controlling factor and these were discussed in some detail. The influence of the surface condition on permeability was explained, and data were given to show the effect of temperature on permeability.

RHINES[60] has written a general review of a number of industrial processes which are influenced by gas-metal diffusion which may have either beneficial or detrimental effects. The information for that paper was gathered by consulting people who were closely associated with the various practices in which gas-metal diffusion plays a part. Some of the phenomena discussed were: hydrogen embrittlement, subsurface oxidation, and rejection of dissolved gases on solidification of molten metal.

In 1948, BARRER[61] reviewed the factors determining the formation of interstitial compounds and some of their properties. While the formation of compounds is not really within the scope of the present paper, this discussion by BARRER is extremely interesting and valuable. It is noteworthy that BARRER, unlike many authors who have written about gas-metal systems, regards the interstitial hydrogen dissolved in a metal to be substantially unionized. The paper deals with the exothermal formation of solutions of hydrogen in metals, equilibrium between interstitial and free gas, and rate-controlling processes.

In the same year, three papers appeared which apply to this discussion. LEPP[62] studied the absorption of gases in metal by applying physico-chemical laws and thermodynamic analysis. SACHSE[63] gave detailed abstracts of studies (conducted in Germany during the late war) of the diffusion of hydrogen and deuterium in palladium, and of hydrogen in iron. SCHAFER[64] reviewed the work of SIEVERTS done during the war years on the solubility of hydrogen and deuterium in nickel, niobium, and palladium, and on the effect of hydrogen on the transition points in some iron alloys.

A monograph recently published by SMITH[2] deals entirely with the problem of hydrogen in metals. This monograph covers the subject with remarkable completeness, and the bibliography of some 1470 references is very useful. However, very early in his book, SMITH questions the possibility of the existence of solid solutions in "endothermic metal-hydrogen systems," and throughout the remainder of the book the discussion is dominated by his "rift" hypothesis. It is claimed that the various experimental phenomena can best be explained

by the opening or closing of these "rifts" which are apparently lattice defects, or widened lattice intervals. One is impressed by the fact that the rift hypothesis is entirely qualitative. The concept of rifts might also be questioned if the specificity of diffusion rates and solubility data are considered. The present writer must agree with RANSLEY,[65] who said of SMITH'S monograph: "Indeed, it is not unfair to say that in a good many instances the author has strained the facts very severely in an attempt to make them conform to this theory, when a simple explanation based on lattice solution and diffusion would appear to be perfectly adequate." There are, of course, cases where the rift hypothesis may very well apply, and where it is, in fact, the preferred explanation. However, the explanation of such a formidable mass of empirical results by one mechanism does not yet seem possible. Since the publication of this book, SMITH has written three papers incorporating the rift hypothesis in each.[66, 67, 68]

As part of a book on vacuum techniques, DUSHMAN[10] included a section (Chapter 9) on gases and metals. Although much of the experimental data in this section was published before 1940, the review is valuable because it does bring much of the available information under one cover. DUSHMAN, among other things, considered the diffusion and solubility of hydrogen, oxygen and nitrogen.

In 1950, BASTIEN[69] discussed the various stages in the permeation of a gas through a metal wall. The consideration, which deals particularly with the iron-hydrogen system, includes adsorption and diffusion phenomena, the effects of temperature and pressure, and effects of gas-metal equilibria.

Recent work by BHAT and LLOYD[70] has shown that the effect of cold-work is to decrease the rate of hydrogen diffusion in steel at room temperature. It was pointed out that diffusion results obtained by electrolytic experiments are only qualitative; this might account for many of the disagreements between solubility and diffusion measurements in other hydrogen-metal systems. For reasons which will not be discussed here, the authors believed that, although variation in the lattice spacing could affect diffusion rates, probably the occluded molecular hydrogen stored at crystallite boundaries is the major factor affecting the diffusion process. The correlation of cold-work and rate of hydrogen diffusion with changes in hardness and density are also discussed. This work is related to the studies conducted by ANDREW on hair-line cracks and embrittlement in steel (q.v.).

ISENBERG[71] has written a brief note to demonstrate why hydrogen ionizes when dissolved in a metal. On a semi-classical model the conclusion was reached that hydrogen should ionize in all metals.

A series of studies by GULBRANSEN and ANDREW[72–75] on the reactions of zirconium, titanium, niobium and tantalum with oxygen, nitrogen

and hydrogen provides an excellent review of the work which has been done on these metals. The data found by GULBRANSEN and ANDREW on the kinetics of the gas-metal reactions add considerably to the knowledge of these metals. At present it is difficult to follow the latest work on titanium and zirconium, for so much experimental work is in progress. The series noted might provide a useful starting point for a reader particularly interested in this work.

A recent review by RHINES,[9] although very short, summarizes some of the experimental facts of gas-metal diffusion processes. This seems to be the most recent review of this field.

Although not directly applicable here, it is interesting that equilibrium relations in certain (titanium) alloy systems have been investigated by measuring hydrogen pressures in equilibria with very small amounts of hydrogen in solution in the alloys. MCQUILLAN[76–79] has studied the alpha → beta transformation in various titanium samples, and the effect on the transformation of alloying elements.

EFFECTS OF DISSOLVED GASES ON THE PHYSICAL AND MECHANICAL PROPERTIES OF METALS

In addition to such effects as hydrogen embrittlement of various metals, and porosity caused by gases coming out of solution, there are several phenomena of interest which are directly attributable to dissolved gases.

*Internal Friction**

One of the most interesting of these phenomena is the effect which absorbed gases have on the relaxation behaviour of metals. A comprehensive theoretical treatment of relaxation phenomena, to which the reader is referred, has been given by ZENER.[80] One type of relaxation effect is internal friction. It has been found that, under suitable conditions, there will be a maximum absorption of energy by a metal when an applied vibrational stress has a frequency of the same period as that of some mechanism taking place within the metal itself. Such mechanisms as reversible grain boundary movement, stress-induced diffusion, and thermal diffusion have been postulated as causes for the marked dissipation of energy which has been observed in some systems.

The relationship between magnetic and mechanical relaxation phenomena was first mentioned by SNOEK.[81] It had been pointed out by YENSEN[82] that small quantities of both carbon and nitrogen have a marked influence on both the magnetic and plastic properties of iron. It was thought that the slow decrease with time of the magnetic

* This section on Internal Friction is presented only as a brief description of this phenomenon which is one manifestation of the effect of gases on metals. A comprehensive review of Internal Friction in Metals by A. S. NOWICK is presented elsewhere in this volume.

permeability in iron at room temperature was due to the segregation of nitrogen. Very little exact knowledge was available, however, until SNOEK[83] found that such phenomena as magnetic after-effect and time decrease of permeability depend on the presence of carbon or nitrogen in solid solution. If these solute atoms are precipitated from the lattice the phenomenon of magnetic after-effect is no longer present. RICHTER[84] then found that samples of iron which had displayed a magnetic after-effect also showed an elastic after-effect. Both effects were temperature-dependent, and the absolute values of the relaxation times were nearly equal. Finally, SNOEK[85] was able to show that both carbon and nitrogen are capable of producing an elastic after-effect (relaxation peak) in iron, thus correlating the magnetic and mechanical phenomena. SNOEK mentioned that there was a true nitrogen solubility at room temperature and that this could be measured by means of the elastic after-effect.

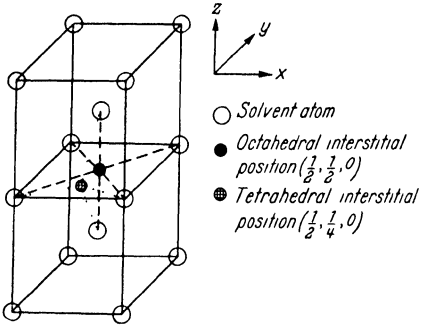


Fig. 1. Examples illustrating the octahedral and tetrahedral positions in a body-centred cubic lattice. The closest atoms to these positions are indicated by dashed and dotted lines (KÊ[86]).

By 1941 SNOEK was able to account for these effects on the basis of the interstitial positions in a body-centred cubic lattice which have tetragonal symmetry, the tetragonal axis being parallel to any one of the three principal axes. In an unstrained material, the equilibrium distribution of the solute atoms among the three types of interstitial positions (given by the crystallographic indices $(\frac{1}{2}00)$, $(0\frac{1}{2}0)$, and $(00\frac{1}{2})$) is quite random. However, if the material is elastically stressed in the x direction, there will be a preferred distribution of solute atoms in the x-type positions $(\frac{1}{2}00)$.

The above explanations apply to the case of annealed material. The results of SNOEK's work with cold-worked iron were more complex and his explanation only tentative, and admittedly unsatisfactory.

This relaxation peak discovered by SNOEK appears to be characteristic of all body-centred cubic metals. More recently, KÊ has studied three peaks in tantalum, corresponding to carbon, nitrogen and oxygen in interstitial solid solution.

KÊ[86] has elaborated on the theory developed by SNOEK, and has done experiments, using tantalum, which have extended the empirical knowledge of the phenomenon considerably. His observations are the first experimental evidence that oxygen forms an interstitial solid solution with a metal. KÊ has emphasized the important observation

that the internal friction peak for oxygen cannot be interpreted in terms of a single relaxation time with a unique activation energy. The significance of this will be discussed later.

Kê has proposed a simple picture of this relaxation effect which may be described with reference to Fig. 1, which represents a body-centred cubic lattice. If solute atoms occupy the octahedral interstitial positions in the lattice, then the lattice will be subjected to tetragonal deformation with the tetragonal axis along one of the ⟨100⟩

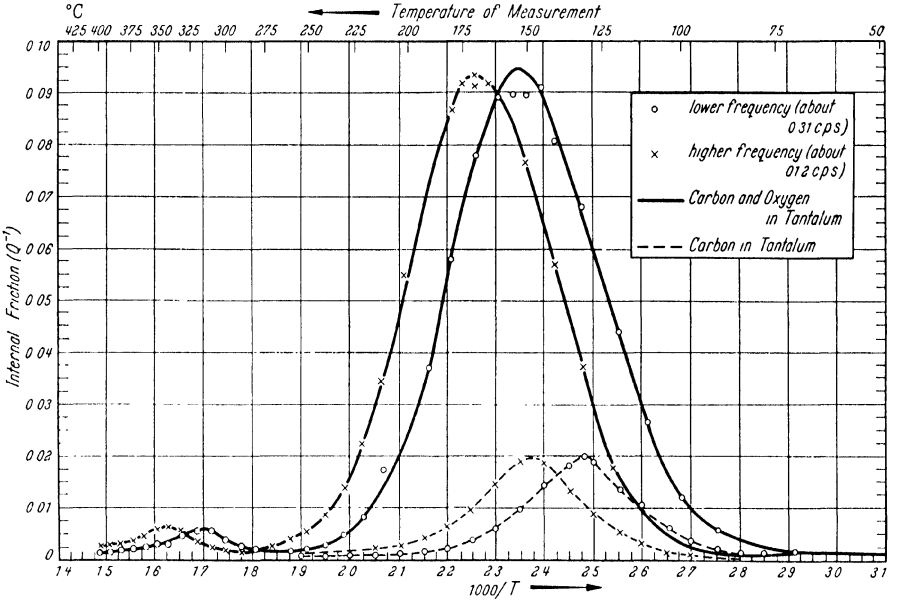


Fig. 2. Internal friction curves for original tantalum specimen and for the oxygen-enriched specimen. Both were determined with two frequencies of vibration (Kê[86]).

directions, depending on which position the solute atom has occupied. If a tensile stress is applied along one of the ⟨100⟩ directions, then the solute atom will prefer to occupy a position such that the tetragonal axis mentioned above will be along the same direction. Thus, under fluctuating stress, solute atoms will "diffuse" in an attempt to maintain this state. Under suitable conditions of temperature and applied frequency this "diffusion" process will absorb large amounts of the applied vibrational energy.

Since all interstitial positions are equivalent in the ⟨111⟩ direction then a tensile stress applied in that direction should produce no relaxation effect. Dijkstra[87] has confirmed this by showing that the relaxation strength along a ⟨111⟩ direction in single crystals of alpha-iron is only about 5 per cent of that along a ⟨100⟩ direction.

In his experimental work, Kê[86] has determined that the oxygen content in solution in tantalum is greater than 0·06 weight per cent at 170°C.

In Fig. 2 are shown the internal friction peaks determined at two frequencies of vibration caused by carbon (0·013 per cent) in tantalum. The higher peaks in the same figure are caused by the composite effect of carbon and oxygen which diffused into the metal on annealing in an oxygen atmosphere. The contribution due to oxygen alone was determined by subtracting the carbon curve value of internal friction from the composite value at the same temperature. The internal friction peak due to oxygen alone was found to occur at about 20 C° higher than that due to carbon.

In Fig. 1 it is shown that there may be two kinds of interstitial positions for the solute atoms, and these are shown with their crystallographic indices. Each kind of position shown may take up one of three types of position, each having the tetragonal axis along one of the principal axes, as mentioned before. Despite the different kinds of position which the solute atom may occupy, the mechanism of preferential distribution will be the same in each case. Kê has emphasized that the distortion produced when the solute atom is in the octahedral position is much higher, and the internal friction peak, therefore, greater.

Since carbon atoms occupy the octahedral positions in martensite,[88] it is assumed by Kê that it is easier for the carbon atom to displace the two nearest neighbours of the octahedral position than it is to move the four atoms equidistant from the tetrahedral position. Regardless of the fact that the octahedral position is smaller than the tetrahedral with respect to nearest neighbours, this view is held to be correct. Since tantalum has an atomic radius greater than iron, the available interstitial space in the tantalum lattice will be larger, and the carbon atoms can enter the octahedral position. Kê was able to demonstrate that the carbon peak at 150°C has a single relaxation time and a unique heat of activation. This indicates that the carbon is distributed in only one kind of interstitial position. If this is the tetrahedral position (for the 150°C peak) then there should be another, larger, peak associated with the octahedral position. This was not found, so it is assumed that the octahedral position was the preferred location for the carbon atom.

The characteristic internal friction peak for the metal containing oxygen in solution showed that there is no single relaxation time, and no unique activation energy. It is assumed from this that the oxygen, with an atomic radius smaller than carbon, may go into either interstitial position. Kê[86] has indicated how the diffusion mechanism from one position to another will determine whether or not a single relaxation time is associated with the process.

Kê[89, 90] also measured the relaxation strengths of tantalum specimens containing small amounts of carbon and oxygen. Experiments at temperatures ranging from 350° to 700°K indicated that a linear relationship exists between the reciprocal of the relaxation strength and the absolute temperature. Since this plot is a straight line passing through the origin, it is concluded that the interaction between solute atoms is negligibly small, and that the critical temperature for self-induced preferential distribution of solute atoms is close to absolute zero. This linear relationship is held to be in agreement with the previously postulated theory of stress-induced preferential distribution of solute atoms. No spontaneous ordering can occur in the solutions of carbon and oxygen in tantalum.

It had been reported in an earlier paper[86] that an internal friction peak was found to occur in tantalum (with a vibration frequency of about 1 c.p.s.) at 350°C, as well as those at 150°C and 170°C, the latter two being associated with carbon and oxygen respectively. Kê[91] later showed that the 350°C peak is associated with the presence of nitrogen in tantalum. It is seen that the temperature of the peak is abnormally high, and experiments were undertaken to decide whether the phenomenon was due to stress-induced diffusion as for carbon and oxygen, or to viscous behaviour of the boundaries between tantalum and tantalum nitride. It was found that tantalum, annealed in nitrogen at 1200°C, and quenched from 900°C, had an initially high internal friction at 350°C but this decreased to a much lower value after 20 hours. Similarly annealed material quenched from 600°C did not have as high an initial value of internal friction and this also decreased to the same equilibrium value after 20 hours' annealing at 350°C. During this time, tantalum nitride needles were precipitated from the supersaturated solid solution of nitrogen in tantalum. From this it was concluded that the internal friction observed could not be due to nitrogen in precipitated form, since then the internal friction should have increased while precipitation proceeded.

This precipitation of nitrogen from the solid solution with tantalum, as indicated by the internal friction measurements, is similar to that observed by Dijkstra[87] in the cases of carbon and nitrogen precipitating from the solid solutions with alpha-iron.

Kê found the heat of activation associated with the 350°C absorption peak to be about 44,000 cal/mole, which is much larger than the values for oxygen in tantalum (29,000 cal/mole) or for nitrogen in alpha-iron,[92] which is 20,000 cal/mole. The observed nitrogen-tantalum peak could not be interpreted in terms of a single relaxation time with a unique heat of activation. Thus the same mechanism of solution is held to be true for nitrogen as was postulated for oxygen in tantalum, i.e. the nitrogen may be in either the octahedral or tetrahedral positions.

The extremely high heat of activation, and with it the relatively high temperature at which the peak occurs, was not explained. It was suggested that these phenomena may be due to a larger distortion in the lattice.

Kê[92] has also done work on iron and found three internal friction peaks to occur at 20°, 225° and 490°C. The 20°C peak is believed to be of the same origin as that observed by Snoek.[81] The mechanism causing this peak, according to Kê, Snoek and also Dijkstra,[87] is the anelastic effect caused by nitrogen in solution which has been described for the case of carbon in tantalum. The 225°C peak which had been observed previously by Snoek[81] and West[93] was found by Kê to be caused by the stress-induced diffusion of nitrogen atoms, as in other cases, but "within some peculiar type of stress regions created in the specimen after it was subjected to cold work." No account of this mechanism has been given but the heat of activation for this diffusion is known to be 32,000 cal/mole and the diffusion distance is in the order of a few atomic diameters. The 490°C peak was found by Kê to be due to grain boundary relaxation in alpha-iron.

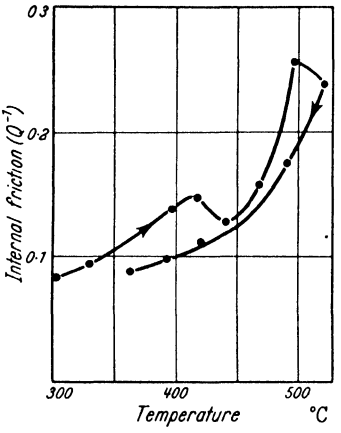


Fig. 3. Effect of heating in air on the grain-boundary relaxation in silver. (The steadily rising internal friction which remains after the peak at 400°C has disappeared, results from another cause) (Nowick[94]).

In a recent review on stress relaxation across interfaces, Nowick[94] has indicated that small amounts of impurities, microscopically undetectable, cause blocking of the interface relaxation when lodged in the grain boundaries. These impurities may cause partial or complete disappearance of the grain boundary internal friction peak. This type of peak is not due to the same causes as the one which has been discussed in this section thus far. It is suggested that the reader see Nowick's paper[94] and the references contained therein for a complete picture of this phenomenon. Oxygen, probably in the form of oxides, is quite effective in decreasing the grain boundary peak in copper and silver. Fig. 3 shows how heating polycrystalline silver in air can reduce the grain boundary peak almost to zero. A similar effect has been observed for copper. The picture of this effect which Nowick[94] has reiterated is that the foreign particles "create new corners at which slip cannot take place," so that the total relaxation may be smaller than in the pure substance. According to his hypothesis, ". . . it is to be expected that, as impurities are added systematically, the grain boundary peak

should not only decrease in magnitude but also move toward lower temperatures because the effective size of the interfaces is shortened and, therefore, becomes smaller."

Some of the more recent work on internal friction phenomena due to gases in solution has used the internal friction measurements as a quantitative indication of the amount of carbon or nitrogen in solution in alpha-iron. That the magnitude of the internal friction is strictly proportional to the concentration of either the carbon or the nitrogen in solid solution, for small solute concentrations, was predicted and verified by DIJKSTRA[87] for the case of alpha-iron. In a later paper DIJKSTRA[95] used this relationship to follow the precipitation of either nitrogen or carbon from solid solution, and also to measure their solubility in alpha-iron as a function of temperature. It is pointed out that the internal friction at the maximum of the internal friction peak (Q_{max}^{-1}) has been related to resistivity which has, in turn, been related to the solute element concentration. This has shown that Q_{max}^{-1} is nearly equal numerically to the weight per cent concentration of either carbon or nitrogen in solid solution. The present writer feels that it is not necessary to review DIJKSTRA'S work which is very interesting, but slightly beyond the scope of this review. As well as the facts above, DIJKSTRA stated also that hydrogen had no effect on the internal friction of alpha-iron in the temperature region between — 40° and 100°C.

Despite the fact that FAST[96] had reported that the nitrogen peak in iron becomes broader and appears at a higher temperature (nearer the carbon peak) due to the presence of manganese in iron, JOSEFSSON and KULA[97] have applied the internal friction peak effect to the study of ageing and related phenomena in commercial mild steel. This work has confirmed the fact that the presence of impurities, such as manganese (and perhaps phosphorus), does change the position of the nitrogen peak and broadens the internal friction curve. Also it was confirmed that the carbon peak was not changed in position, although it may have been broadened by the addition of manganese to the iron. Ageing curves (as a function of internal friction) have been given for four steels which were each treated in several ways. The removal of nitrogen from solution by means of an aluminium addition, and the tendency for nitrogen to remain in supersaturated solid solution, even at slow cooling rates, in certain steels was also shown.

Finally, the most recent work which has been published is a brief note by MARINGER and MUEHLENKAMP,[98] who have found an internal friction peak at 1300°F, and 1 c.p.s., in molybdenum, caused by the stress induced diffusion of nitrogen atoms in solid solution.

Other effects, similar to those which have just been described, have been reported in the literature. KÖSTER[99] studied an unexpected decrease in the modulus of elasticity and increase in damping capacity

with the introduction into several metals of solute atoms of small solubility. It was found that the anomalous behaviour was due to small amounts of oxygen held in the copper and silver being studied. KÖSTER decided that, from knowledge of the nature of damping, the effect can be ascribed to stress-induced diffusion of the solute atoms. Later, ZENER[100] was able to show that the presence of residual elastic strain energy is associated with a reduction of the elastic moduli, in agreement with KÖSTER'S observations. It is proposed that, as a solute atom passes over the potential energy barrier between two stable interstitial positions, the lattice acquires additional strain energy which is associated with a lowering of the elastic moduli. Two brief summaries of papers by MARX[101] and KOEHLER[102] add to the information available on the effect of hydrogen in copper. In each case the hydrogen has a marked effect on the elastic properties of the metal.

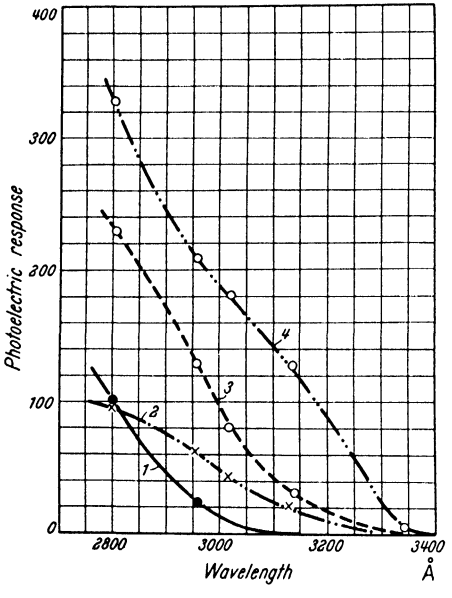


Fig. 4. Photoelectric activity of zirconium (RENTSCHLER and HENRY[104]).

Work-function

Some studies have been carried out on the effect of gases on the photoelectric work-function of some metals. When such metals as thorium, uranium, calcium, caesium, and barium react with oxygen, the photoelectric threshold is shifted toward a longer wavelength. RENTSCHLER and HENRY[103] first studied the metals just mentioned and later[104] studied zirconium, titanium, hafnium, and thorium. In the latter paper it was pointed out that when oxygen, for instance, is present as an oxide at the grain boundary of a metal, a sputtered film from this metal forms a surface from which the true photoelectric threshold characteristic of the pure metal may be measured. It is also to be expected that, if the gas is in solution in the metal, surfaces sputtered from different batches of the metal will yield different values of the photoelectric sensitivity, due to variations in the gas content of the different batches. RENTSCHLER and HENRY predicted that this would be more a difference in relative sensitivity to different wavelengths than in a given threshold wavelength. It was, in fact, found that, unlike the other metals mentioned above, different photo-cells of zirconium and titanium responded

differently to various wavelengths even though the photoelectric threshold was the same.

RENTSCHLER and HENRY[104] were able to show the effects of various treatments with gas on the photoelectric response of zirconium. When a small amount of oxygen reacted with the active metal surface the threshold wavelength increased. Fig. 4 shows the photoelectric activity of zirconium in different conditions. Curve 1 is the response of pure

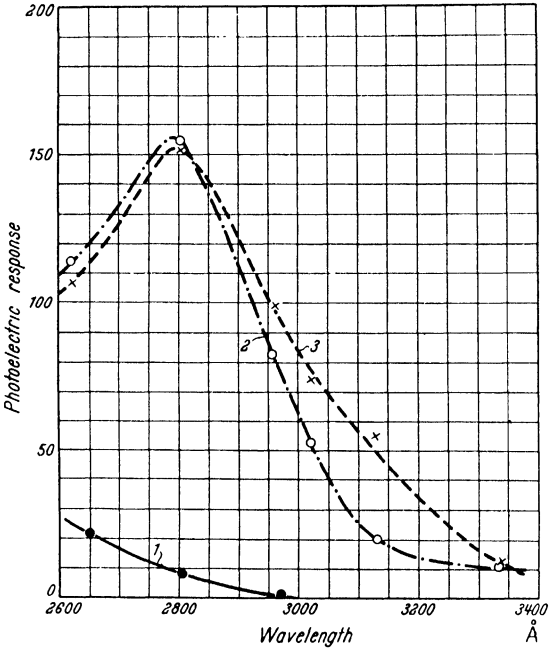


Fig. 5. Photoelectric activity of titanium (RENTSCHLER and HENRY[104]).

zirconium metal. Curve 2 is the response from a surface sputtered from zirconium previously heated to 400°C in hydrogen. Curves 3 and 4 show the response from surfaces sputtered from zirconium heated in nitrogen and oxygen respectively. It was found that the response from a surface sputtered from an oxygen-bearing wire does not depend on the time of sputtering. Thus the lower work function is independent of the film thickness provided the sputtering is done from metal with the gas uniformly dissolved throughout the metal. Also the threshold wavelength was found to be independent of the oxygen content (provided there was some in solution), but the shape of the response curve is dependent on the oxygen content. These facts were interpreted to mean that the film consists partly of pure metal with no gas, and partly of metal containing gas in solution, the former having a low threshold and the latter a higher threshold wavelength.

Fig. 5 shows the same type of thing for titanium where the effect of the dissolved gases is even more marked than for zirconium. Curve 1 represents the photoelectric response for pure titanium; Curve 2 is the response for a solution of nitrogen in titanium; and Curve 3 the response for oxygen-bearing titanium.

In an interesting, but only qualitative experiment, RENTSCHLER and HENRY found that titanium with oxygen in solution is not as volatile as the pure metal. This effect has also been found with zirconium.

It was found that thorium behaves like zirconium and titanium but only in so far as the gas increases the threshold wavelength. Both oxygen- and nitrogen-treated surfaces have a response similar to that of Curve 2, Fig. 4, i.e. a response lower, at low wavelength, than that of the pure metal. Since the thorium behaved like zirconium and titanium where the gas has reacted only with the surface of the metal, it was assumed that the gas does not dissolve as readily throughout the thorium. No explanation was given of the reason for this effect.

Recently, ANDERSON[105] has studied the work functions and ageing characteristics of copper. The work functions of fourteen surfaces were measured by means of contact differences of potential with respect to barium surfaces of a known work function. The surfaces were prepared from spectroscopically pure copper which had been subjected to 40 vacuum fusions and fractional distillations. The result of some 200 measurements showed that the first films to be deposited were not reproducible, and drifted rapidly with time to a lower work function. Later deposits were more reproducible, and the rate and magnitude of drift decreased. For the specimens 9–14 there was good reproducibility and no drift. At no time was there any drift to a higher work function. The variability is due to dissolved gases which are retained by the copper despite vacuum fusions, and are reabsorbed into the condensed film. Since the gas, presumed to be oxygen, lowers the work function, it is concluded that the gas is in solution and not merely adsorbed on the copper. This conclusion is reached since oxygen adsorbed on a clean tungsten surface raises the work function, and copper would be expected to behave similarly if it had an adsorbed gas film. However, it must be emphasized that there was no proof of the presence of oxygen in the copper. FAST[58] agrees that the adsorption of oxygen causes an increase in the work function.

A number of papers on the influence of hydrogen on photoelectric sensitivity have been listed by SMITH.[2] Except for those reported in this paper there appears to be no recent work on this subject.

Lattice Expansion

Recent work on the lattice expansion of metals due to the presence of dissolved gas in the lattice appears to be almost non-existent. A

relatively recent paper, which has some bearing on the subject, is by HUME-ROTHERY and ANDREWS.[106] In studying the lattice spacing and thermal expansion of copper, they accidentally oxidized some copper at 771°C. This metal was cooled to room temperature, reheated to 870°C for one hour and again cooled to room temperature. X-ray analysis of this showed oxide lines, but the copper lattice spacing was the same as that of pure copper. At higher temperatures the lattice spacing of the oxidized metal was only slightly higher than that of the pure metal, the difference at 770°C being 0·0004 Å. No attempt was made to examine equilibrium conditions.

A large part of the available information on metals which dissolve only small amounts of gas is that for electrodeposited chromium. HÜTTIG and BRODKORB[107] report that it has a body-centred cubic lattice which has a larger parameter than that of chromium which has been melted, and this was taken to indicate the entry of hydrogen into solid solution. WOOD[108] later found, along with a marked line broadening, a very slight displacement of the lines of a diffraction pattern. He concluded that a small part of the hydrogen in chromium is in solid solution, but that the major portion is adsorbed on the walls of "crystallites" which cause the diffuseness of the diffraction lines. HUME-ROTHERY and WYLLIE[109] confirmed the formation of a body-centred cubic structure, and recognized the possibility of "abnormal lattice spacings due to co-deposited hydrogen" but did not study the effect in detail.

Recent work has been done on electrolytic chromium by POTTER and LUKENS,[110] who measured the effect of hydrogen on the lattice parameter. As in other cases, the electrodeposited metal gave a normal body-centred cubic diffraction pattern with the lattice parameter somewhat enlarged compared to gas-free metal, and with a broadening of the lines. This was taken to indicate an enlargement of the lattice with variations in the lattice parameter of 0·3 to 0·4 per cent, causing the line broadening. On degassing at 500°C, the parameter decreased somewhat and the lines were less diffuse. The sample degassed at 1000°C showed a normal lattice with sharp lines of almost the correct spacing. Similar effects had been observed with manganese[111] on degassing at 300°C and 500°C, respectively.

In discussing this paper, SNAVELY briefly described a theory which he had developed from experimental evidence reported in another recent publication.[112] It was suggested that unstable chromium hydrides were deposited, with the hydrogen atom in interstitial positions forming either a wurtzite- or fluorite-type structure with the chromium, depending on the hydrogen content. Since these hydrides are unstable at ordinary temperatures they dissociate rapidly after deposition to form atomic hydrogen and body-centred cubic chromium. A large volume change (15 per cent) accompanying this reaction is believed to

be related to the hardness, brittleness, and cracking characteristics of this type of electrodeposit. "Microfissures" are formed during the volume change and the hydrogen, which has a very low solubility in the chromium lattice, is "occluded" in the fissures. These fissures are apparently the cause of the line broadening in the diffraction patterns. POTTER and LUKENS had suggested the possibility of a supersaturated interstitial solution of hydrogen in chromium. However, SNAVELY pointed out that the body-centred cubic chromium lattice has interstices which are about 1·58 Å in least diameter, while the hydrogen atom is less than 1 Å in diameter. SNAVELY[110, 112] considered that "an interstitial atom must be sufficiently large to contact all surrounding atoms if a stable or even metastable interstitial solid solution" is to form, and, therefore, the possibility of chromium absorbing hydrogen is small.

SMITH[2] reports work by two investigators who found lattice distensions in iron and nickel which had been electrodeposited from solutions containing arsenic. SMITH claims that the observed increase in parameter, accompanied also by line broadening, is due to the arsenic.

It must be appreciated that the gas content of an electrodeposited metal is not an equilibrium gas content, and may cause strains not present in an annealed metal. It seems that no lattice parameter increase has been found in those metals which dissolve only small amounts of hydrogen when in equilibrium with the gas phase. This is really to be expected if one considers the size of the hydrogen atom compared to the relatively large interstitial space available for the small number of hydrogen atoms absorbed. However, SMITH[2] uses this absence of true lattice expansion as an argument in favour of his "rift" hypothesis.

POTTER and HUBER[111] have studied the effect of hydrogen on the lattice parameter of electrolytic manganese. It was found that the manganese lattice is distended, the average lattice parameter enlargement being about 0·0003 per cent per c.c. of gas per 100 g of metal. A widening of the lines in the x-ray diffraction patterns of gas-bearing specimens indicated a deviation from the average parameter caused by non-uniform distortion of the lattice. This non-uniform lattice was attributed to non-uniform gas distribution. The effect of various degassing treatments on the metal indicated that the parameters were not affected unless the metal is heated under a reduced pressure at over 300°C. At 500°C and over, the lattice returned to its normal size and the formerly broadened diffraction lines became sharp. It is pointed out that although the hydrogen in solution is only 0·02 per cent by weight, under normal conditions it would occupy a volume twenty times that of the metal containing it. The authors state further: "It is not unreasonable to expect that this highly compressed gas would have some effect on the structure of the metal containing it." The concept of this

dissolved gas being merely "highly compressed gas" is hardly in keeping with other information. The non-uniformity in the lattice parameters mentioned above is not proportional to the amount of gas in the metal. The deviation from the mean lattice parameter (line broadening) is zero at zero gas content, high at 95 c.c. per 100 g of metal, and decreases with increasing gas content up to the observed gas content of 615 c.c. per 100 g of metal. It is shown that this indicates that the deviation might again be zero at a value of about 695 c.c. hydrogen per 100 g manganese, when the metal unit cell would contain exactly 2 gas atoms. Since there was some diffuseness in the lines at 350 c.c. of hydrogen per 100 g of manganese it was concluded that the average distribution of 1 atom of gas per unit cell was not uniform and that the cells could take two atoms each.

In another paper[113] it is inferred on the basis of related phenomena described in the literature, that, owing to the absorption of hydrogen by nickel, systematic lattice defects arise in the electrolytically deposited metal. These defects amount locally to about a 13 per cent change in the lattice constant. These local lattice distortions cause the fine-grained structure and hardness of the deposit. In another experimental study[114] it is shown that hydrogen participates directly in the formation of the structure of the electrodeposit of nickel. One other abstract refers to a paper[115] which describes an experimental spectrographic study of the state of hydrogen in a metal, the action of x-rays and the structure of the lines of the $K_{\alpha 1-\alpha 2}$ doublet. It is claimed that hydrogen in the form of protons is a factor in determining the structure of the doublet and also the properties of the material.

Considering those metals which dissolve larger amounts of hydrogen SMITH[2] in various parts of his monograph discusses the effect of hydrogen on the metal lattice. Most of the experimental work has been done on palladium, but information is given for other metals which absorb hydrogen readily. Representative of the studies on palladium is the work of OWEN and JONES[116, 117] which will be summarized here. By making measurements of the palladium lattice parameter under isothermal conditions and at varying pressures, the hysteresis effect when the pressure of the gas is taken through a cycle was determined. Fig. 6 shows the results which were obtained at 100°C. Those parts of the isothermals in which only a single component existed could be reproduced by simply altering the pressure. This reproducibility disappeared, however, when a second phase was present.

It was found that the hydrogen solubility changed rapidly when the beta phase appeared or disappeared. The measurements made are interpreted as meaning that the system consists of two solid solutions at the temperatures considered, namely 100° and 120°C. The hydrogen-rich beta phase is closely associated with the combination Pd_2H which

is capable of dissolving more hydrogen. The beta phase was always distorted no matter what precautions were taken to obtain it without distortion. It had been suggested earlier that this phase had an orthorhombic lattice, but this was not confirmed. It was found too that the distortion of the lattice remained after removal of the hydrogen. OWEN and JONES state that, "The presence of the hydrogen, even though it enters the palladium lattice interstitially, causes a permanent displacement of the palladium atoms which if the concentration is high enough, cannot be removed without subjecting the material to a

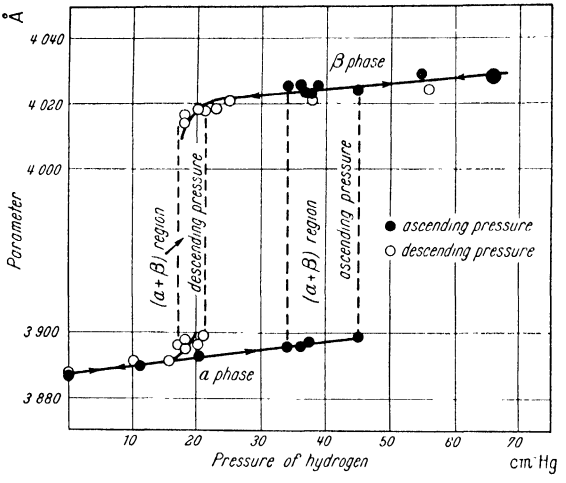


Fig. 6. Effect of hydrogen pressure on lattice parameter of palladium (OWEN and JONES[116]).

temperature well above that required to remove the hydrogen from the metal. This would suggest that at certain concentrations the gas may not be in its normal state when it leaves the metal."

Recently CLARK[118] has reported a new determination of the lattice constants of alpha (hexagonal close-packed) titanium. Comparing his results with earlier work, CLARK found agreement with FAST's work but not with others. CLARK is able to show how the C_0 value is markedly increased by oxygen and nitrogen which form interstitial solid solutions, while the a_0 value remains essentially unchanged. Since the only points of disagreement with earlier work were in the C_0 values, CLARK concluded that the other materials tested must have been contaminated with oxygen or nitrogen. Fig. 7 shows CLARK's results indicating the effect of the two gases on the lattice.

BUSK and BOBALEK[119] found a similar effect in the case of a magnesium + 10 per cent aluminium alloy charged with hydrogen in an alkaline solution. In this case both parameters had increased

significantly, and it was concluded that the hydrogen had dissolved interstitially.

The most recent discussion on the lattice parameter has been presented by SMITH.[68] X-ray evidence of so-called "rifts," or widened lattice spacings in work-hardened metal, as compared with the relatively continuous and uniform lattice of fully annealed metal is discussed in connexion with hydrogen solubility. SMITH claims that the characteristic line-broadening associated with the absorption of hydrogen is due to widened lattice spacings, caused initially by plastic deformation, but which may be extended by hydrogen absorption. The development, manifestations, and effects of the "rifts" are described in this paper.[68]

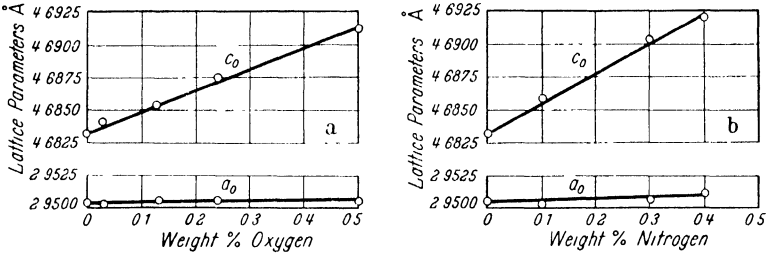


Fig. 7. The effect of the addition of small amounts of oxygen and nitrogen on the lattice parameter of alpha titanium (CLARK[118]).

Studies of the relationships between lattice parameter, hardness, resistivity, and volume changes have been carried on for a number of years by a French group. PORTEVIN, CHAUDRON and MOREAU[120] showed that the electrodeposited metals are much harder, have lower conductivity and have broader x-ray diffraction lines than annealed metals. By using the cold bombardment method for degassing, it was shown that the removal of hydrogen from electrodeposited iron altered neither the hardness, the lattice parameter, nor the diffuseness of the lines. Annealing after cold degassing caused no appreciable removal of gas, but did cause a decrease in parameter, a sharpening of the x-ray lines, and the expected softening of the metal. A special type of cold-working by the hydrogen atoms was postulated, the metal remaining in a strained condition after removal of the gas. Cathodically charged Armco iron was found to have a decreased tensile elongation and increased resistivity. Similar changes were found in the properties of aluminium after the same treatment.

In studying the palladium-hydrogen system, MICHEL[121] found that, on changing the composition from pure Pd to Pd_4H by the addition of hydrogen, the Brinell hardness increased from 47 to 170, while the lattice parameter increased from 3·882 to 3·885 Å, the hydrogen atoms occupying the body-centred positions in a face-centred cubic palladium

lattice. The beta phase, PdH, was found to have a Brinell hardness of only 70. On heating PdH at 310°C, hydrogen was slowly evolved and the hardness rose to 170 Brinell after about 2 hours. After that time the alloy with a composition corresponding to Pd_4H had a hardness which slowly fell to a value limited by the annealing temperature. When the PdH was degassed by cold bombardment, the lattice remained distorted and the hardness was unaffected.

The results which had previously been obtained for iron and aluminium were reproduced in the case of palladium and tantalum by CHAUDRON, PORTEVIN and MOREAU.[122] The results obtained for palladium were as follows:

Composition	*Ratio of Resistivities*	*Lattice Parameter*	*Brinell Hardness*
Pure annealed Pd	1	3·882	45
Beta phase	1·69	4·020	77
Beta phase degassed (ionic bombardment)	1·69	4·020	77
Beta phase degassed and annealed at 600°C	1	3·882	45
Alpha phase	1·45	3·885	170
Alpha phase degassed (ionic bombardment)	1·45	3·885	170
Alpha phase degassed and annealed at 600°C	1	3·882	45

MICHEL, BÉNARD and CHAUDRON[123] followed the changes of the lattice parameter and diffuseness of the x-ray diffraction lines as palladium was electrolytically charged with hydrogen. Since the beta phase appeared as soon as hydrogenation started, and from subsequent observations, it was concluded that a superficial layer of the beta phase is formed and the hydrogen diffuses from this phase into the metal to form the alpha solution. When the inner portion of the metal has reached a composition corresponding to Pd_4H, then the beta phase starts to advance into the body of the metal by diffusion of an excess of hydrogen over that amount required for the alpha phase. On exposure to air, hydrogen was found to leave the beta phase, the lattice parameter decreasing to a minimum value of 4·01 Å (at room temperature).

Finally MICHEL[124] carried out a thermal investigation on the hydrogen-palladium system by dilatometric and differential analysis. The results obtained were interpreted as being in agreement with the results of the previous x-ray analysis. The diagrams in Fig. 8, taken from MICHEL'S paper, show the dilatometric curves for both the alpha and beta phases. On heating in nitrogen, the alpha phase starts to contract at about 200°C, and the beta phase starts at about 100°C. In each case the contraction is complete at 300°C. This is shown in parts *a* and *b* of Fig. 8. The differential thermal analysis of the beta

phase, heated in nitrogen, is shown in part *d*. Starting at room temperature there is an absorption of heat due to the removal of hydrogen, and this is followed by a marked exothermic effect due to the contraction of the lattice at about 300°C. The same experiment repeated on a specimen which has been degassed by ionic bombardment will yield

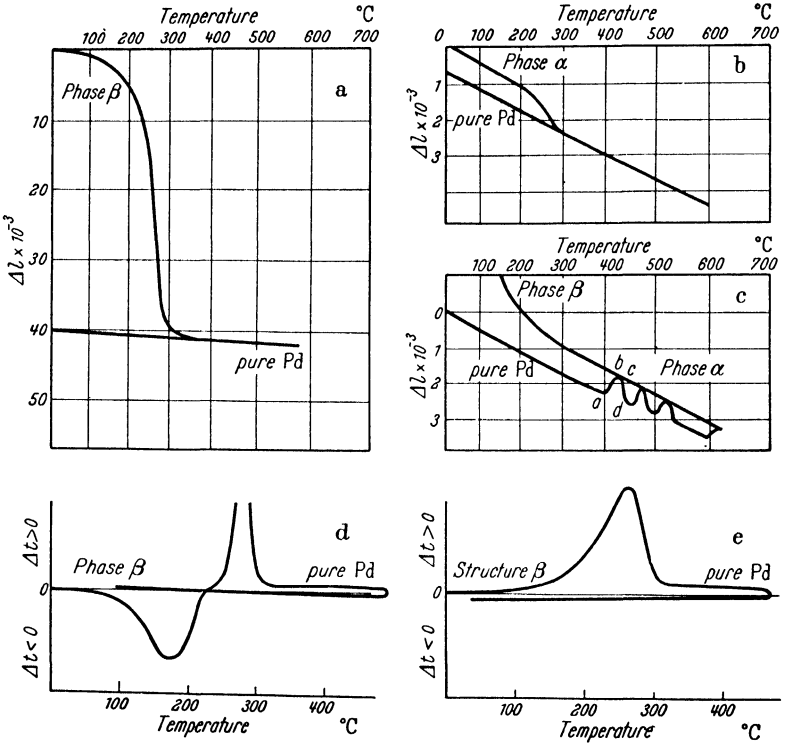


Fig. 8. Thermal and dilatometric relationships in the palladium-hydrogen system (MICHEL[124]).

the curve shown in part *e*. The exothermic effect persists, but no heat is required to remove the hydrogen.

PORTEVIN[125] briefly expressed the same ideas that the distortion of the crystal lattice persists after degassing, because of a "cold-working" effect by the gas atoms. The temporary introduction of gas atoms may be used to modify the resistivity of the metal.

Magnetic Properties

Although the present author has found no report of any effect of oxygen or nitrogen on the magnetic properties of metals, the effect of hydrogen is well established for palladium and iron. It has been found that the strong paramagnetism of palladium decreases, on absorbing

hydrogen, from the value of the susceptibility for pure palladium to a final value which is reached at a concentration of H/Pd = 0·55 to 0·66. The susceptibility remains approximately the same at concentrations higher than the rather uncertain limiting composition. It should be noted that the decrease in susceptibility is not due to the formation of the beta phase, since the addition of a small amount of hydrogen will lower the susceptibility prior to any possible beta formation. SMITH[2] has reported several explanations of this effect, the simplest probably being that of MOTT and JONES[126] who concluded that the hydrogen electron goes into the *d* shell, thereby decreasing the number of vacancies in the palladium *d* shell and hence the paramagnetism. Since the palladium atom has 0·55 holes per atom in the *d* shell, and since it had been shown that the magnetic susceptibility of the palladium was reduced to zero by the addition of about 0·55 hydrogen atoms to each palladium atom, MOTT and JONES also concluded that the dissolved hydrogen atom must be completely ionized. Since some investigators report the H/Pd ratio to be as high as 0·66 when the loss of paramagnetism is complete, SMITH[2] concludes that one-sixth of the hydrogen atoms remain uncombined and are contained "either in the interstices of the distended lattice of the beta phase or outside the lattice proper in the rifts. . . ."

SIEVERTS and DANZ[127] have shown that the effect of deuterium on the magnetic susceptibility of palladium is exactly the same as that of hydrogen. Thus it may be assumed that the deuterium atom also yields its electron to the *d* shell of palladium thereby rendering it diamagnetic.

VOGT[128] did experiments which indicate at least roughly that the addition of gold to a palladium-hydrogen system causes a similar decrease in the magnetic susceptibility as that brought about by the hydrogen. This lowering of the susceptibility is complete at 50 to 60 atomic per cent gold, and at this composition the hydrogen was found to be insoluble. VOGT concluded that gold and hydrogen have equal effects on the paramagnetism of palladium.

The most recent information which the writer has been able to find in the literature is a paper by FITZWILLIAM, KAUFMANN and SQUIRE[129] who studied the effect of hydrogen on the magnetic properties of titanium and zirconium. An interesting point is that, on absorbing hydrogen, the paramagnetic susceptibility of titanium increased. The reason for this was not explained. Only one composition (1·2 atoms of hydrogen per atom of titanium) was apparently considered. In the original zirconium specimen, an iron content of 0·11 per cent played no part in the magnetic susceptibility except at liquid hydrogen temperatures. However, on addition of hydrogen and with suitable heat treatment, "the iron appeared to precipitate out of solution with the zirconium in that it gave a ferromagnetic component to the magnetization." It is believed that the iron must have actually formed clusters in

order to cause the observed effects. However, the authors were able to separate the effect due to hydrogen alone, and that due to iron, on the susceptibility and they have presented a curve showing the paramagnetic susceptibility of zirconium as a function of hydrogen concentration. The curve is similar to that for the palladium-hydrogen system but differs in that, instead of being a linear decrease, the curve is concave toward the axes, and a constant value of susceptibility is reached at a positive, rather than a diamagnetic, value as shown in Fig. 9. It is

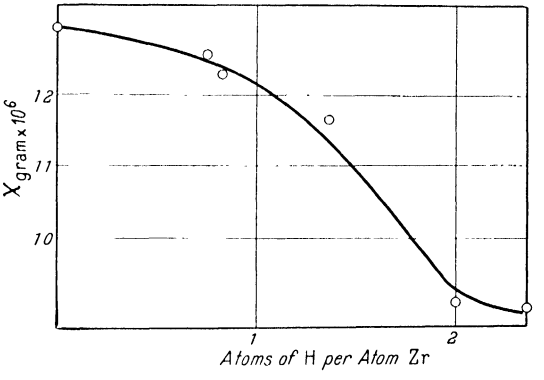


Fig. 9. Magnetic susceptibility of the system Zr–H as a function of hydrogen concentration at 295°K (FITZWILLIAM, KAUFMANN and SQUIRE[129]).

interesting to note that the Curie point of the precipitated iron was found at 585°K, instead of near 1000°K, as in the lattice of pure iron. Other experiments relating structure, temperature, hydrogen content, and susceptibility were also described.

The effect of interatomic distances on susceptibility has been applied by MICHEL in an attempt to explain the magnetic behaviour of the palladium-hydrogen system. In 1939 MICHEL and GALLISSOT,[130] using

TABLE 3

Composition	a_0	$\chi \times 10^6$	R/R_0	*Brinell Hardness*
Pure Pd	3·882	6·1	1	47
Pd_4H (alpha Pd)	3·885	5·1	1·45	170
Pd_4H degassed by bombardment	3·885	5·1	1·45	170
PdH (beta Pd)	4·048	0·01	1·68	70
PdH degassed	4·048	0·01	1·68	70
Pd_4H or PdH degassed and then annealed	3·882	6·10	1	47

R = electrical resistance

R_0 = resistance of pure Pd

the cold-bombardment method of degassing developed by CHAUDRON and MOREAU, determined the respective effects of hydrogen and interatomic distance in the variation of the susceptibility of palladium. The experimental results are given in Table 3 on page 135. These results seem to show that the magnetic susceptibility is unaffected by degassing. The authors attribute the large variations in susceptibility to the modification of the interatomic distances by the dissolved hydrogen.

The effect of hydrogen on the magnetic properties of those metals which absorb only small amounts of the gas is not so well explained as in the systems just discussed. SMITH[2] has reviewed the effects of hydrogen on the magnetic properties of iron and manganese, and concludes that the influence of hydrogen is probably due to its effect in cold-working the metal and also its purifying effect. It is indicated[2] that the susceptibility (and permeability) of iron at low field strength (about 10 oersteds) increases with decreasing hydrogen content, in much the same way as in the palladium-hydrogen system.

Extremely high permeabilities have been reported for iron annealed near 1500°C in wet hydrogen for 30 minutes, and then cooled rapidly to a temperature just below the A_3-point of pure iron, and held there for some time. The author attributed the high permeability to the high purity of the metal, due to the hydrogen treatment, and not to any residual hydrogen content. SMITH[2] claims that the effect is due to the annealing only, the atoms (of metal) being "no longer in the altered states in which they presumably exist in the vicinity of a rift." WALL[131] confirmed the improvement in the magnetic properties of iron by the hydrogen treatment, and found similar improvements in the case of cobalt and nickel.

WHEELER[132] found that the absorption of hydrogen by manganese lowered the susceptibility of alpha-manganese at 450°C, and beta-manganese at 850°C, but above 1200°C raised that of face-centred tetragonal or molten Mn. At no time was the metal reported to become ferromagnetic.

SMITH[2] concludes that the modification of the magnetic properties of those metals which do not dissolve large amounts of hydrogen must be due to some cold-working of the metal. At the same time it is recognized that a similar cold-working should take place in Pd, and this would be expected to alter the linear dependence so well explained by the electron theory of metals. This anomaly does not seem to have been explained.

Finally, SMITH[2] has attempted to show a reciprocal relationship between paramagnetic properties and capacity to absorb hydrogen. It is stated that all highly magnetic metals absorb measurable amounts of hydrogen, while elements with χ less than $0{\cdot}8 \times 10^{-6}$ do not, with

five exceptions, absorb hydrogen. SMITH has proposed that the relation between paramagnetism and hydrogen content may be due to some process which precedes a final arrangement. This process is postulated to be the "penetration of the metal" which factor is determined by "the rift-state." It is suggested that the "rifts" cause "changed electronic configurations" which in turn alter the susceptibility. Earlier, in explaining the sudden decrease in the magnetic permeability of iron after being cathodically charged with hydrogen for some time, SMITH said:

"Since saturation is that concentration of a phase which is in equilibrium with a certain pressure, we shall be making little more than a change of terminology if we suppose the critical stage to be that at which the pressure of the rift hydrogen, steadily building up toward the high effective pressure of electrolysis at high current density, reaches a value at which the lattice lets go and sudden enlargement of the whole rift system takes place, the resulting increase in interatomic spacings causing alterations of electron configurations in the atoms, which are, in turn, perhaps responsible for the changes in magnetic properties."

The present writer finds it very difficult to visualize a metallic lattice, in the so-called elastic range, "letting go".

SEEMAN[133] repeated some experiments done earlier by SHIMIZU on the effect of gas on the susceptibility and electrical conductivity of gold-silver alloys before and after heating *in vacuo*, heating in nitrogen, and remelting *in vacuo*. The above treatments had no effect on the magnetic properties of either the alloys or the pure metals. SEEMAN stated that the effect of oxygen on the susceptibility of diamagnetic and paramagnetic alloys is still not accurately known, but for precise determinations complete degassing of the specimens was recommended.

MAHN[134] has stated that dissolved gas has a marked influence on the value of the magnetism and, therefore, the alloys (the thallium-lead system was used as an illustration) must be rigorously annealed *in vacuo* after each measurement.

Electrical Resistivity

The present writer has been able to find very little information describing the effect of dissolved oxygen or nitrogen on the electrical resistivity of a pure metal. For those metals which dissolve only small amounts of hydrogen there appears to be no recent experimental evidence to show a certain effect. SMITH[2] has briefly reviewed the literature available on this subject up to 1946 and the reader is referred to SMITH'S supplementary notes (pp. 258, 262, 265, 267, 268, 275) for other references. Although the evidence seems to be rather uncertain, and the experimental work was done many years ago, it seems that the

resistivity of iron will increase a small amount on absorbing hydrogen from the gas phase. If the iron is charged with hydrogen cathodically then the resistance will increase to a limiting value of hydrogen content and thereafter starts to decrease. This effect is similar to that which occurs when palladium absorbs hydrogen, and will be described later. When nickel absorbs hydrogen from the gas phase, experimental results indicate any one of the following effects: (*a*) an increase, (*b*) a decrease, and (*c*) no effect on the resistivity. When the nickel was cathodically charged with hydrogen the effect is also doubtful. One report has indicated a decrease in resistivity due to "protonic conduction." The same remarks as have been made for nickel seem to apply also to the case of platinum. No effect of any kind on the resistivity of copper and silver has been found due to hydrogen.

As in the case of all other physical properties of the palladium-hydrogen system, the electrical resistivity seems to have been studied quite thoroughly. As in the case of other properties, SMITH[2, 135] has used the experimental evidence in the case of palladium-hydrogen to support his view that a large proportion of the hydrogen present in palladium is concentrated in "rifts." His arguments will not be presented here, but in preference to them the writer will review some recent work by WRIGHT[136] which is in agreement with the earlier x-ray work discussed by OWEN.[137] It is pointed out that the exceptional properties of the transition metals are due to the partly filled *d* shell. It was stated in the discussion of magnetic susceptibility that palladium has approximately 0·6 unoccupied states per atom in the *d* band. Generally it is accepted that the hydrogen dissolved in palladium is almost completely ionized, and its electrons enter this partially filled *d* band. Thus the palladium *d* shell is filled when the H/Pd ratio is 0·6. This seems to be true for palladium in equilibrium with the gas phase, but SMITH[2] has shown several cases where the H/Pd ratio is higher than 0·6 when the palladium is electrolytically charged with hydrogen. This has been explained by SMITH to be due to hydrogen being held in the widened lattice spacings of the metal lattice. WAGNER,[138] in calculating the hydrogen solubility from thermodynamic principles, has assumed that this excess hydrogen is located interstitially in solid solution, the electrons entering the *s* band of the palladium. WRIGHT has presented some theoretical ideas proposed by LACHER and WAGNER to account for the form of the concentration-pressure isotherms of hydrogen in palladium. It was concluded by WAGNER that the existence of a hysteresis loop is due to a failure to establish thermodynamic equilibrium. The relatively low temperatures of observation allowed only restricted exchange of palladium atoms between the two phases of the palladium-hydrogen system. Experimental studies of the hysteresis effect in the palladium-hydrogen system indicated that reversible

isotherms are obtainable only by using palladium in a finely divided state. Quite contrary to the definite opinions of SMITH[2, 135], WRIGHT has interpreted the work of SIEVERTS and his co-workers,[24, 26, 139, 140] and of OWEN and his co-workers,[116, 137, 141] to indicate that "the mixed phase or (alpha + beta) regions within the hysteresis loop are characteristic of a system in which the thermodynamic equilibrium is not completely established. The x-ray studies also indicate that the new phase, alpha or beta, exists finely dispersed throughout the old phase, the establishment of a new phase occurring, presumably, by the process of nucleation and growth." LACHER'S formula, which explains the hysteresis effect, is based on assumptions which are supported by the above conclusions. SMITH has claimed these assumptions to be invalid in the light of experimental evidence which appears to indicate that the two phases of the palladium-hydrogen system cannot exist interdispersed. WRIGHT'S experiments have definitely shown that the (alpha + beta) phase does exist over a small range of pressure.

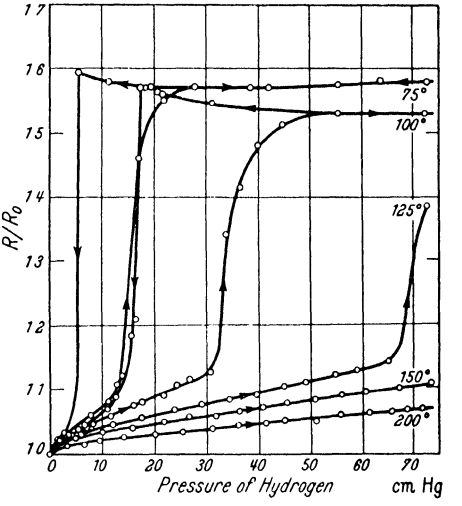


Fig. 10. Resistance-pressure isotherms for the palladium-hydrogen system (WRIGHT[136]).

WRIGHT determined resistance isotherms at several temperatures, allowing sufficient time after each pressure change for the resistance to become almost constant. The results of his measurements are shown in Fig. 10. It was shown that the limits of the hysteresis loops were quite close to the corresponding limits for the occlusion isotherms. The transitions from one phase to another have been described on the basis of x-ray, solubility, and electrical resistance measurements. If one accepts that the nuclei of a new phase are dispersed throughout the old phase, then the hysteresis effect may be explained by the contribution of the surface between the two phases to the free energy of the system. WRIGHT has described this mechanism:

"The vapour pressure of the dispersed beta-phase particles would be greater than the pressure they would exert in bulk, hence they could not form without some degree of supersaturation in the alpha phase. Similarly during evolution of hydrogen, an undersaturated beta phase

would be necessary for the formation of alpha-phase nuclei. Over a certain range of pressure up to the point where the isotherms rise (or fall) sharply, it must be assumed that the two phases co-exist in a state of metastable equilibrium. Beyond this stage, the degree of supersaturation in the alpha phase (or undersaturation in the beta phase) is such that the system becomes unstable and the new phase is

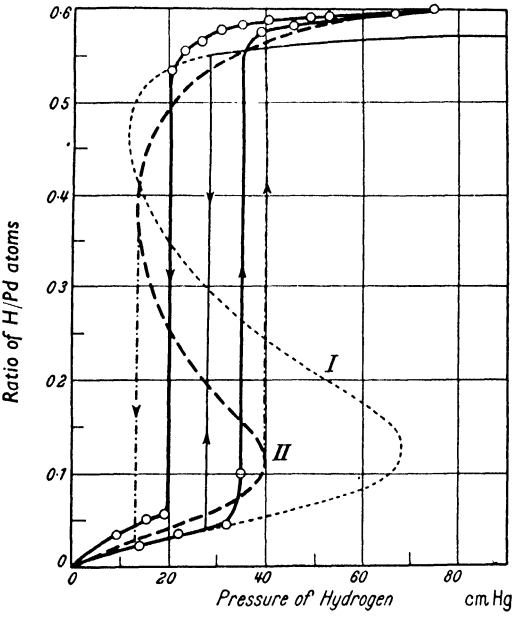


Fig. 11. Concentration-pressure isotherms at 100°C.
Experimental: SIEVERTS and DANZ, ———
Theoretical: I. LACHER - - - - - -, equilibrium ———
II. WAGNER — — —, path followed between upper and lower isotherms · — · — · — (WRIGHT[136]).

rapidly formed in fairly large blocks out of the finely dispersed particles."

The role of the hydrogen in each of the stages of the 100°C hysteresis loop shown in Fig. 10 has been clearly outlined by WRIGHT. The hydrogen in solid solution in the lattice will ionize, and the proton or hydrogen ion will increase the resistance by disturbing the periodicity of the lattice, while the electron will enter the unfilled *d* band of the palladium. These are opposing effects, the former tending to increase the resistivity, and the latter to decrease the resistivity. From the close correspondence between the concentration isotherm (Fig. 11) and the resistance isotherm, WRIGHT has deduced that the scattering centres provided by the hydrogen ions have a greater effect and, therefore, increase the resistance of the palladium-hydrogen system. It was

also indicated that the increase in resistance due to cold working is small compared with this alloying effect, so that lattice distortion at the phase boundaries will be relatively unimportant in altering the resistance. Finally, the effect on the resistivity of a new phase in the system will be greatest when the phase particles are of comparable dimensions to the wavelength of the conduction electrons.

In the first stage of the resistance-pressure isotherm at 100°C, the relatively greater effect due to decreased periodicity of the lattice caused by the hydrogen ions entering the metal, over the effect of the electrons entering the *d* shell, will result in an increase in the resistivity. When the beta phase starts to form at about 27 cm Hg, there is a slight increase of resistance due to strains at the phase boundaries. At about 31 cm Hg, the system transforms completely to the beta phase. If this transformation were brought about by very small increases in hydrogen pressure with alternate long time periods allowed for equilibrium conditions to be reached, the vertical portion of the curve would become almost parallel to the resistance axis. In the beta phase, the *d* shell is filled and an increase in pressure appears to cause no further marked change in the resistivity up to normal pressure. By comparing Fig. 10 with Fig. 11, one sees that there is an increase in resistance as the hydrogen content is decreased, i.e. as the pressure over the beta phase is lowered from normal pressure to 17 cm Hg. Since the alpha phase does not form until the pressure is decreased to about 19 cm Hg, there can be no increase in resistance due to strains, but WRIGHT has attributed it to an increase in the number of positive holes in the *d* band of palladium. As the pressure is decreased to 17 cm Hg, the system changes to the alpha phase. With the removal of hydrogen the resistance drops, since the fall due to reduced hydrogen scattering is a more important factor than the removal of electrons from the *d* shell which tends to increase resistance. Under ideal conditions this decrease in resistance would be shown as a line parallel to the resistance axis. This seems to the present writer to be a reasonable explanation of the phenomena which have been observed.

Some time ago ANDREWS[142] used electrical resistance measurements in conjunction with pressure-volume relationships to study the reaction of nitrogen, oxygen and hydrogen with tantalum. Fig. 12 shows the variation of resistance at room temperature with gas content of the metal. Unfortunately, no theory was suggested to explain these curves. It was found that in nitrogen, absorption increased with increased gas pressure until about 100 relative volumes were absorbed. At this point the equilibrium pressure dropped as absorption increased, although this had no effect on the resistivity. It was noted that the resistance at room temperature rises less rapidly with hydrogen concentration than for either oxygen or nitrogen. The reason given was: "It can be

assumed that this is due to the small atomic volume of the hydrogen, so that a given number of molecules dissolved in the metal would strain the lattice less than larger atoms." Probably ANDREWS meant that a given number of *atoms* of hydrogen would strain the lattice less than oxygen or nitrogen. It is now generally accepted that the dissolved gas is not present as molecules. The sudden change in slope of the oxygen curve was attributed to the formation of an oxide compound, but no further explanation was given.

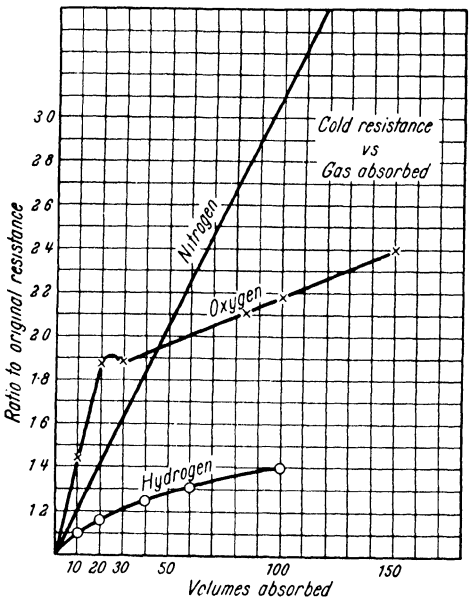


Fig. 12. Resistance at room temperature vs. gas content. Resistance given as the ratio to the resistance of the gas-free filament (ANDREWS[142]).

Later, WEBBER,[143] basing his experimental treatment on the results of ANDREWS' work, showed that gas content (particularly oxygen and nitrogen since the specimens were annealed in air) is not a major factor affecting the superconductivity of tantalum. In this case the gases were absorbed from the gas phase.

In more recent work GOLIK, LAZAREV and KHOTKEVICH[144] have found that the temperature range of supraconductive transition is broadened as a result of electrolytically charging tantalum with hydrogen. The broadening increases with the amount of hydrogen in the tantalum. At any given temperature between 1·85° and 4·2°K, the electrical resistance ratio $R_T/R_{4\cdot2}$ was found to increase with the amount of occluded hydrogen. At temperatures down to 1·86°K, it was found that tantalum specimens could be made to lose supraconductivity altogether by saturation with hydrogen. The upper limit of the transition range remained unchanged, all samples showing a marked decrease in resistivity at that point. The condition of the specimens appeared to control the amount of hydrogen which could be occluded. For example, 0·340 mg of hydrogen was required to suppress supraconductivity in a new tantalum sample. This wire was heated 10 hours at 1700°C *in vacuo*, and its supraconductivity was completely restored. After this anneal, only 0·005 mg of hydrogen was required to suppress supraconductivity. In contrast to tantalum, it was observed that

the supraconductivity of niobium was held at 4·2°K after charging with hydrogen. It was concluded that this effect of hydrogen on tantalum could not be due to simple expansion of the lattice, but must be due to the formation of solid-solution alloys.

DRAVNIEKS[145] used the change in electrical resistance as a means of measuring the rates of reaction of zirconium and copper with both ionized and non-ionized gases. The study was mainly one of surface reactions such as oxidation. The resistance changes noted in the copper specimens were taken to be caused by the formation of a cuprous oxide film which is practically non-conducting compared with copper. The conductivity of the zirconium changed as the reaction with gas proceeded due to: (*a*) the formation of a zirconium-gas solid solution of lower conductivity than the pure zirconium, and (*b*) the formation of an oxide scale of zirconium. The change of resistivity was used merely as a measuring device, and again no theory was proposed to account for the effect.

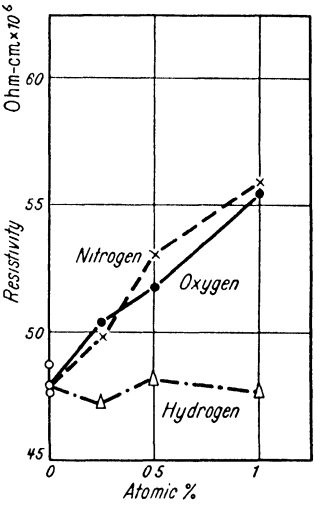


Fig. 13. Effect of oxygen, nitrogen and hydrogen on the electrical resistivity of iodide titanium (JAFFEE and CAMPBELL[146]).

Recently JAFFEE and CAMPBELL[146] have reported that both oxygen and nitrogen raise the resistivity of titanium to the same extent per atomic per cent. Hydrogen up to one atomic per cent does not seem to change the resistivity. The results obtained are shown in Fig. 13.

There seem to have been relatively few studies reported in the literature on the effect of gases on the electrical resistivity of pure metals other than palladium. Such work might prove to be very significant in showing the nature of the interaction of gases with metals.

One might expect dissolved oxygen to lower the conductivity of pure copper. It will be shown later in this section that oxygen will increase the conductivity of slightly impure copper by removing the impurities from solid solution. SMART, SMITH and PHILLIPS[147] have shown that there is a very small decrease in the conductivity of pure copper due to dissolved oxygen. Their results are shown in Table 4 on page 144.

It is seen that there does seem to be a loss in conductivity due to the dissolved oxygen, but the relationship between the data obtained at the two annealing temperatures is inconclusive.

In the course of work on internal oxidation, MEIJERING and DRUYVESTEYN[148] determined the effect on the electrical resistance of annealing

TABLE 4

Conductivity of Samples Annealed for One Hour at Indicated Temperatures and Cooled by Rapid Quenching

Sample No.	*Description*	*Conductivity per cent*	
		Annealed at 500°C	*Annealed at* 850°C
A	Oxygen-free	102·32‡	
B	Oxygen-bearing (saturated at 850°C)	102·23†	102·10†
C	Oxygen-bearing (Check run)	102·21†	102·23†
D*	Oxygen-free	102·39‡	
E*	Oxygen-bearing	102·19†	

* Measurements on these samples were made in another laboratory.
† Annealed in nitrogen.
‡ Annealed in hydrogen.

silver- and copper-base alloys in oxygen. Table 5 gives in semi-quantitative terms the results of this work. Factors such as oxidation temperature, and time of annealing before or after oxidation affect the resistivity. These factors are not considered in the table so the results are semi-quantitative.

TABLE 5

Resistivity in 10^{-6} Ω cm at 0°C of Annealed (ρa) and Oxidized (ρo) Silver and Copper Alloys

Description of Alloy	ρa	ρo	$\frac{\rho o - \rho a}{\rho a} \cdot 100$
Silver + 0·33 per cent magnesium	2·05	2·32	+ 13 per cent
Silver + 0·15 ,, ,, beryllium	1·56	1·92	+ 23 ,, ,,
Silver + 0·26 ,, ,, aluminium	3·45	2·39	− 31 ,, ,,
Silver + 0·60 ,, ,, titanium	2·39	1·94	− 19 ,, ,,
Silver + 0·79 ,, ,, zinc	2·05	1·65	− 20 ,, ,,
Pure silver*	1·75	1·57	− 10 ,, ,,
Copper + 0·75 per cent magnesium	2·30	1·72	− 25 ,, ,,
Copper + 0·10 ,, ,, beryllium	2·08	1·70	− 18 ,, ,,
Copper + 0·69 ,, ,, beryllium	2·95	1·76	− 40 ,, ,,
Copper + 0·81 ,, ,, aluminium	3·17	1·78	− 44 ,, ,,
Copper + 1·1 ,, ,, titanium	8·45	1·78	− 79 ,, ,,
Pure copper	1·61	1·62	0 ,, ,,

The percentages are weight per cent. The silver-zinc alloy was annealed in nitrogen to cut down evaporation of zinc.

* Evidently not very pure.

In the case of the copper alloys, the resistance after oxidation of the alloy approaches that of pure copper. Only the non-hardening zinc

alloy in the silver-base group decreased to the resistivity of pure silver. The alloys with titanium and aluminium had a resistivity well above that of pure silver and those of magnesium and beryllium increased above the value for the annealed alloy.

The two main factors which influence the resistance have been described by MEIJERING and DRUYVESTEYN as:

"(1) Internal strains cause generally an increase of the resistance as the lattice is distorted irregularly by these strains. This effect causes a decrease of the resistance by annealing. TAMMANN and DREYER[149]* found that by annealing after 98 per cent rolling the resistance of pure

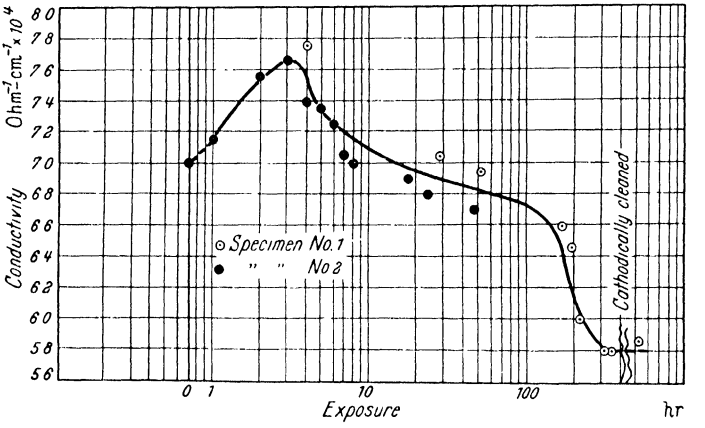


Fig. 14. Effect of atmospheric exposure on electrical conductivity of a Mg-Al-Zn alloy. Such exposure results in the addition of hydrogen to the magnesium lattice (BUSK and BOBALEK[119]).

silver decreases by 5 per cent, the decrease of pure rolled copper after annealing being only 1·6 per cent.

"(2) A small number of foreign ions, substituted on the normal lattice places, increases the resistance, this effect being for copper and silver alloys the larger the higher the valency of the alloying element,[150]* as ions with a high electric charge will scatter the electron waves considerably. When precipitation occurs and larger particles are formed the resistance will decrease. It is uncertain whether the scattering of electron waves in a silver lattice is highest on magnesium ions or on magnesium oxide molecules."

Thus oxidation of the alloys may be expected to involve both these factors to some degree. It was suggested that the lattice distortion on oxidation would probably increase, causing an increased resistivity, while the electron scattering near the solute ions probably decreases, tending to cause a decreased resistivity. If, on oxidation, very large

* This number refers to the list of references at the end of the present paper.

particles of oxide are precipitated then the strain in the lattice will not be sufficient to increase the resistivity. If the scattering of electrons near the solute ions is large as in the case of trivalent or quadrivalent ions like aluminium and titanium, then there will be a marked tendency for the resistivity to decrease when the solute element is oxidized. In silver alloys with magnesium and beryllium in solution, the increase in resistance, due to increased distortion in the lattice after oxidation, appears to be a more effective factor than the tendency to decrease the resistivity by decreasing the electron scattering. In all other cases the effect of the decrease of electron scattering outweighs the distortion effect.

Also considering alloys, BUSK and BOBALEK[119] determined the effect of hydrogen on the conductivity of magnesium alloys. The results of their investigation are shown in Fig. 14. The last point on the diagram was obtained after the specimen was saturated with hydrogen after it was charged for 4 hours in an alkaline solution. Prior to saturation the specimen was exposed to a normal atmosphere for the times shown, during which period hydrogen was absorbed by the lattice as indicated in another experiment. No suggestions were given to account for the curve. It was shown that the initial rise in conductivity was not due to the precipitation of a second phase during the initial stages of gassing.

Miscellaneous Effects

Other interesting effects caused by gases in various metals are discussed in the remainder of this section. Since each effect is a relatively isolated phenomenon, only a small amount of study having been devoted to each one, this section has been separated from the others.

The analogy between the solid solution of gases in metals, and liquid solutions has been demonstrated by experiments showing that the electrolysis of these solid solutions is quite possible. COEHN and his co-workers studied the hydrogen-palladium system[151] and SEITH and KUBASCHEWSKI[152] reported a similar effect in the iron-carbon system. In each case the non-metallic element moves in the solid metal as a positive ion under the influence of the electric field. This agrees with the idea that hydrogen is ionized to form an electron and a positive hydrogen ion when it enters the metal. More recently DE BOER and FAST[50] have found that under suitable conditions of temperature and applied potential difference, the oxygen in solution in zirconium wires could be made to diffuse through the metal lattice as a negative ion. In one experiment a wire with 5 atomic per cent of dissolved oxygen was heated to 1640°C by passing direct current through it. After some hours the temperature had dropped at the negative pole and had risen to a white heat at the positive end of the wire. Subsequent x-ray analysis indicated a concentration of 16 atomic per cent of oxygen at

the positive end and a gradual transition to almost pure zirconium at the negative end. In another experiment the oxygen was driven back and forth at will, simply by reversing the polarity of the applied potential. It was predicted by BARRER[3] that oxygen would behave as a negative ion.

In the same work[50] it was shown that the lattice constants a and c of the hexagonal zirconium lattice both increase up to 10 atomic per cent of oxygen. Beyond that point only the c value increases. Measurements were made on specimens containing up to 30 atomic per cent oxygen. This greater increase of c is similar to the case for titanium reported more recently by CLARK,[118] although CLARK measured only to 0·5 per cent of dissolved gas. The specific gravity of the zirconium-oxygen samples was measured and it was found that this increased over 5 per cent while the volume increased less than 2 per cent after absorption of 30 atomic per cent oxygen. A calculation of the specific gravity was made assuming that all the oxygen is contained in the interstices of the zirconium. The calculated values agreed very well with the experimental data. From all this information, and from previous work by FAST, it was concluded that all oxygen atoms are contained in solid solution at interstitial positions. More recently, HALL, MARTIN and REES[153] have confirmed by x-ray analysis that in samples containing up to 36 atomic per cent oxygen, the oxygen was in true solid solution in the zirconium.

JAEGER, ROSENBOHM and VEENSTRA[154] have reported that the specific heat of silver freshly melted in air, or heated in air at 940°C for 5 to 10 hours was higher than that of silver which had been melted in a vacuum, or had been stored for several months after melting in air. Their results showed a marked change in specific heat when gas-free silver was heated in air above 650°C. This effect was ascribed to the presence of dissolved oxygen.

Later, KELLEY[155] determined the specific heat of pure tantalum, and of tantalum containing small amounts of hydrogen in solution over the temperature range 80°K to 295°K. The tantalum containing hydrogen behaved anomalously, the specific heat having a high maximum in the region 220° to 265°K, depending on the amount of hydrogen in solution. The height of the maximum and the difference in measurable heat content between pure tantalum and that containing hydrogen were found to be about proportional to the amount of hydrogen in solution. The temperature at which the maximum occurs increases as the hydrogen content increases. When specimens containing hydrogen were degassed the specific heat returned to the normal value for pure tantalum. KELLEY has postulated that differences in entropy between hydrogenated and gas-free tantalum may be attributed to differences in

the number of possible ways of arranging the hydrogen ions in the tantalum lattice at the higher and lower temperatures.

The effects of gases on the mechanical properties of metals will not be considered in great detail here. It is felt that much of the experimental work has been done on metals having metallic impurities which would make the results of the work insignificant. In some cases unsuspected internal oxidation may have occurred which would seriously distort the results of the experiment. In all cases where a second phase is formed (oxide or nitride) the hardening or embrittling effect is well known and need not be discussed. The effect of hydrogen, oxygen or nitrogen in true solution in those metals which only dissolve small amounts of the gas does not seem to be really well known. Of those metals which dissolve relatively large amounts of gas, titanium has been the subject of the most intensive study.

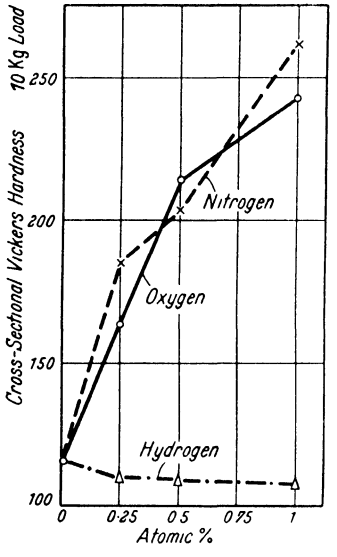


Fig. 15. Average cross-sectional hardness values for alloys of iodide titanium with oxygen, nitrogen and hydrogen (JAFFEE and CAMPBELL[146]).

JAFFEE and CAMPBELL[146] have studied the effect of oxygen, nitrogen and hydrogen on the mechanical properties of iodide-refined titanium. Fig. 15 shows a plot of the average cross-sectional hardness of the specimens versus their gas contents. The results indicate that nitrogen may harden titanium to a slightly greater extent than oxygen, while hydrogen actually appears to decrease the hardness slightly. This apparent decrease was felt to be due to hardness variations in the original titanium rod. Fig. 16 shows that, for the range of solubility considered, nitrogen increases the tensile strength more than oxygen, and correspondingly decreases the ductility more. The apparent high ductility found with 0·5 per cent of hydrogen was felt to be due to specimen variation and not to any effect of the gas.

FINLAY and SNYDER[156] have recently studied the effect of nitrogen, oxygen and carbon on the mechanical properties of titanium. This work is similar to that discussed above but is more complete. In general this more recent study has confirmed the results obtained by JAFFEE and CAMPBELL. In contrast to the substitutional alloys, where the solute element with the lowest solubility has the greatest strengthening effect per unit concentration, it was found here that the interstitial

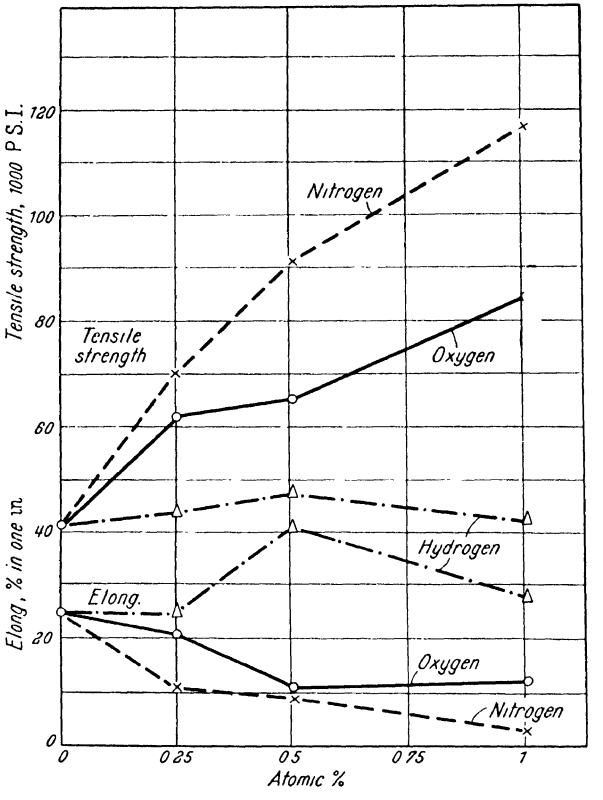


Fig. 16. Tensile strength and elongation of alloys of iodide titanium with oxygen, nitrogen and hydrogen (JAFFEE and CAMPBELL[146]).

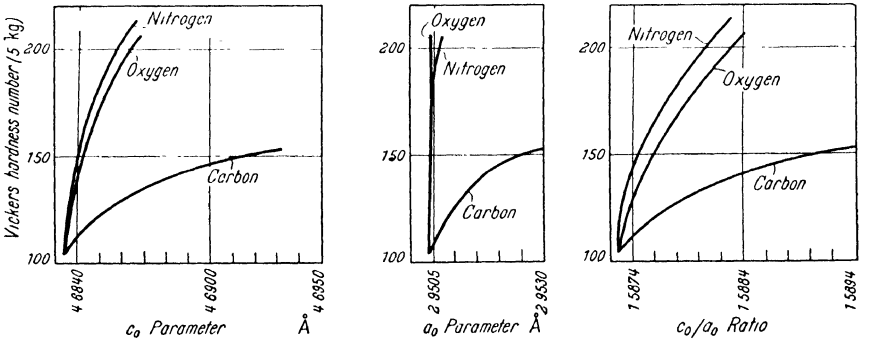


Fig. 17. Correlation of Vickers hardness with the lattice parameters and C_0/A_0 ratios of annealed, high-purity, binary, titanium-base alloys containing nitrogen, oxygen and carbon (FINLAY and SNYDER[156]).

element with the lowest solubility of the three studied (carbon) has the lowest strengthening effect. However, it was also noted that the element (oxygen) with the highest solubility does not have the greatest hardening effect. Also JAFFEE and CAMPBELL showed that hydrogen, with a greater solubility in titanium than carbon, oxygen or nitrogen, has almost no strengthening effect. There does not appear to be sufficient information available to make a generalization at this time. FINLAY and SNYDER have studied the relationship between the amount of hardening and the extent of lattice-parameter increase effected by additions of the interstitial solute elements. This relationship is shown in Fig. 17. It was suggested that oxygen and nitrogen may behave similarly in their interactions with the bonding electrons in the alpha solid solution, thus accounting for the correlation between mechanical properties and parameter changes.

A recent paper by WAIN and COTTRELL[157] describes one of the most interesting effects which gases have on metals. In this work it was shown that zinc crystals containing nitrogen have a sharp yield point when sufficiently strain-aged, while no yield point was found in nitrogen-free zinc. References are given in the paper to others which have shown that nitrogen also causes a yield point in iron, cadmium single crystals, and molybdenum polycrystals. OROWAN had observed a similar effect in zinc which he termed "thermal hardening," but he did not attribute the yield point to the presence of a specific solute element. Unlike cadmium crystals, the zinc crystals would not absorb sufficient nitrogen from the atmosphere during growth, so the nitrogen had to be introduced in a preliminary operation, by melting the zinc under a flux or by bubbling the gas through the zinc. It was definitely shown that the effect was due to nitrogen present internally and not merely as a surface film. Stress-strain curves, made after various ageing treatments at room temperature, show that the elastic limit decreases due to recovery and the yield point does not appear for several hours. Almost complete recovery precedes the return of the yield point, which appears after ageing for 8 or 9 hours at room temperature. Ageing for longer time periods causes an increase in the yield point. At higher temperatures the yield point appeared after shorter ageing times. From experimental results it was deduced that the process is thermally activated, with an activation energy of 21,000 cal/mole. The orientation of the crystals was found to have no effect on the magnitude of the yield point. The theory presented to explain this effect was that: ". . . the yield point is caused by the segregation of solute atoms to dislocations. The attraction of the dislocations to the segregated atoms provides a bond which has to be broken by a larger force than is necessary to maintain

freed dislocations in motion; the material thus gives way suddenly and softens, at the start of plastic flow, producing a sharp yield point. The theory also explains the observed removal of the yield point by plastic overstrain and its return on strain-ageing. A freshly strained specimen contains freed dislocations and does not show a yield point, but on ageing these dislocations become anchored by the migration of solute atoms to them and the yield point returns."

It is believed that, in each of the metals in which the yield phenomenon has occurred, the nitrogen is present in interstitial solution so that it distorts the lattice of the metal and diffuses very rapidly, conditions which must be fulfilled to agree with the theory suggested. It was predicted that marked yield points may not be found in face-centred cubic metals "because of the difficulty of anchoring screw dislocations in this type of lattice."

It was previously pointed out that one consequence of the theory of nitrogen atoms migrating to dislocations (in alpha-iron) is that the electrical resistance should decrease by a small, predictable amount on strain-ageing. This effect was examined by COTTRELL and CHURCHMAN,[158] and was found to exist.

SUBSURFACE OXIDATION OR INTERNAL OXIDATION

Although, as mentioned earlier, surface oxidation will not be discussed, the phenomenon of internal oxidation is considered to be within the scope of the present review.

Generally, if an alloy contains a solute element with greater affinity for oxygen that the solvent metal, the minor constituent may oxidize below a surface oxide formation, in an area where the partial oxygen pressure is too low to cause oxidation of the major element. This subsurface oxide will not flake off as may the surface layer of oxide of the solvent metal.

Although this effect was observed by TURNER[159] as far back as 1912, relatively little has been published until recently. Although they both observed internal oxidation in commercial copper during annealing, BLAZEY[160] and later WYMAN[161] did not consider this to be as important as other factors causing embrittlement. C. S. SMITH[162] is responsible for the term "subscale" which is used to denote the layer of subsurface oxide, and he appears to have been the first to make quantitative observations of the phenomenon. FRÖHLICH[163] has remarked on the occurrence of internal oxidation of several copper alloys but did not consider the problem in detail. LEROUX and RAUB[164] found a precipitation structure of both cupric and cuprous oxide in silver-copper alloys. The cupric oxide occurred in alloys poor in copper, and the cuprous oxide in those rich in copper. Most of the experimental work in this field has been done with copper as the solvent metal, but MEIJERING[165]

describes the properties of alloys of silver, copper, nickel, and iron respectively, with small amounts of other metals, after treatments leading to internal oxidation.

In one paper, RHINES[60] has classified the modes of occurrence of subsurface oxidation as one or a combination of three types: "(1) general precipitation of separate particles of the oxide of the alloying element within a matrix of the major metal, (2) deposition of relatively thin layers of the oxide of the alloying element along the grain boundaries or special planes of the major metal, or (3) envelopment of particles of the major component in a matrix of the oxide of the alloying element." The first type, which is generally indicated when one mentions "internal oxidation," occurs in dilute alloys at high temperatures. The second type occurs at lower temperatures with greater as well as smaller amounts of the alloying metal. The third type can only occur when there is a relatively large amount of the oxidized metal present. In the same paper RHINES has described some very interesting practical aspects of subsurface oxidation.

The actual mechanism of internal oxidation seems to be quite simple. As prerequisites for internal oxidation to occur, RHINES[166] has indicated that there must be a solubility of oxygen in the alloy and the oxygen must be able to diffuse fairly rapidly into the alloy. The minor constituent must form an oxide which is more stable than the oxide of the major component, and which is insoluble in the alloy. If one considers the system as a ternary alloy, then the constitution of the system must allow a quasibinary equilibrium to occur between the solid solvent and the oxide particles at the oxidation temperature.

Fig. 18 shows typical internal oxidation. At the bottom of the picture one may see a relatively thin dark layer which probably consists mainly of cuprous oxide in this case, with an oxide of manganese. This formation is not actually described by RHINES, but would be analogous to one of a copper-silicon alloy described in an earlier paper. The second, and larger, region consists of copper-manganese solid solution plus a precipitate of a manganese oxide. It is this layer which is the area of internal oxidation. Finally, the region at the top of the photograph is simply a copper-manganese solid solution. By oxidizing at a low oxygen pressure a subscale may be formed without oxidizing any of the solvent metal, thus leaving an apparently clean surface even though a considerable depth of oxidation has occurred. This effect is shown in Fig. 19.

To explain the mechanism of internal oxidation RHINES[166] has used the diagram shown in Fig. 20.

As has been noted earlier in this paper, there is a considerable solubility of oxygen in copper. Thus it is possible for the diffusing oxygen in the metallic lattice to come into contact with the alloying

Fig. 18. Typical Internal Oxidation. Alloy of copper + 0·4 per cent manganese oxidized at 1000°C in air for 8 hours × 50 (RHINES[9]).

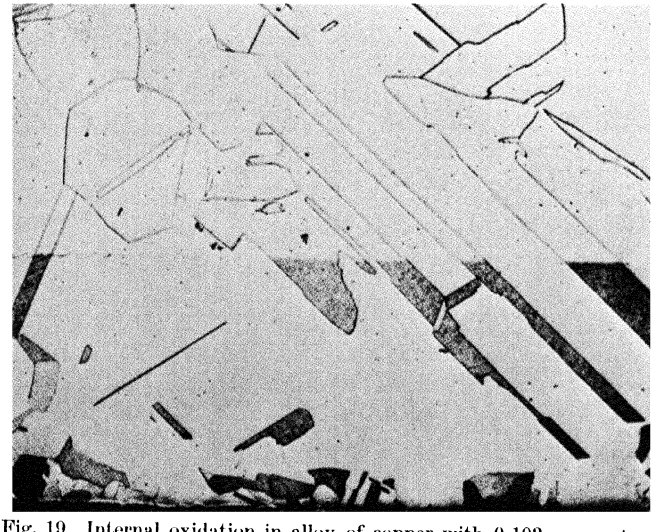


Fig. 19. Internal oxidation in alloy of copper with 0·103 per cent silicon × 50. Sample oxidized at low oxygen pressure has no external scale. Internal oxidation is marked by zone crossing lower half of picture (RHINES[168]).

element. If the oxide of the alloying element is more stable than cuprous oxide, under the conditions of temperature and oxygen pressure imposed, then the solute metal oxide will form. This preferential oxidation of the alloying element may take place beneath the external oxide scale or, if the external oxygen pressure is not sufficient to form a scale, it may extend to the apparently clean surface (Fig. 19). Once the solute metal oxide has formed it will not remain in solution very long but will separate from the lattice as a precipitate. As the solute metal is exhausted by this precipitation, the oxygen supplied to the surface of the subscale (either from the external atmosphere or from the copper oxide layer) diffuses farther into the alloy and thus produces a zone of

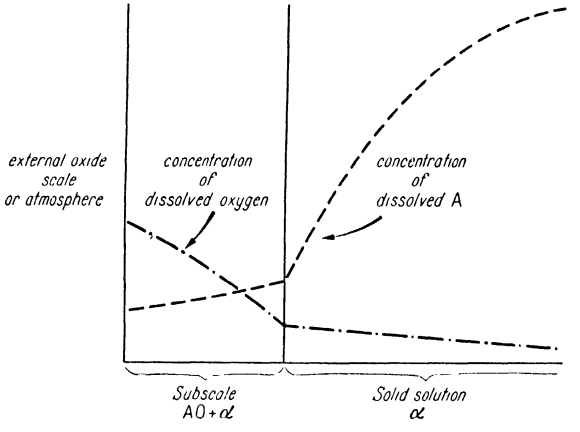


Fig. 20. Schematic representation of distribution of dissolved oxygen and alloying element A in subscale and unoxidized solid solution (RHINES[166]).

ever-increasing thickness. At the same time, the solute metal component must diffuse in the opposite direction toward the advancing front of scale to replace the decreasing concentration of solute atoms being precipitated as oxide particles. Thus RHINES[9] points out that the rate of subscale growth depends, among other conditions, upon the diffusion rates of both oxygen and the reacting solute element. To confirm this theory, it has been found that in many alloy systems, the square of the subscale depth is a linear function of the time of reaction at constant temperature. It was shown in another paper[167] that, if melting does not occur, internal oxidation of copper alloys occurs between 750° and 1000°C as a relatively simple diffusion process involving the movement of gas and metal atoms. This process was described by mathematical expressions. The oxidation process investigated at 600°C was found to be more complex. DARKEN[168] has also discussed the mechanism of subscale formation. He has shown how the

rate of growth and composition of the zone may be predicted from a knowledge of diffusion constants, phase relations, and boundary conditions.

RHINES found that the particle size of the precipitate depends upon the chemical nature and concentration of the solute element and the temperature of oxidation. The more stable oxides tend to form fine precipitates, while the less stable ones precipitate as coarse particles. High temperature and high concentrations of the alloying element both promote large particle size. Usually the oxide particles are well distributed throughout the zone of oxidation.

The precipitate distribution depends on the system and the temperature of oxidation. In some cases the density of oxide particles is greater near the metal surface, in other cases the reverse is true, depending on the relative diffusion rates. As the temperature of oxidation is lowered the oxide particles tend more to occur at the grain boundaries rather than in the uniform distribution characteristic of high temperature oxidation.

RHINES has also found the *rate* of formation of the zone of oxidation to decrease with time, with an increased content of the solute element, with lower temperatures and with decreased oxygen pressure. The time to form a given thickness of subscale increases almost directly with increase in the concentration of the solute metal.

In ternary alloys, where the two solute elements form oxides of different solubilities, double zones of internal oxidation occur. The outer zone contains the oxides of both alloying elements, while the inner zone contains only the more stable oxide of the two. By altering the alloy composition one may change the dimensions of the subscale. By changing the oxygen pressure selective oxidation of the solute element forming the most stable oxide may be promoted.

In 1947 RHINES[169] summarized his work, and in the review mentioned that the surface hardening of alloys due to the presence of uniformly distributed oxide precipitate does not seem to be possible. However, in 1945, CHASTON[170] had already noted what he called "precipitation-hardening effects" and used associated properties to develop his theory of water vapour embrittlement of silver (see section on Embrittlement). CHASTON was able, by careful oxidation, to obtain a marked increase in the Vickers hardness of silver alloys with small amounts of aluminium or zinc in solution. It was noted that there is an optimum solute content required to give the maximum hardening effect for a given oxidation treatment. Larger amounts of the solute in solution appear to hinder the diffusion of oxygen into the subsurface layers of metal. It was found that the depth of the oxide subscale was greater in alloys of low solute element content, and the depth could be increased by increasing the oxidation time. Commercially pure silver could be

hardened by heating in air followed by heating in an inert atmosphere containing zinc vapour. The zinc vapour apparently diffused into the silver to combine with the dissolved oxygen, thus causing a hardening effect. Spectroscopically pure silver, when solid, could not be hardened in this way, since it will not, according to CHASTON, dissolve oxygen (see section on Embrittlement). Similar hardening was obtained by heating silver in air containing zinc vapour. In this case the zone of internal oxidation formed by heating silver for 4 hours in air containing zinc vapour was found to extend 0·04 in. below the metal surface. Shortly after RHINES' review was published, MEIJERING and DRUYVESTEYN[148, 165, 171] reported their finding substantial surface hardening caused by internal oxidation in certain binary alloys. This hardening effect has recently been confirmed by SMITH and DEWHIRST.[172]

In their paper MEIJERING and DRUYVESTEYN point out that "dispersion hardening" by blocking the glide planes with minute particles is effectively carried out by internal oxidation. By heating strips of silver alloys for 2 hours in air at 800°C they obtained the following results:

Silver + 0·3 weight per cent magnesium + 170 Vickers (kg/mm², load: 2½ kg)
Silver + 0·4 weight per cent aluminium + 160 Vickers
Silver + 0·2 weight per cent beryllium + 135 Vickers

Similar strips annealed in nitrogen or hydrogen had hardnesses of from 30 to 50 Vickers.

The success of a hardening operation depends on the difference in stability of the oxides of the two components of the metal alloy. Experimental results have shown that although the oxidation of a given percentage of a solute element will increase the hardness of silver, the same treatment may not harden copper to the same extent. Similarly an oxidation treatment which hardens a copper-base alloy may not harden a nickel-base alloy due to the greater stability of nickel oxide. The difference between hardening and non-hardening alloys has been explained on a thermodynamical basis by MEIJERING and DRUYVESTEYN.

With respect to the precipitation of the oxide particles, on which depends the resultant hardening effect, MEIJERING and DRUYVESTEYN have shown that at very high magnifications (1200 ×) no particles were visible. The oxide will be first formed in a molecular state, the molecule being immobile in the alloy lattice. After some time period this molecule may dissociate (the rate decreasing with increase in stability of the oxide) and the atoms of the oxide will move to "form greater (and, therefore, more stable) complexes." The less stable oxides produce coarser particles which have practically no hardening effect.

X-ray measurements showed that the lattice spacing of silver containing 0·36 per cent magnesium, and silver containing 0·26 per cent aluminium increases considerably by oxidation to a value larger than

that of either pure silver or of the unoxidized alloy. The influence of dissolved oxygen in the silver is quite small. This result was interpreted to indicate that a considerable part of the magnesium and aluminium oxides are dissolved in almost molecular dispersion in the silver lattice. Those alloys of silver with zinc and cadmium which did not harden had a spacing about the same as that of pure silver, which indicated the formation of larger oxide particles. Copper alloys which have been hardened retain a lattice spacing almost equal to that of pure copper.

Electrical resistance measurements made by MEIJERING and DRUYVESTEYN show that the resistivity of the oxidized copper alloys decreases to a value approaching that of pure copper. In the silver alloys, the hardened specimens had a resistivity higher than pure silver (in two cases, with magnesium and beryllium, the resistivity increased on oxidation) and only the non-hardenable zinc alloy decreased to the value of pure silver. This high resistivity of the hardened metals has been assigned to two causes: (*a*) the silver lattice may be distorted, and (*b*) the electric field near a foreign ion or atom may differ from that near a silver ion. On oxidation the distortion effect is increased by the oxide molecules on the lattice. In cases other than magnesium and beryllium cited above, the slight decrease in resistivity is caused by annealing which may in some cases be larger than the effect caused by the oxide particles.

Generally it was found that serious embrittlement is caused in polycrystalline specimens which have been hardened to about 140 Vickers or more. The elongation in tension is often reduced to 1 or 2 per cent. However, since the brittleness was of an intercrystalline nature, single crystals hardened by oxidation remained quite ductile (28 per cent and 21 per cent elongation).

In defence of his failure to recognize the hardening effect, RHINES[9] attributes the cause to the fact that he used a "micro-character scratch hardness tester" while other investigators used micro-indentation measurements. From this he concludes that there is no matrix hardening but an "aggregate hardening effect, which is less likely to be reported by a scratch test." This conclusion is quite at variance with the x-ray measurements made by MEIJERING and DRUYVESTEYN[148] which indicated a lattice expansion of the alloys which had been hardened.

On the basis of unpublished work done by THOMAS,[173] RHINES[9] states that: "It has been supposed, largely upon the basis of studies by FRÖHLICH,[163*] that internal oxidation is not to be expected in alloys in which the minor element is noble with respect to the major element. An exception to this has been reported recently by THOMAS,[173*] who studied the oxidation behaviour of copper-palladium and copper-platinum alloys. Subscales were found in alloys containing as little as

* This number refers to the list of references at the end of the present paper.

5 per cent of the noble metal. The oxide of the subscale was, in these cases, composed of cuprous oxide containing the noble metal in solid solution." From the brief description given by RHINES it does not seem as if this phenomenon is one of internal oxidation in the usual sense of the term. It has been usually accepted that the minor component of a binary alloy must behave according to the requirements set out earlier in this section, i.e. it must oxidize preferentially to the major constituent etc. This is not the case here, where the major component is oxidized. It has not been made clear by RHINES how the noble metals are in any way responsible for this "internal oxidation." This phenomenon may be some secondary effect due to the presence of the noble metals, but they are not directly responsible for the "internal oxidation."

GAS EVOLUTION ON SOLIDIFICATION OF METALS

Although this review is not concerned with gases in molten metals, it may be well to consider briefly the effect which gases, evolved during solidification of metals, have on the resultant solid metals. The evolution may be caused simply by the fact that molten metal at the freezing point will absorb much larger quantities of gas than the same solid metal at the melting point. On solidification of the metal, the excess of gas comes out of solution, forms bubbles, and is trapped as minute cavities or blowholes in the casting. Hydrogen generally appears to be the major offender in causing defects of this type, particularly in iron, copper, nickel, and aluminium. Oxygen in silver causes a similar effect, known as "sprouting."

If the metal contains dissolved oxygen as well as hydrogen, then on solidification of the metal, these gases may react under certain conditions to form steam. This "reaction unsoundness," as it is called, has been studied by ALLEN and his co-workers. In a very detailed study, ALLEN[174] described the porosity due to the simple evolution of hydrogen from copper, and went on to say that when oxygen was present as well as hydrogen the voids were smaller, more numerous, and occurred in the crystal boundaries. The amounts of oxygen and hydrogen required to cause this unsoundness were very small, of the order of 0·002 per cent and 0·0001 per cent by weight respectively. ALLEN concluded also that: "The phenomenon results from a reaction, taking place during the process of freezing, between cuprous oxide and minute residual traces of hydrogen, whereby steam is generated. . . . Cuprous oxide and hydrogen can exist together in molten copper." Sulphur would have the same effect as hydrogen in reducing the oxide of copper. ALLEN suggested that the "reaction unsoundness" phenomenon is more important commercially than simple gas evolution. In a later paper ALLEN and STREET[42] described similar effects in copper-nickel alloys. It was concluded that unsoundness could be caused by an amount of

hydrogen below that needed for pure hydrogen unsoundness provided that the oxygen content was greater than about 0·001 per cent for alloys containing less than 25 per cent nickel, and about 0·01 per cent for alloys of more than 30 per cent nickel. This critical oxygen content is related to the solid solubility of oxygen in the alloys. ALLEN[175] also studied the effect of pressure on the evolution of gases from metals, using the silver-oxygen system as an example.

It was stated in the introduction to the present paper that there would be no discussion on degassing. However, for the reader who may be concerned with this problem, there is a section in the supplementary reference list which gives a few of the many papers which have been written on porosity, unsoundness and degassing. The very brief description of the problem which has been given here is not intended to be a review of the latest work as much as it is an introduction to a commercially important aspect of the gas-metal problem. To the person interested in physical metallurgy there may not be as much interest in this problem as to the foundryman. At any rate further details may be found in the papers referred to in the supplementary list. Since many of the papers discuss both porosity and degassing, no attempt has been made here to indicate which papers deal with any particular study.

EMBRITTLEMENT OF METALS BY GASES

The phenomenon of embrittlement of metals by gases has long been a problem in the metal industry and seems to re-occur as an unsolved problem at various times. Here the mechanism of gas embrittlement will be discussed, and the causes mentioned. The remedies will not be considered here, but may be found in some of the papers which are referred to below.

At the outset, the writer wishes to emphasize a point which was mentioned by MARTIN and PARKER.[176] The term "hydrogen embrittlement," which is correctly used in referring to iron and steel, for example, becomes a misnomer when it refers to the common embrittlement of copper and silver. This latter phenomenon will be referred to here as "water vapour embrittlement" in recognition of the fact that both hydrogen and oxygen are necessary to cause embrittlement in solid copper. Water vapour embrittlement will be discussed first, and true hydrogen embrittlement will be considered later.

In a series of papers, WYMAN,[177] in discussing copper embrittlement, has shown that alternately annealing in oxidizing and reducing atmospheres causes water vapour embrittlement. It was thought that during the oxidizing cycle, oxygen diffuses into the metal and forms cuprous oxide. It will be shown later that compound formation is not at all necessary for embrittlement. During the reducing anneal,

hydrogen combines with the oxygen to form water vapour which ruptures the grain boundaries.

WYMAN'S investigations led to some rather important conclusions concerning commercial annealing practice:

"1. Neither pure copper when oxidized or deoxidized nor tough-pitch copper up to 0·06 per cent oxygen is embrittled by hydrogen at temperatures of 400°C or below, at four hours.

"2. (*a*) The rate of oxygen penetration into 'pure' copper decreases markedly below 800°C.

"(*b*) The rate of hydrogen penetration into oxygen-bearing copper decreases markedly below 700°C. Consequently, up to about 800°C, the coppers containing no oxygen are superior to tough-pitch coppers for resistance to the embrittling action of oxidation-reduction cycles. . . .

"5. The elimination of oxygen by purifying process or by means of deoxidation of the copper is advantageous in increasing the resistance to the embrittling action of reducing gases."

In the final paper of WYMAN'S series it is shown that a steam atmosphere *alone*, even for periods of 8 hours at temperatures of about 1000°C, is not sufficient to embrittle the copper. Any iron in the furnace will reduce the water vapour and make hydrogen available for the embrittlement of the copper, but this action is only important at temperatures above 700°C. It was concluded that (with oxidized copper) hydrogen will cause embrittlement even at 500°C if there is sufficient gas. WYMAN feels that the sources of gas are dirt or grease on the copper, volatile products from enamel, dirt, grease, lubricants, or other materials carried into the furnace extraneously.

RHINES and ANDERSON[178] have definitely established that water vapour embrittlement is caused in copper containing oxygen, either in solid solution, or as precipitated cuprous oxide, when hydrogen is present in the annealing atmosphere. The embrittlement was found to be more severe when cuprous oxide particles were present. RANSLEY[179] had shown that pure dry carbon monoxide will not diffuse into copper, and removes oxygen from the copper without causing embrittlement. This was confirmed by RHINES and ANDERSON who found that copper, annealed in carbon monoxide, was not embrittled on subsequent heating in hydrogen. In this same paper it was shown that hydrogen could reduce foreign metal-oxides, present as impurities in the copper, thus causing embrittlement. The oxides which, when reduced, caused the most severe embrittlement were those of indium, manganese, nickel, phosphorus, tin, and zinc. The oxides of columbium, gallium, lithium, magnesium, and zirconium were the least detrimental of those investigated, and the oxides of aluminium, arsenic, beryllium, bismuth, boron, calcium, cerium, cobalt, chromium, iron, lead, antimony, and titanium had an intermediate effect. Associated with RHINES' studies

of internal oxidation (q.v.), a minor type of embrittlement has been found to occur when the oxides of the subscale were reduced by the hydrogen. Another type of embrittlement, not mentioned above, was also found when copper, containing small amounts of arsenic, antimony, or bismuth, was annealed in hydrogen. This effect was thought to be due to the formation of the volatile hydrides of these metals creating a pressure sufficient to rupture the alloy.

Many separate reports are to be found in the literature on the embrittling action of water vapour which is usually produced within the metal on annealing in a hydrogen-bearing atmosphere. Most of the investigations have been purely empirical and in many cases the results are applicable only to the particular conditions under which the work was done. In many cases rather complex alloy systems have been used, so that there were several variables to be considered. This, of course, often leads to invalid assumptions and erroneous conclusions. The present writer would prefer to review just a few of the outstanding papers, including those discussed above. Although the following remarks apply particularly to copper and silver, it must be recognized that similar phenomena must be expected wherever one is treating a metal which will dissolve both oxygen and hydrogen in suitable quantities.

In 1926, SMITH and HAYWARD[180] presented evidence which was interpreted as showing that a sintering action occurs in copper when heated above 800°C after having been embrittled by a hydrogen treatment. It was observed that cast copper containing less than 0·06 per cent oxygen could be deoxidized in a relatively short time by annealing at a high temperature in hydrogen. However, with copper containing larger amounts of oxygen (0·10 per cent), it was noted that, after annealing at very high temperatures, the copper had much better electrical and mechanical properties than similar copper annealed at only 800°C. From metallographic and other evidence it was suggested that at the lower annealing temperature, the hydrogen diffuses into the copper and reacts with oxide particles forming steam, which blows the metal apart. This initial rupture will allow more rapid penetration of the hydrogen, and the copper oxide eutectic areas rapidly become spongy networks of microscopic cracks. In the case of high-oxygen copper, it is claimed that the stress created by the steam formation strains the metal which, on heating, recrystallizes and during this process the cracks are healed. With low-oxygen copper, the oxide particles form in globules along the grain boundaries. It is suggested that the hydrogen can reduce this oxide with little effect on the copper lattice. Since no strain is present, the copper will not recrystallize or sinter, so the metal remains brittle.

An excellent paper by RANSLEY[179] indicates that the mechanism of

embrittlement suggested above is unlikely. From data available at that time, RANSLEY made the following conclusions:

"(*a*) Compound gases do not diffuse through solid copper. Dissociation of molecules adsorbed at the surface may, however, be followed by penetration of the constituent atoms.

"(*b*) The embrittlement of copper in reducing atmospheres may in all instances be ascribed to penetration by hydrogen. When embrittlement occurs in gases such as carbon monoxide, hydrogen is present as an impurity, or is produced either by dissociation at the surface, or by a side reaction, e.g. $CO + H_2O \rightleftharpoons CO_2 + H_2$.

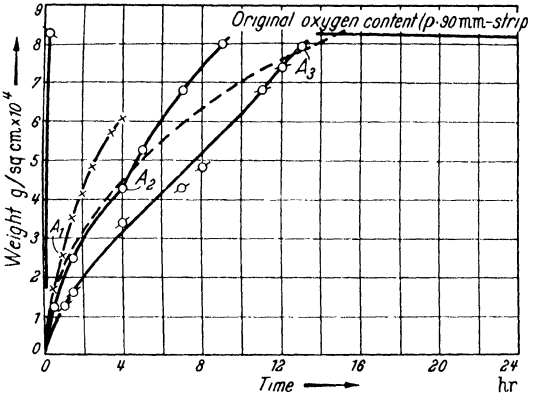


Fig. 21. Removal of oxygen from copper at 850°C in hydrogen.
○ 760 mm pressure × 10·0 mm pressure ○ 5·0 mm pressure
ø 2·0 mm pressure
Broken line indicates theoretical diffusion of oxygen (RANSLEY[179]).

"(*c*) Reaction between gases in solution with the formation of a new phase cannot occur in the lattice. Reaction can only occur at an interface, i.e. in the present case at the copper surface, at an oxide inclusion, or possibly at a grain boundary."

Since cracking does not necessarily originate at oxide particles, and since the effect always seems to be intergranular, it was suggested that the hydrogen and oxygen can exist in equilibrium in the lattice as separate atoms, but they diffuse to the grain boundaries and there react to form water vapour which ruptures the metal.

On the basis of results published by PILLING, RANSLEY has suggested a mechanism which depends also on the diffusion rate of oxygen in the copper, a point neglected by SMITH and HAYWARD. At low hydrogen pressures (less than 26 mm Hg), it was noted that the rate of deoxidation of copper at 700°C in hydrogen depended on the rate of diffusion of oxygen in the copper, and not on the hydrogen pressure. This meant that the deoxidation was a surface process. Further experiments were

performed by RANSLEY, who deoxidized copper at 850°C with hydrogen at various pressures. The time-loss of oxygen curves obtained are shown in Fig. 21. At 760 mm pressure the metal was rapidly deoxidized and became very brittle. At lower pressures there was no significant change in the specimens until the conditions marked A_1, A_2 and A_3 were reached. At these points the specimens suddenly developed blisters. This blistering was attributed to the formation of water vapour beneath the ductile, deoxidized skin of the metal. RANSLEY has pointed out that blistering only took place when the solid deoxidation curves crossed the broken curve, which indicates the normal rate of diffusion of oxygen to the surface at 850°C. These curves indicated to RANSLEY that, at a given hydrogen pressure, as long as the supply of oxygen available at the surface was sufficient to react with the hydrogen, then there would be no embrittlement. As the deoxidized layer becomes progressively thicker, however, there is a progressive decrease in oxygen diffusion, and finally the hydrogen will be able to enter the metal and cause water vapour formation and consequent rupture of the metal. RANSLEY concluded that: "The removal of oxygen from copper in hydrogen will thus be a function of the oxygen content of the copper, the hydrogen pressure, the temperature, and the time of heating." As was stated earlier, RANSLEY'S work has also definitely shown that copper may be completely deoxidized, without deleterious effects on the mechanical properties, by annealing in carbon monoxide free from hydrogen. Carbon monoxide does not diffuse into solid copper.

MARTIN and PARKER[176] found that silver, containing dissolved oxygen, suffered a decline in strength and ductility when heated in hydrogen. This effect is analogous to that in copper noted by RHINES and ANDERSON,[178] who found that water vapour embrittlement could occur without the presence of cuprous oxide particles. MARTIN and PARKER also found that, due to the rapid absorption of oxygen by silver, it is possible to embrittle silver by first annealing it in hydrogen and then in air. It was also found that when the silver specimens were annealed in air, a fine-grained structure was present near the surface but larger grains formed in the body of the metal. This grain-growth inhibition was apparently due to some property of the surface material only. When a layer of the surface metal was machined off the specimen prior to annealing, it was found that the fine grain structure did not form. It was concluded that the original specimen was not homogeneous, but no further tests were made. As was expected, specimens of gold and platinum showed no embrittlement on annealing in hydrogen after heating in air. The embrittlement of silver by heating in hydrogen did not occur below 500°C.

CHASTON[170] studied the inhibition of grain growth mentioned above, and also noted by RHINES and GROBE.[181] In the experimental work on

commercially pure silver (99·98–99·99 per cent silver) CHASTON observed that annealing in an oxygen-free atmosphere at temperatures above 600°C caused rapid growth, while the same treatment in air caused impeded grain growth in the areas to which oxygen had diffused. By heating this silver in air at relatively low temperatures (500°C) where grain growth is slow even *in vacuo*, enough oxygen may diffuse into the metal to prevent any grain growth even at 800°C. On the other hand, it was also observed that grain growth occurred to an equal degree in spectroscopically pure silver when it was annealed at 600°–800°C in air as when annealed in hydrogen. It was found that an addition of 0·01 per cent copper, iron or lead to this silver restored the grain growth inhibition. CHASTON attributed the grain growth inhibition to the particles of the metallic impurity oxides formed by internal oxidation.

The results of the work of MARTIN and PARKER were not in agreement with SMITHELLS[182] who claimed that silver saturated with oxygen could not be embrittled by hydrogen "because oxygen diffuses so rapidly out of silver that it has all escaped before the hydrogen can diffuse in." Actually embrittlement by oxygen diffusion into hydrogen-saturated silver was found to be a minor effect, contrary to the opinion of SMITHELLS, who felt that the oxygen would diffuse in much faster than the hydrogen would diffuse out. However, CHASTON confirmed the possibility of embrittling silver by heating first in oxygen and then in hydrogen, the reaction between the gases occurring at temperatures above about 300°C. By various experiments it was learned that the depth and severity of embrittlement was controlled by the treatment in oxygen, and that the variations in the hydrogen treatment had no effect, provided the metal was above 400°C.

CHASTON was not able to confirm the observation by MARTIN and PARKER that oxygen-free silver, annealed first in hydrogen, could be embrittled by heating in oxygen. CHASTON believes that the effect noted by MARTIN and PARKER was due to small amounts of oxygen retained in the silver prior to the hydrogen anneal. CHASTON was not able to produce embrittlement in silver with less than about 0·01 per cent of total impurities. The exact limiting amount of impurity which will restrain grain growth or promote embrittlement was not determined. Another important observation was that these effects are not affected by the rate at which silver is cooled after annealing in air. CHASTON has claimed that this last observation is an argument against the possibility that dissolved oxygen may be an important factor in the embrittlement reaction. For this, and other reasons, a new mechanism of embrittlement was proposed involving the concept of a subscale formation of impurity metal oxides in the silver. CHASTON suggests that this finely dispersed phase would (*a*) hinder grain growth, (*b*) react

with hydrogen to embrittle the silver, and (*c*) harden the metal (see section on Internal Oxidation). It was considered that such a formation would account for all the observations. One weakness in the idea, suggested by CHASTON himself, is that the oxygen may be removed from the silver by a vacuum treatment below the melting point. One would not expect this to occur if relatively stable oxides had been formed. However, it was claimed that: "It is not inconceivable, however, that in these conditions the very finely divided oxides distributed through silver may be capable of reduction to the metallic state."

Another effect, not strictly embrittlement, has been studied by EBORALL and RANSLEY.[183] When an aluminium-magnesium alloy containing 7 per cent of magnesium, is heated with water vapour in the temperature range between the solidus and the liquidus, the metal blisters and blackens. The blistering is due to the absorption of hydrogen which is released at surfaces in the interior of the metal when the concentration of hydrogen in the metal has reached a certain value. At lower temperatures (400°C) it was thought improbable that the hydrogen alone would cause blistering. However, if severe flaws were present in the metal, they could be enlarged considerably by annealing in a moist atmosphere.

ZAPFFE has considered the above type of problem[184] and has suggested the concept of the "hydrogen potential." A hypothetical system is considered to consist of a closed chamber divided into two compartments by a diaphragm of the metal being studied. Steam, at a constant partial pressure of 1 atm, is admitted to one compartment. If the metal membrane is initially saturated with oxygen, the steam will react with it to form metal oxide and hydrogen. Some of the atomic hydrogen will diffuse through the membrane forming molecular hydrogen in the second compartment, some will remain in solution in the metal, and some will form molecular hydrogen in the steam. The limiting equilibrium hydrogen pressure which is equal in both compartments is called the "hydrogen potential," and is a measure of the hydrogen absorbed by the metal. The case is also considered where the membrane is not initially saturated with oxygen. It is shown from this concept, for example, that nickel has a very low "hydrogen potential" which is related to the fact that nickel does not react with steam. The "potentials" of aluminium and magnesium, which react with steam, are much higher. The measure of hydrogen potential is related to the tendency of the metal to react with water vapour and, therefore, to form blisters.

Several years ago PFEIL,[185] when experimenting with samples of almost pure iron containing crystals of about 1 in. in diameter, found that when such a specimen had been pickled for 30 minutes in 20 per

cent sulphuric acid, the crystals could be broken away from each other at the crystal boundaries. Apparently the metal was so brittle at the boundaries that fracture occurred without any marked deformation in the body of the individual crystals. Since that time, many experiments have been done to study the detrimental effects of hydrogen in iron and steel—particularly embrittlement and hair-line crack formation—caused by prolonged soaking in hydrogen atmospheres, pickling, or electrolytic plating. A survey of the extensive literature has shown that little is known concerning the fundamental factors involved in these problems. Many of the papers are purely descriptive and deal with relatively complex systems which are not suited to a study of the problem. The writer appreciates the importance to industry of knowing how to overcome these effects in certain special cases; however, one feels that it is more important to answer the fundamental, "Why?" Fortunately, in more recent years, there have been a few studies devoted to answering the question.

Some of ANDREW's[186] earlier papers dealt with the effect on mechanical properties and hair-line crack formation of hydrogen introduced into laboratory specimens of steels. Later, SYKES, BURTON and GEGG[187] attempted to interpret this work in terms of industrial practice. The hydrogen contents of a number of steels were determined at different stages of the steel-making process. It was confirmed that a hydrogen content greater than 2 c.c. per 100 g would reduce the steel's ductility. This is significant since it was also found that all the steels produced contained 4–6 c.c. of hydrogen per 100 g. Depending on the thermal history of the specimen, high hydrogen contents may or may not cause hair-line cracks. Although other factors may modify the phenomenon, it was definitely concluded that steel with a high hydrogen content is more liable to crack during heating and cooling than hydrogen-free steel. It was not suggested that hydrogen alone is the cause of crack formation. In fact, in slowly cooled material, free from macro-stresses, the hair-line cracks were attributed to the combined effects of hydrogen, local stresses, and low ductility in a preferred direction. The reader is referred to the paper for complete details of this interesting work, and for the puzzling problems which had to be left unanswered.

Recently ANDREW[188] has published the results of a comprehensive study of the relationship between hydrogen and transformation characteristics in steel. Crack formation and embrittlement were found to be closely associated with each other, and with the general theme of the study, but the amount of hydrogen required for embrittlement is less than that needed for hair-line crack formation. In each case the crack formation is due to "disruptive hydrogen pressure being built up in internal cavities or voids." This pressure would be caused by the diffusion of atomic hydrogen into any imperfection in the lattice which

would hold more than one hydrogen molecule. It was also suggested that the reaction between hydrogen and iron carbide might create a pressure of methane sufficient to rupture the lattice. Finally, ANDREW emphasized that hydrogen diffusivity and solubility are the two main factors determining the behaviour of a given steel in response to hair-line cracks. Unless the effect on these factors of structure and stresses in the steel is known, full understanding of the phenomenon cannot be reached. The results of some work on diffusion have now been published by CHANG and BENNETT,[189] who investigated the effects of chromium, nickel and molybdenum on the rate of diffusion of hydrogen in iron. It is to be hoped that this is the first of a series of experimental studies carefully planned to reveal the fundamental reason for the phenomena.

In another paper, ANDREW and LEE[190] discuss the hydrogen phenomena in steel on the basis of internal pressures caused by the hydrogen. It is suggested that at high temperatures the hydrogen is retained in the lattice in atomic form in solid solution, however, at lower temperatures the hydrogen may be occluded in molecular form at the crystalline boundaries. If the hydrogen is precipitated from solution at high temperatures, plastic yielding will relieve the stresses. At lower temperatures the precipitating hydrogen creates internal stresses in the metal. Once the pockets of molecular hydrogen have been formed, further diffusion is impossible. It is claimed that the concept of molecular hydrogen concentration at the lattice imperfections (crystallite boundaries) must be feasible since it could explain why the composition and thermal history of the steel control the amount of hydrogen present in it. It was shown how the amount of hydrogen retained by a steel specimen increased with the increasing hardness of the specimen produced by rolling. This cold work would increase the volume of crystallite boundaries, and, therefore, the available volume for molecular hydrogen precipitation. Just as a plea was made for more fundamental research on diffusion, so a need was expressed for a clearer understanding of the effects of composition and treatment on the crystalline substructures.

HOBSON and SYKES[191] have confirmed a correlation between the ductility and the hydrogen content of low-alloy steels. Hydrogen was found to have little effect on the shape of the stress-strain curve, but fracture was caused before elongation and reduction in area had reached normal values. As in the other papers mentioned, it is stressed that the hydrogen content alone does not determine whether or not hair-line cracks will form, but it is a major factor.

The author wishes to thank DR. J. N. PRATT, who read the first draft of this chapter and made many helpful suggestions. He also wishes to

acknowledge the assistance of *The National Conference of Canadian Universities* and *The International Nickel Company of Canada, Limited*, who provided a fellowship for a research programme of which this survey is a part.

References

[1] Smithells, C. J. *Gases and Metals* John Wiley, New York: and Chapman and Hall, London 1937

[2] Smith, D. P. *Hydrogen in Metals* University of Chicago Press 1948

[3] Barrer, R. M. *Diffusion In and Through Solids* Cambridge University Press 1941

[4] Eastwood, L. W. *Gas in Light Alloys* John Wiley, New York 1946

[5] Nikuradse, A. and Ulbrich, R. *Das Zweistoffsystem Gas-Metall* Verlag von R. Oldenbourg, Munich, Germany 1950

[6] Fast, J. D. *Philips Technical Review* 6 (1941) 365; 7 (1942) 74; also in *Philips Techn. Rundschau* 6 (1941) 369; 7 (1942) 73

[7] Smithells, C. J. *J. Roy. Soc. Arts* 86 (1938) (4472) 936; (4473) 951; (4474) 971

[8] —— *Metal Treatment* 1 (1935) 165

[9] Rhines, F. N. *Trans. A.S.M.* 43A (1951) 174

[10] Dushman, S. *Scientific Foundations of Vacuum Technique* John Wiley, New York 1950

[11] Kubaschewski, O. *Z. Elektrochem.* 44 (1938) (2) 152

[12] Johns, I. B. *U.S. Atomic Energy Commission Publ.* 1946 (MDDC–290) 1

[13] Seeliger, R. *Naturwiss.* 1942 (30–31) 461

[14] Baukloh, W. and Kayser, Z. *Z. Metallkunde* 27 (1935) 284

[15] Bartholomeyczyk, W., Funk, W. and Seeliger, R. *Z. Physik* 117 (1941) 651

[16] Lumpe, W. and Seeliger, R. *Z. Physik* 121 (1943) (9–10) 546

[17] Russell, A. S. *Metal Progress* 55 (1949) 827

[18] Sieverts, A. and Hagen, H. *Z. phys. Chem.* 169 (1934) 237

[19] Lobley, A. G. and Jepson, D. *J. Inst. Metals* 35 (1926) 213

[20] Steacie, E. W. R. and Johnson, F. M. G. *Proc. Roy. Soc.* 112 (1926) 542

[21] Toole, F. J. and Johnson, F. M. G. *J. phys. Chem.* 37 (1933) 331

[22] Wells, A. F. *Structural Inorganic Chemistry* Oxford University Press 1950

[23] Sieverts, A. *Z. physik. Chem.* 155 (1931) (A) 299

[24] ——, Zapf, G. and Moritz, H. *Z. physik. Chem.* 183 (1938–39) (A) 19

[25] Martin, E. *Arch. Eisenhüttenwesen* 3 (1929–30) 407

[26] Sieverts, A. and Zapf, G. *Z. physik. Chem.* 174 (1935) 359

[27] Smithells, C. J. *Metals Reference Book* Butterworths Scientific Publications, London 1949

[28] de Boer, J. H. and Fast, J. D. *Rec. trav. chim.* 55 (1936) 459

[29] Fast, J. D. *Metallwirtschaft* 17 (1938) (24) 641

[30] Sieverts, A. and Zapf, G. *Z. anorg. Chem.* 229 (1936) (2) 161

[31] —— and Brüning, H. *Arch. Eisenhüttenwesen* 7 (1933) 641

[32] Norton, F. J. and Marshall, A. L. *Trans. Amer. Inst. Min. Met. Engrs.* 156 (1944) 351

[33] Iwasé, K. and Fukusima, M. *Nippon Kinzoku Gakkai-Si* (*Trans. Inst. Metals, Japan*) 1 (1937) (8) 289

[34] Hägg, G. *Z. phys. Chem.* 7 (1930) 339

[35] Smithells, C. J. and Rooksby, H. P. *J. Chem. Soc.* (1927) 1882

[36] Sieverts, A. and Krumbhaar, W. *Ber. Dtsch. chem. Ges.* 43 (1910) 893

[37] Röntgen, P. and Braun, H. *Metallwirtschaft* 11 (1932) 459; 471

[38] Chaudron, G. *Bull. Assoc. Tech. Fonderie* 11 (1937) (1) 4

[39] MOREAU, L. *Bull. Assoc. Tech. Fonderie* 12 (1938) (11) 294
[40] —— and CHAUDRON, G. *Compt. rend.* 224 (1947) (11) 829
[41] HANSON, D., MARRYAT, C. and FORD, G. W. *J. Inst. Metals* 30 (1923) 197
[42] ALLEN, N. P. and STREET, A. C. *J. Inst. Metals* 51 (1933) 233
[43] RHINES, F. N. and MATHEWSON, C. H. *Trans. Amer. Inst. Min. Met. Engrs.* 111 (1934) 337
[44] PHILLIPS, A. and SKINNER, E. N., Jr. *Trans. Amer. Inst. Min. Met. Engrs.* 143 (1941) 301
[45] HERTY, C. H. and GAINES, J. M. *Trans. Amer. Inst. Min. Met. Engrs.* 80 (1928) 142
[46] SEYBOLT, A. U. and MATHEWSON, C. H. *Trans. Amer. Inst. Min. Met. Engrs.* 117 (1935) 156
[47] MERICA, P. D. and WALTENBERG, R. G. *Trans. Amer. Inst. Min. Met. Engrs.* 71 (1925) 715
[48] SIMONS, J. H. *J. Phys. Chem.* 36 (1932) 652
[49] GROSHEIM-KRISKO, K. W., HOFMANN, W. and HANEMANN, H. *Z. Metallk.* 36 (1944) (4) 91
[50] DE BOER, J. H. and FAST, J. D. *Rec. trav. chim.* 59 (1940) 161
[51] GULDNER, W. G. and WOOTEN, L. A. *J. Electrochem. Soc.* 93 (1948) (6) 223
[52] EHRLICH, F. *Z. anorg. Chem.* 24 (1911) 53
[53] GIBB, T. R. P. *J. Electrochem. Soc.* 93 (1948) (5) 198
[54] RUNDLE, R. E. *Acta Cryst.* 1 (1948) (4) 180
[55] EVANS, J. W. *Ph.D. Thesis* Royal College of Science, Imperial College, London 1951
[56] SIMONS, J. H. and HAM, W. R. *J. Chem. Physics* 7 (1939) (10) 899
[57] STEARN, A. E. and EYRING, H. *J. Physical Chem.* 44 (1940) 955
[58] FAST, J. D. *Philips Technical Review* 6 (1941) 365
[59] —— *Philips Technical Review* 7 (1942) 74
[60] RHINES, F. N. *Trans. Amer. Inst. Min. Met. Engrs.* 156 (1944) 335
[61] BARRER, R. M. *Discussions Faraday Soc.* (Physical Chemistry of Process Metallurgy) 1948 (4) 68; discussion 117
[62] LEPP, H. *Vide* 3 (1948) 433; 463
[63] SACHSE, H. *F.I.A.T. Rev. German Sci.* 1939–1946: *Inorganic Chem.*—I (1948) 48
[64] SCHAFER, H. *F.I.A.T. Rev. German Sci.* 1939–1946: *Inorganic Chem.*—V (1948) 107
[65] RANSLEY, C. E. *Met. Abs.* 16 (1949) 858
[66] SMITH, D. P. *Phil. Mag.* 39 (1948) (vii) 477
[67] —— *J. Appl. Phys.* 20 (1949) 1186
[68] —— *Science* 113 (1951) 348
[69] BASTIEN, P. *Metaux et Corrosion* 25 (1950) 248
[70] BHAT, U. V. and LLOYD, H. K. *J. Iron and Steel Inst.* 165 (1950), 382
[71] ISENBERG, I. *Phys. Rev.* 79 (1950) (ii) 736
[72] GULBRANSEN, E. A. and ANDREW, K. F. *J. Electrochem Soc.* 96 (1949) 364
[73] —— —— *Trans. Amer. Inst. Min. Met. Engrs.* 185 (1949) 515
[74] —— —— *Trans. Amer. Inst. Min. Met. Engrs.* 185 (1949) 741
[75] —— —— *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 586
[76] MCQUILLAN, A. D. *Proc. Roy. Soc.* 204 (1950) (A) 309
[77] —— *J. Inst. Metals* 79 (1951) 73
[78] —— *J. Inst. Metals* 79 (1951) 371
[79] —— *J. Inst. Metals* 80 (1952) 363
[80] ZENER, C. *Int. Conf. on the Physics of Metals*, Amsterdam 1949
[81] SNOEK, J. L. *Physica* 8 (1941) 711

[82] YENSEN, C. F. D. in F. Bitter *Introduction to Ferromagnetism* New York 1937 Chap. IV
[83] SNOEK, J. L. *Physica* 6 (1939) 161
[84] RICHTER, G. *Ann. Physik*, Leipzig 26 (1937) 605; 32 (1938) 683
[85] SNOEK, J. L. *Physica* 6 (1939) 591
[86] KÊ, T. S. *Phys. Rev.* 74 (1948) (ii) 9
[87] DIJKSTRA, L. J. *Philips Research Reports* 2 (1947) 357
[88] PETCH, N. J. *J. Iron and Steel Inst.* 147 (1943) 221; LIPSON, H. and PARKER, A. M. B. *ibid.* 149 (1944) 123
[89] KÊ. T. S. *Phys. Rev.* 74 (1948) (ii) 16
[90] —— *Phys. Rev.* 74 (1949) (ii) 1219
[91] —— *Phys. Rev.* 74 (1948) (ii) 914
[92] —— *Trans. Amer. Inst. Min. Met. Engrs.* 176 (1948) 448
[93] WEST, W. A. *Trans. Amer. Inst. Min. Met. Engrs.* 167 (1946) 192
[94] NOWICK, A. S. *Trans. A.S.M.* 44 A, (1952) 248
[95] DIJKSTRA, L. J. *Trans. Amer. Inst. Min. Met. Engrs.* 185 (1949) 252
[96] FAST, J. D. *Iron and Coal Trades Review* 14 April 1950 837
[97] JOSEFSSON, Å. and KULA, E. *Trans. Amer. Inst. Min. Met. Engrs.* 1952 (in *J. Metals* [Feb. 1952] 161)
[98] MARINGER, R. and MUEHLENKAMP, G. T. *Trans. Amer. Inst. Min. Met. Engrs.* 1952 (in *J. Metals* [Feb. 1952] 149)
[99] KÖSTER, W. *Z. Metallk.* 39 (1948) 9
[100] ZENER, C. *Phys. Rev.* 75 (1949) (ii) 345
[101] MARX, J. *Phys. Rev.* 75 (1949) (ii) 1626
[102] KOEHLER, J. S. *Phys. Rev.* 75 (1949) (ii) 1626
[103] RENTSCHLER, H. C. and HENRY, D. E. *J. Opt. Soc. Amer.* 26 (1936) 30
[104] —— —— *Trans. Electrochem. Soc.* 87 (1945) 289
[105] ANDERSON, P. A. *Phys. Rev.* 76 (1949) (ii) 388
[106] HUME-ROTHERY, W. and ANDREWS, K. W. *J. Inst. Metals* 68 (1942) 19
[107] HÜTTIG, G. F. and BRODKORB, F. *Z. anorg. allgem. Chem.* 144 (1925) 341
[108] WOOD, W. A. *Phil. Mag.* 23 (1937) 984
[109] HUME-ROTHERY, W. and WYLLIE, M. R. J. *Proc. Roy. Soc.* (A) 181 (1943) 331
[110] POTTER, E. V. and LUKENS, H. C. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 699
[111] —— and HUBER, R. W. *Phys. Rev.* 68 (1945) (ii) 24
[112] SNAVELY, C. A. *Trans. Electrochem. Soc.* 92 (1947) 537
[113] VOZDVIZHENSKY, G. S. *Zhur. Priklad. Khim.* 20 (1947) 813
[114] —— *Zhur. Priklad. Khim.* 20 (1947) 1255
[115] KRASNIKOV, A. I. *Izvest. Akad. Nauk U.S.S.R.* 1946 (Tekhn) (1) 133
[116] OWEN, E. A. and JONES, J. I. *Proc. Phys. Soc.* 49 (1937) 603
[117] —— —— *Proc. Phys. Soc.* 49 (1937) 587
[118] CLARK, H. T. *Trans. Amer. Inst. Min. Met. Engrs.* 185 (1949) 588
[119] BUSK, R. S. and BOBALEK, E. G. *Trans. Amer. Inst. Min. Met. Engrs.* 171 (1947) 261
[120] PORTEVIN, A., CHAUDRON, G. and MOREAU, L. *Compt. rend.* 204 (1937) 1252
[21] MICHEL, A. *Bull. Assoc. Tech. Fonderie* 12 (1938) 302
[22] CHAUDRON, G., PORTEVIN, A. and MOREAU, L. *Compt. rend.* 207 (1938) 235
[23] MICHEL, A., BÉNARD, J. and CHAUDRON, G. *Bull. Soc. Chim. France*, 12 (1945) (v) 336
[24] —— and Mme. MICHEL *Compt. rend.* 221 (1945) 551
[25] PORTEVIN, A. *Metal Progress* 34 (1938) 694
[26] MOTT, N. F. and JONES, H. *The Theory of the Properties of Metals and Alloys* Oxford University Press, London 1936

[127] SIEVERTS, A. and DANZ, W. *Z. physik. Chem.* 38 (1937) (B) 61
[128] VOGT, E. *Ann. Physik* 14 (1932) 1
[129] FITZWILLIAM, J., KAUFMANN, A. and SQUIRE, C. *J. Chem. Phys.* 9 (1941) 678
[130] MICHEL, A. and GALLISSOT, M. *Compt. rend.* 208 (1939) 434
[131] WALL, T. F. *Engineer* 163 (1937) 397
[132] WHEELER, M. A. *Phys. Rev.* 49 (1936) (2) 642
[133] SEEMAN, A. J. *Rev. Met.* 47 (1950) 485
[134] MAHN, (MLLE.) F. *Rev. Met.* 46 (1949) 363
[135] SMITH, D. P. *Phil. Mag.* 39 (1948) 477
[136] WRIGHT, P. *Proc. Phys. Soc.* (A) 63 (1950) 727
[137] OWEN, E. A. *Phil. Mag.* 35 (1944) 50
[138] WAGNER, C. *Z. phys. Chem.* 193 (1944) (A) 386
[139] SIEVERTS, A. and BRÜNING, H. *Z. phys. Chem.* 163 (1933) (A) 409
[140] —— and DANZ, W. *Z. phys. Chem.* 38 (1937) (B) 46
[141] OWEN, E. A. and WILLIAMS, E. ST. J. *Proc. Phys. Soc.* 56 (1944) 52
[142] ANDREWS, M. R. *J. Amer. Chem. Soc.* 54 (1932) 1845
[143] WEBBER, R. T. *Phys. Rev.* 72 (1947) (ii) 1241
[144] GOLIK, V. R., LAZAREV, B. G. and KHOTKEVICH, V. I. *Zhur. Eksper. i Teoret. Fiziki* 19 (1949) 202
[145] DRAVNIEKS, A. *J. Phys. Colloid Chem.* 55 (1951) 540
[146] JAFFEE, R. I. and CAMPBELL, I. E. *Trans. Amer. Inst. Min. Met. Engrs.* 185 (1949) 646
[147] SMART, J. S. JR. ,SMITH, A. A. JR. and PHILLIPS, A. J. *Trans. Amer. Inst. Min. Met. Engrs.* 143 (1941) 272
[148] MEIJERING, J. L. and DRUYVESTEYN, M. J. *Philips Research Reports* 2 (1947) 81; 260
[149] TAMMANN, G. and DREYER, K. L. *Ann. Phys.* (Leipzig) 16 (1933) 111
[150] LINDE, J. O. *Ann. Phys.* (Leipzig) 10 (1931) 52; 14 (1932) 353; 15 (1932) 219
[151] COEHN, A. *Z. Elektrochem.* 35 (1929) 676; COEHN, A. and SPECHT, W. *Z. Physik* 62 (1930) 1; COEHN, A. and JÜRGENS, H. *Z. Physik* 71 (1931) 179
[152] SEITH, W. and KUBASCHEWSKI, O. *Z. Elektrochem.* 41 (1935) 551
[153] HALL, M. N. A., MARTIN, S. L. H. and REES, A. L. G. *Trans. Faraday Soc.* 41 (1945) 306
[154] JAEGER, F. M., ROSENBOHM, E. and VEENSTRA, W. A. *Proc. K. Akad. Wetensch.* Amsterdam 36 (1933) 291
[155] KELLEY, K. K. *J. Chem. Physics* 8 (1940) 316
[156] FINLAY, W. L. and SNYDER, J. A. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 277
[157] WAIN, H. L. and COTTRELL, A. H. *Proc. Phys. Soc.* (B) 63 (1950) 339
[158] COTTRELL, A. H. and CHURCHMAN, A. T. *J. Iron and Steel Inst.* 162 (1949) 271
[159] TURNER, T. *J. Inst. Metals* 8 (1912) 248
[160] BLAZEY, C. *J. Inst. Metals* 46 (1931) 353
[161] WYMAN, L. L. *Trans. Amer. Inst. Min. Met. Engrs.* 104 (1933) 141; 111 (1934) 205
[162] SMITH, C. S. *Min. and Met.* 11 (1930) 213; 13 (1932) 481; *J. Inst. Metals* 46 (1931) 49
[163] FRÖHLICH, W. *Z. Metallk.* 28 (1936) 368
[164] LEROUX, J. A. and RAUB, E. *Z. anorg. Chem.* 188 (1930) 205
[165] MEIJERING, J. L. *Pittsburgh Internat. Conf. on Surface Reactions,* (Proceedings) (1948) 101
[166] RHINES, F. N. *Trans. Amer. Inst. Min. Met. Engrs.* 137 (1940) 246
[167] —— JOHNSON, W. A. and ANDERSON, W. A. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 205

[168] DARKEN, L. S. *Trans. Amer. Inst. Min. Met. Engrs.* 150 (1942) 157
[169] RHINES, F. N. *Corrosion and Material Protection* 4 (1947) 15
[170] CHASTON, J. C. *J. Inst. Metals* 71 (1945) 23
[171] MEIJERING, J. L. and DRUYVESTEYN, M. J. *Phys. Soc. Rep. Conf. on Strength of Solids* (1948) 140
[172] SMITH, G. C. and DEWHIRST, D. W. *Australian Inst. Metals* Melbourne Branch, Phys. Met. Div. Preprint, 1950: and *Research* 2 (1949) 492
[173] THOMAS, D. E. *Thesis*, Carnegie Inst. of Technology 1950
[174] ALLEN, N. P. *J. Inst. Metals* 43 (1930) 81
[175] —— *J. Inst. Metals* 49 (1932) 317
[176] MARTIN, D. L. and PARKER, E. R. *Trans. Amer. Inst. Min. Met. Engrs.* 152 (1943) 269
[177] WYMAN, L. L. *Trans. Amer. Inst. Min. Met. Engrs.* 104 (1933) 141; 111 (1934) 205; 137 (1940) 291
[178] RHINES, F. N. and ANDERSON, W. A. *Trans. Amer. Inst. Min. Met. Engrs.* 143 (1941) 312
[179] RANSLEY, C. E. *J. Inst. Metals* 65 (1939) 147
[180] SMITH, C. S. and HAYWARD, C. R. *J. Inst. Metals* 36 (1926) 211
[181] RHINES, F. N. and GROBE, A. H. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 318
[182] SMITHELLS, C. J. *Royal Society of Arts Cantor Lectures* 1938
[183] EBORALL, R. and RANSLEY, C. E. *J. Inst. Metals* 71 (1945) 525
[184] ZAPFFE, C. A. *Trans. A.S.M.* 40 (1948) 315
[185] PFEIL, L. B. *Proc. Roy. Soc.* (A) 112 (1926) 182
[186] ANDREW, J. H. and others *J. Iron and Steel Inst.* 1942 (II) 193; 203; (1946) (I) 67
[187] SYKES, C., BURTON, H. H. and GEGG, C. C. *J. Iron and Steel Inst.* 156 (1947) 155
[188] ANDREW, J. H., LEE, H., LLOYD, H. K. and STEPHENSON, N. *J. Iron and Steel Inst.* 156 (1947) 208
[189] CHANG, P. L. and BENNETT, W. D. G. *J. Iron and Steel Inst.* 170 (1952) 205
[190] ANDREW, J. H. and LEE, H. *Institute of Metals*, Symposium on Internal Stresses in Metals and Alloys, (1948) 265; and various parts of the discussion 432
[191] HOBSON, J. D. and SYKES, C. *J. Iron and Steel Inst.* 169 (1951) 209

SUPPLEMENTARY REFERENCE LIST OF PAPERS ON POROSITY, UNSOUNDNESS, AND DEGASSING

ALLEN, N. P. and HEWITT, T. *J. Inst. Metals* 51 (1933) 257
AMES, B. N. and KAHN, N. A. *Trans. Amer. Found. Assoc.* 55 (1947) 558
ARCHBUTT, S. L. *J. Inst. Metals* 33 (1925) 227
BAKER, W. A. *Shrinkage and Gas Effects in the Casting of Non-ferrous Metals and Alloys* B.N.–F.M.R.A. Research Report No. R.R.A. 661, London, (1944)
—— *Found. Trade J.* 82 (1947) 229; 317; 324
—— and CHILD, F. C. *J. Inst. Metals* 70 (1944) 349; discussion 557
BEVER, M. B. *Iron Age* 161 (1948) 90
BRADASCHIA, G. *Bol. Assoc. Brasil. Metals* 4 (1948) 123
BROWN, H. *Light Metal Age* 5 (1947) 16
—— *Foundry* 75 (1947) 88; 199; 203; 206
CHASTON, J. C. *Metal Treatment* 11 (1944–45) 212; 213
CLAUS, W., BRIESEMEISTER, S. and KALÄHNE, E. *Z. Metallk.* 21 (1929) 268
EASTWOOD, L. W. *Gas in Light Alloys* John Wiley, New York; Chapman and Hall, London 1946
—— and KURA, J. G. *Foundry* 75 (1947) (5); (6); (7); (10)

EDWARDS, C. A. *Trans. Amer. Inst. Min. Met. Engrs.* 117 (1935) 13
ERICKSON, J. L. *Light Metal Age* 2 (1944) 19
ESELGROTH, T. W. *Met. Progress* 55 (1949) 817
FREAR, C. L. *Foundry* 76 (1948) 70; 142; 145
GERNER, W. *Neue Giesserei (Tech.-Wiss. Beihefte)* (1950) (2) 57
JOHNSON, H. H., ARNER, L. H. and SCHWARTZ, H. A. *Trans. A.S.M.* 37 (1946) 309
KAUFMANN, A. R. and GORDON, E. *Met. Progress* 52 (1947) 387
LEPP, H. *Met. Ind.* (Lond.) 47 (1935) 315; 341
LEVANTO, K. I. *Metallurgia* 37 (1948) 269
LORIG, C. H. *Trans. Amer. Found. Assoc.* 42 (1934) 152
MILLER, H. J. *Metal Treatment* 1 (1935) 191; 201
MORSE, R. S. *Ind. Eng. Chem.* 30 (1947) 1064
OPIE, W. R. and GRANT, N. J. *Foundry* 78 (1950) 104; 209
PEARSON, T. F. and BAKER, W. A. *J. Inst. Metals* 67 (1941) 231
PELL-WALPOLE, W. T. *J. Inst. Metals* 70 (1944) 127
—— *Found. Trade J.* 78 (1946) 87; 105; 147; 163; 209; 235; 240
—— *Metallurgia* 37 (1948) 119
*PHILLIPS, A. J. *Trans. Amer. Inst. Min. Met. Engrs.* 171 (1947) 17
PRYTHERCH, W. E. *J. Inst. Metals* 43 (1930) 73
RANSLEY, C. E. and NEUFELD, H. *J. Inst. Metals* 74 (1948) 599
ROHN, W. J. P. *J. Inst. Metals* 42 (1929) 203
SAVAGE, R. E. and TAYLOR, H. F. *Trans. Amer. Foundrymen's Soc.* 58 (1950) 393
SEYBOLT, A. U. *Met. Progress* 50 (1946) 1102
SIMS, C. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 1321
STEPHENS, H. A. *Australasian Eng.* 43 (1943) 18; 69; 71
SWAGERMAN, A. and LEUTSCHER, H. *Metalen* (Internat. Gieterij Congres 1949) (1950) 89
VERÖ, J. A. *Acta Techn. Acad. Sci. Hungar.* 1 (1950) 130
WARD, L. A. *Trans. Amer. Found. Assoc.* 42 (1934) 173
WHITTENBERGER, E. J. and RHINES, F. N. *Trans. Amer. Inst. Min. Met. Engrs.* 1952. In *J. Metals* April (1952) 409
COMMITTEE OF THE A.F.A. NON-FERROUS DIVISION *Trans. Amer. Found. Assoc.* 42 (1934) 160
Symposium on Metallurgical Aspects of Non-ferrous Metal Melting and Casting of Ingots for Working, Monograph and Report Series No. 6, The Inst. of Metals, London

In addition to references quoted in the text of this paper, the following literature references are concerned with various aspects of the general subject:

ARMBRUSTER, M. H. *J. Amer. Chem. Soc.* 65 (1943) 1043. (H–Fe, N, Steel)†
DARDEL, Y. *Metal Ind.* 76 (1950) 203 (H–Al)†
—— *Trans. Amer. Inst. Min. Met. Engrs.* 180 (1949) 273; and discussion 185 (1949) 868 (Al–H)†
LANGE, W. On the Absorption of Oxygen by Liquid Metals *Z. Metallk.* 30 (1938) 274
LAURENT, A. *Metaux Corrosion-Usure* 17 (1942) 113
MOREAU, L. *Metaux et Corrosion* 16 (1941) 125 (Ni–Cr–H)†
POTTER, E. V. and LUKENS, H. C. *Trans. Amer. Inst. Min. Met. Engrs.* 171 (1947) 401. (H–Mn)†
RANSLEY, C. E. and NEUFELD, H. *Nature*, Lond. 159 (1947) 709 (Al, H)†
—— —— *J. Inst. Metals* 74 (1948) 599; discussion 781. (Al–H)†

* This is the most comprehensive study of this particular subject which has been published in recent years and is highly recommended.

† Indicates the general subject of the discussion.

RUNDLE, R. E. A New Interpretation of Interstitial Compounds—Metallic Carbides, Nitrides, and Oxides of Composition MX *Acta Cryst.* 1 (1948) 180

SCHMAHL, N. G. Solid (or Molten) Gaseous Equilibria: Oxide Systems *F.I.A.T. Rev. German Sci.* 1939–1946; *Inorganic Chemistry*—V, (1948) 12

TYLECOTE, R. F. Review of Published Information on the Oxidation and Scaling of Copper and Copper-Base Alloys *J. Inst. Metals* 78 (1950–51) 259

4

THE THEORY OF SINTERING

G. A. Geach

POWDERS of most metals will sinter when heated to about three-quarters of their absolute melting points while protected from oxidation or other gaseous attack. Permanent adherence occurs between the particles, the whole mass becomes less porous and its other properties change towards those of the dense metal. A great many studies of these changes have been published, but in almost every case the cold metal powder was mechanically compressed before sintering. This, of course, is the usual procedure when sintering is applied in industrial powder metallurgy. Each such study has taken into consideration only some of the factors which can affect the progress of sintering. Published work is reviewed by KIEFFER and HOTOP[46] and, comprehensively, by GOETZEL.[28] Pressing increases the initial areas of contact between particles and reduces the porosity before sintering; it is usual to reach a far lower porosity when sintering a compressed powder than if a loose powder is used. Only very few studies of the sintering of unpressed metal powders have been published (SCHLECHT, SCHUBARDT and DUFTSCHMID[69]; DELISLE[18]; SHALER and WULFF[73]) but from these and from other related work there is no suggestion that the basic course of sintering is modified by compression.

Most metals are known to sinter, although those of quite low melting points, and a few metals of normal melting points but complicated structure and, perhaps, largely covalent binding, may be exceptions. Thus, germanium does not sinter even at 850°C (90 per cent of its melting point). HÜTTIG and HENNIG[41] have suggested it to be a general rule that all metals sinter appreciably at about three-quarters of their absolute melting points if interference by extraneous factors is avoided. This tends to give too high sintering temperatures for metals of normal or high melting points; if it applies to metals of low melting points cadmium and lead should sinter at 200°C and tin at about 100°C. No sintering at all has been observed in tin. SAUERWALD[68] could find no adherence between zinc particles at 360°C, although WOLF[80] observed coherence when zinc was extruded through a die at 300°C. KIKUCHI[48] compared the hardness of metal powder pressings after heating them to different temperatures and, as would be expected, found the hardnesses of the worked metals to fall after increasing the temperature. In the case of zinc, however, the hardness rose slightly again after a minimum at 400°C. Silver also had a minimum hardness at 700°C, which is a

conventional sintering temperature for this metal. It is probable that with both metals some bonding of the actual particles commenced at the temperatures of minimum hardness.

The extent to which oxide surface films interfere with the sintering of the low melting-point metals is not clear; available data indicates that interference of this origin is at most only partial. Lead, for example, has oxides which are reduced by carbon monoxide at very low temperatures, but lead powder does not sinter perceptibly when heated to within 5°C of its melting point for 7 hours in atmospheres of hydrogen or carbon monoxide, or in vacuum. Comparison of the behaviour on heating of zinc powder prepared by filing sticks of zinc under carbon tetrachloride, and never exposed to air, with that of similar filings which were left in the atmosphere for an hour before heating, showed a distinctly greater coherence in the unoxidized material after 2½ hours at 380°C (95 per cent of its melting point) but the extent of this sintering was not comparable with that occurring in copper and similar metals at 50 per cent of their melting points.

Many inorganic substances such as silica, refractory oxides and silicates, and halides of high melting points also sinter. So, also, do powdered glasses. Sulphur, silicon, carbon and pure organic materials which have low melting points and contain rather large permanent molecules or structural units seem not to sinter. Sintering appears to be a characteristic of substances which, at the sintering temperatures, contain such small structural units as atoms or simple ions.

Sintering may be considered to consist of two stages: one is the welding of particles which are in contact and the other the elimination of porosity. The welding is not remarkable and it is well known that absolutely clean surfaces of many materials adhere when brought together. The shrinkage, however, implies the rapid movement of matter into the pores and the principal requirement of any theory of sintering is that it explains this flow.

At one time it was suggested that attractive forces between particles of a powder developed at some characteristic sintering temperature (SAUERWALD[66]) but there is no evidence that such forces do depend upon temperature. Heating is required to effect the removal of obstructions between particles and to make it possible for the area of interface to increase and the mass to grow less porous. The union of pieces of cold metals occurs in pressure welding which takes place when clean surfaces of such metals as aluminium, cadmium or copper are squeezed together. A true weld is obtained across the interface. The only preparation required is cleaning the surfaces of oxide and grease; pressing serves to bring clean metal surfaces together over the welding area. It is inconsequent to the formation of a weld whether recrystallization occurs across the interface.

The earlier observation by BEILBY[8] that filings of zinc will adhere in the cold to steel and those of silver to a silver surface may also be a demonstration that the attraction across an interface does not require an elevated temperature.

The cohesive force between small particles of quartz and of sodium pyroborate was measured by BRADLEY,[11] who found it to be about 10 dynes between quartz spheres 0·05 cm in radius. This would be

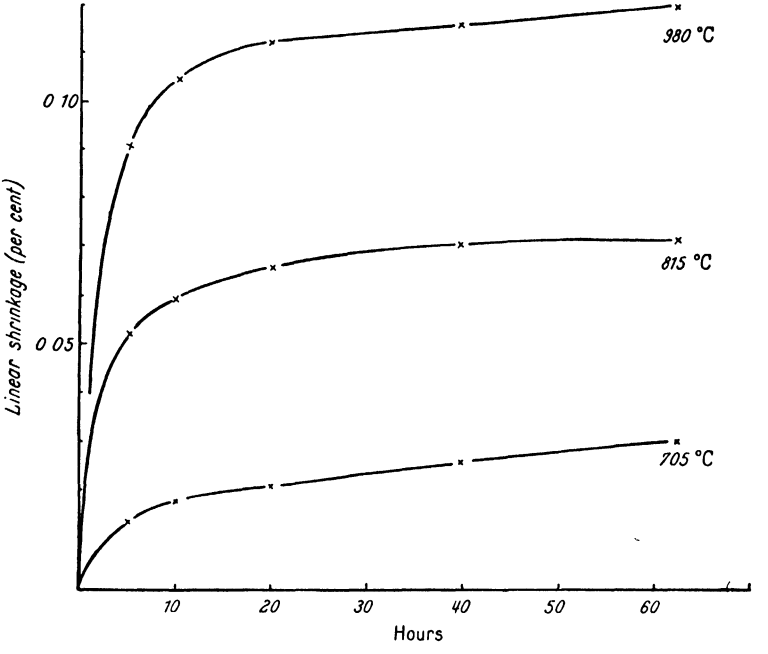


Fig. 1. Changes in length of copper compacts pressed at $1{\cdot}4 \times 10^6$ g/cm^2 and then sintered isothermally in hydrogen[41]

about the order of a tensile strength if the contact was 10^{-4} cm across. BRADLEY found his results very sensitive to the atmosphere about his specimens and so, presumably, to the surface films upon them. The possibility that a direct attractive force between atoms in two particles might act over distances of much more than atomic dimensions has been examined by SHALER.[71] He began with the Sommerfeld model of a metal as positive ions in an electron cloud and, for copper, deduced a force of about 2×10^6 g/cm^2 acting across 120 Å. The effect of temperature would be small. Action across such long distances was questioned by OROWAN,[71–discussion] who pointed out that by such forces a 1 cm diameter copper sphere would support another hanging from it after they had touched.

The properties in which changes during sintering are most simply,

and have been most often, followed are the porosity of the specimen and those such as density or length which are related to it. At any given sintering temperature change is at first rapid and becomes gradually slower.

It is not clear whether, at constant sintering temperature, the porosity of a material tends always to zero as sintering is prolonged or whether it comes to a limiting value. SMITHELLS[75] states that tungsten reaches a fixed density on sintering for about 15 minutes. It appears as though

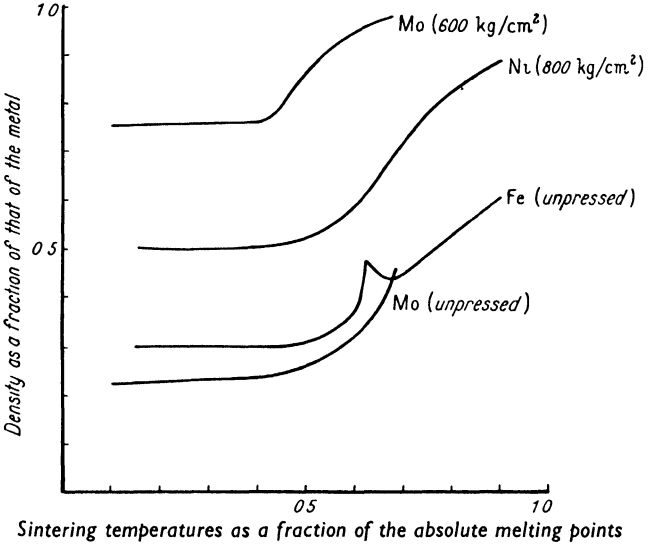


Fig. 2. Densities of different specimens after sintering for fixed periods at different temperatures[30, 52]

crystalline inorganic materials, such as refractory oxides, come to constant porosities, but that glasses have no such limits and tend continuously to zero porosity. The most sensitive measurements on metals show them to become continuously less porous during long experiments; isothermal sintering curves for copper obtained by sensitive dilatometry show this for times up to 60 hours (Fig. 1).

Whether porosity actually tends to a limit or in fact to zero, it does come to a roughly steady value in a few hours at practical sintering temperatures. This value depends very markedly upon that temperature. Much of the published experimental work has reported such apparently limiting densities for different temperatures. Some typical examples of density, as a fraction of that of the compact metal, are plotted against sintering temperature, as a fraction of the absolute melting points, in Fig. 2.

Results for iron are interesting as these show a very marked increase

in porosity at the α-γ transformation temperature. SCHLECHT, SCHUBARDT and DUFTSCHMID[69] sintered unpressed carbonyl iron powder and made metallographic sections after different sintering treatments, as well as measuring the practical limiting densities reached. Increase in pore size after passing the α-γ transformation temperature is clear in their micrographs; there was little rounding of the pores below 1200°C.

DUWEZ[22] has also examined iron and found that the rate of densification at 982°C was only one-thousandth part of the rate at 816°C. Shrinkage at 982°C could not, in fact, be detected after two months.

SHALER[71] made density measurements on specimens of sintered unpressed copper. In 6 hours at 930°C in vacuum the density rose from 50 per cent to 95 per cent that of copper, and the remaining pores were sealed off from the surface. However, in similar experiments with atomized copper powder GEACH and JONES observed quite different rates of sintering, densities after 60 hours sintering at 900°C being 75 per cent that of copper and at 1000°C, 80 per cent. In their specimens no pores were closed off from the surface. Different results from such similar experiments may be largely due to differences in the powders used; in this instance SHALER used copper powder of a less uniform size.

An initial expansion before the shrinkage normal to sintering is often noted. This is usually quite small and most commonly reported during slow sintering at low temperature. Thus it occurs in tungsten, which may increase in length by about 0·2 per cent at 600°C, and DRAPEAU[19] obtained more considerable expansions of this sort (2 per cent in length) in pressed copper specimens heated for 2 to 4 hours at 760°C. The expansion is commonly said, but has not been shown, to be due to gases contained in the metal. That it does not depend, as BALSHIN[7] suggested, on cold work and subsequent recrystallization was shown by SHALER and WULFF,[73] who sintered unpressed copper to a density of 7·1 g/cm^3 by heating for 16 hours at 850°C in argon, and on then heating the same specimen in vacuum for a further 8 hours found the density to fall to 5·3 g/cm^3. In this experiment any possible cold work in the material was annealed out during the first heating.

CLARK and WHITE[14] found that pressings of chromic oxide also expanded during the first part of sintering at 1550°C; the density fell from 2·16 g/cm^3 to 2·10 g/cm^3 in 16 hours, after which it rose again. However, throughout this expansion, the strength of the pressing increased, so that bonding together of particles appeared to go on during the expansion.

Measurements of mechanical properties of sintered powders have practical value, but generally little more; these properties are very sensitive to many factors in addition to sintering. HÜTTIG[39] has examined such properties of sintered metals as the rate of solution in an

acid or the amount of methyl alcohol vapour adsorbed; these properties are all connected with the volume of pores open to the surface of the specimen. KIEFFER and HOTOP[47] have collected and compared many properties of sintered and melted metals.

It is noteworthy that the temperatures at which the changes in different properties of a material are most marked may vary quite widely. This is very clear on comparing electrical resistivity, which is sensitive to the early growth of interparticle welds, with density. Thus, HAUSNER and PINTO[35] examined the sintering of beryllium in vacuum and found that, while the resistivity fell abruptly at about 700°C (170 to 50 microhm-cm), the sharp increase in density took place above 1000°C. This difference in temperatures is quite understandable because it will be the early stages of sintering, when welds between particles develop, that will markedly influence resistivity, but the structural changes then occurring in the general mass will not be great. It is at higher temperatures that pore size is altered rapidly and that great changes occur in density. MYERS[59] studied resistivity changes in tantalum pressed at 7×10^6 g/cm^2 (50 tsi) and sintered for 30 minutes in vacuum. He obtained a marked drop in resistivity at 600°C and smaller ones at 1200° and 2200°C. Density did not change appreciably until this last temperature was reached. MYERS suggested that the earlier stages are due to the removal of gases and impurities, but HAUSNER and DEDRICK[34] supposed the second step due to evaporation of tantalum within pores and its condensation at necks. Resistivity is very dependent upon particle size; thus for copper pressed at $1{\cdot}4 \times 10^6$ g/cm^2 (10 tsi) and sintered $1\frac{1}{2}$ hours in hydrogen, HAUSNER[33] gives the following resistivities:

Resistivities in Microhm-cm

Particle Size	*Sintering Temperatures*		
	600°C	800°C	1000°C
2×10^{-4} cm	5·01	2·42	1·74
44×10^{-4} cm	5·37	3·64	1·97
$44\text{–}74 \times 10^{-4}$ cm	5·6	3·78	2·52

Resistivity values often give a smooth curve when plotted against sintered densities.

Several workers have made metallographic examination of the welds between particles and in general found that no changes are to be detected during early stages of sintering when the compact as a whole gains strength and density. This is in agreement with the fact that crystal growth across the interface is not essential in pressure welding.

LIKHTMAN,[56] however, reports the observation of minute bridges formed between particles of electrolytic copper at temperatures as low as 300°C.

BIER and O'KEEFE[9] found that in specimens of copper reduced from the oxide and pressed at 30 tsi the original particles only ceased to be identifiable on polished sections after sintering at quite high temperatures, for instance after 30 minutes at 900°C. COOK and PUGH[15] examined the changes in the structure of three types of copper powder on sintering. They found that the larger original particles retained their identities after sintering at temperatures between 400° and 800°C, but at 800° to 1000°C the whole specimen acquired a fine-grained equiaxed and porous structure. Recrystallization generally occurred at much lower temperatures in the largest particles.

In electrolytic iron, studied by LIBSCH, VOLTERRA and WULFF,[55] after pressing at $4{\cdot}6 \times 10^6$ g/cm^2 (33 tsi) and sintering for 4 hours in hydrogen, little change could be seen at 500°C (density 6·4 g/cm^3) but at 625°C (density 6·8 g/cm^3) recrystallization in grains and boundaries appeared. At 725°C (density 7·0 g/cm^3) a new structure of uniform equiaxed grains less in size than the original particles had developed, and from 825°C (density 7·2 g/cm^3) to 900°C (density 6·9 g/cm^3) very marked grain growth occurred. Increased porosity near the α-transformation temperature was thought not to be due to the transformation itself, but perhaps to liberation of gases in the metal; the increase appeared to be spaced over a 100°C range about the transformation point. Powders annealed before sintering changed less than similar material that had been strained but not annealed; thus an electrolytic iron powder with particles in a normal strained state increased in density from 3·0 to 3·7 g/cm^3 on heating for 24 hours at 700°C but if the powder used was first annealed no change in density took place under these conditions.

Pores in the iron specimens appeared larger and were more rounded after sintering at 900°C than at lower temperatures. The fact that pores are usually more apparent as sintering proceeds probably arises because flattened porous areas around grains change to more noticeable round shapes. In a pressed powder the pores will tend to be flattened and to extend thinly normal to the pressing direction. As such pores become rounded shrinkage should be more marked normal to than along the pressing direction; this has been observed, for instance, by KAMM, STEINBERG and WULFF[45] for copper. STEINITZ[76] associated changes in the magnetic permeability of pressed iron powders when sintered with the same change in pore shapes. Specimens pressed at $1{\cdot}4 \times 10^6$ to 14×10^6 g/cm^2 (10 to 50 tsi) had permeabilities after sintering for 1 hour at 1150°C which agreed with theory for a mass having the porosity distributed as flat discs. Specimens sintered at 1250° to 1350°C for 24 hours had higher permeabilities than those of the same densities

obtained at lower sintering temperatures, but on re-pressing the permeability fell back to the "flat-pore" value.

Small pores disappear before larger ones; for example RHINES, BIRCHENALL and HUGHES[63] found that the average pore size in copper powders increases on sintering. The number of pores of cross sections of 9×10^{-8} and 12×10^{-6} cm^2 in 0·0293 cm^2 of surface after sintering in hydrogen are given in the table to illustrate these points, and Fig. 3

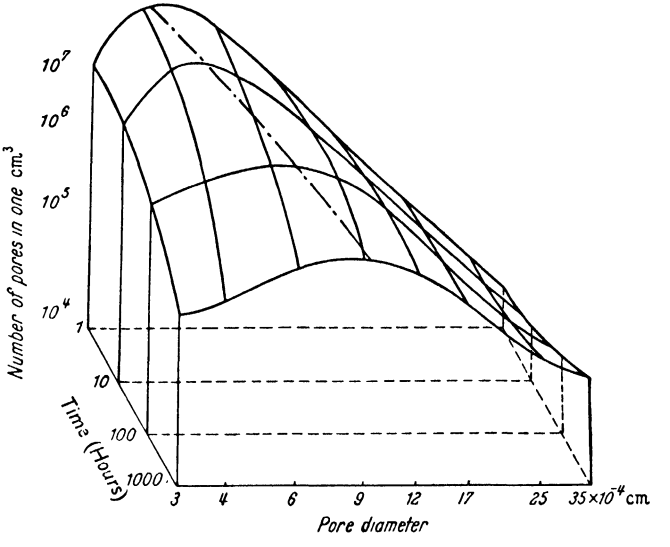


Fig. 3. The population distribution of pores in copper compacts after sintering in hydrogen at 1000°C

shows the general effect. Results are rather more erratic but the general trends are the same when sintering in argon or in vacuum.

Number of Pores in 0·0293 cm^2 of Section

Sintering Temperature °C		*Sintering Time (Hours)*			
		1	10	100	1000
9×10^{-8} cm^2 pores	800	1496	1370	1423	1339
	900	1654	1423	850	757
	1000	1415	1024	392	66
12×10^{-6} cm^2 pores	800	0	1	5	5
	900	16	27	69	65
	1000	65	73	93	113

These workers also reported that regions near the outer surfaces of a specimen appeared to become free from pores before the centre did.

GREENOUGH[29] has made a very interesting observation on the

behaviour of small "blow holes" (about 0·001 cm across) in silver. The specimen was cold rolled sheet in which the holes were randomly distributed. After it had been heated for 19 hours in nitrogen at 900°C almost all the blow holes were collected on the new grain boundaries of the metal. GREENOUGH also noted that the average size of these holes had increased. He suggested that the holes move with boundaries, part of the area of which they eliminate, but it is not at all clear how in this phenomenon material crosses a moving hole. In the same specimens large grooves along boundaries at the surface of the silver and a depression of some grains below the level of others were noted. This implies that holes move along boundaries, intact or as vacant sites; apparently one grain dominates the supply of atoms to move a hole. GREENOUGH questioned the validity of any assumption of random distribution of pores on sintering; he believed that, if sintering involved a plastic flow of matter, a body with pores distributed on planes in this way should increase in density more rapidly than one in which they were randomly distributed.

Experiments in which copper wires from 0·005 to 0·050 cm diameter were wound in close-packed arrangement on bobbins, heated in hydrogen and then sectioned on a diameter for metallographic examination have been made by GEACH and JONES[27] and by ALEXANDER and BALUFFI[2]; these gave pictures resembling an ideal section through contacts between close-packed spherical particles of powder. The former workers found the original boundaries between wires to be surprisingly permanent and to remain generally visible long after the grains within the wires were of the same order of cross section as the wires. The rounding of the "triangular" spaces between wires depended very markedly on wire diameters. Thus at 1000°C the pores between 0·005 cm diameter wires became circular in section in 6 hours, but between 0·013 cm diameter and larger wires they remained clearly triangular. The fifth power of the width of the junctions between wires was roughly proportional to the time of heating and also to the square of the diameter of the wire. Contraction of pore area was slower than that occurring in powders. ALEXANDER and BALUFFI found a cavity to shrink so long as a grain boundary ran into it; when this was no longer the case the cavity would still grow more rounded in shape, but no further change in area occurred. As boundaries disappear more quickly at higher temperatures the pores did not shrink so much, although what change there was took place more quickly. The same authors recorded that the rate of sintering of silver powder is sensitive to the grain size and the consequent concentration of boundaries.

The behaviour of synthetic pores consisting of holes 0·016, 0·022 or 0·026 cm radius and 1 cm long drilled in small bars of copper, when these were heated *in vacuo*, have been examined by POSTLETHWAITE and

SHALER[62]; the specimens stood vertically while heating. The radii fell to minimum values after times of 30 to 70 hours heating and rose again; in some instances there was an initial increase before dropping to the lowest value. Typical values are:

Time (Hours)	0	10	30	40	60	100
	(cm)	(cm)	(cm)	(cm)	(cm)	(cm)
1000°C . .	0·0260	0·262	0·263	0·260	0·254	0·257
1050°C . .	0·0262	0·261	0·259	0·266	—	—

KUCZYNSKI[50, 51] and KUCZYNSKI and ZAVARINE[53] studied sections through the joints between round particles of copper, silver or glass and flat surfaces of the same materials on which they had been heated (Fig. 4). A boundary line at the interface generally remained visible and, in some cases, it appeared to have moved slightly into a copper particle which then stood raised on a slight pedestal above the flat surface. However, as the particles were not strictly round in section it is difficult to be certain of this. The sharp angles at points of contact became rounded as the width of these interfaces grew; KUCZYNSKI measured these widths. These experiments must have been troublesome, as rounded particles had to be found and the sections had to be taken through the centres of the junctions. The glass particles examined were more nearly round than the metals and showed both the rounding of angles and the build up of matter above the original flat surface.

In one of the sections published by KUCZYNSKI the flat surface of the copper block appears to have a shallow "moat" around the interface; this specimen had been heated for 4 hours at 700°C in hydrogen, the copper particle having a radius of about 0·003 cm; however, the available data cannot justify an assumption that this structure is usual if, perhaps, commonly less apparent. KUCZYNSKI[52] obtained most interesting sections through the junctions between a copper particle and a nickel plane or a nickel particle and copper plane (Fig. 5). In each case a very deep groove appears in the copper around the junction. The formation of these grooves would be easy to understand if rapid surface diffusion of the copper atoms might be assumed; the convergence of the groove as it reaches deeper into the copper might also follow, for the mouth of the opening must move out from the original point of contact as the area of the weld increases. It is not apparent why any other mechanism should give vacancies condensing to single large cavities open to the surface of the copper but penetrating deeply into it.

A single layer of 0·03 cm radius copper particles was sintered by DEDRICK and GERDS[17] using temperatures from 700° to 1025°C. The widths of the interfaces between particles were measured and varied

from 0·0036 to 0·0122 cm in experiments with sintering times up to 100 hours, and it was found that the fifth power of the ratio of this diameter to that of the particles gave a straight line which passed through the origin if plotted against sintering time. SHALER[71] has also examined the growth of the junctions between graded copper spheres 0·15 to 0·20 cm in diameter. These were first annealed and then arranged in contact in a single row and heated in hydrogen. SHALER plotted the angle subtended at the centres of particles by the radii of the junctions against sintering times; this gave curves with linear portions for long sintering times but these did not extrapolate through the origin.

The rate of sintering of a material is generally very sensitive to the nature of the atmosphere. Metal powders commonly reach any given density at a lower temperature or, at a given temperature, in less time in hydrogen or other reducing atmosphere than in argon or vacuum. Thus at 980°C copper reached the same density in about 2 hours in hydrogen as after 16 hours in a "vacuum" of 0·5 mm Hg. In the case of beryllium, HAUSNER and PINTO[35] found that after pressing beryllium at $2{\cdot}1 \times 10^6$ g/cm^2 and sintering for 3 hours at 1200°C it had a density of 1·14 g/cm^3 when the sintering was in argon, but of 1·66 g/cm^3 when it was done in air under a pressure of 0·0005 mm Hg; a less marked difference was found when using specimens which had been subjected to higher pressures. A similar sensitivity of refractory oxides, such as beryllium oxide, to the atmosphere in which they are sintered may, it has been suggested, be due to easier diffusion following the production of vacant oxygen sites by surface reaction with the atmosphere. Zinc oxide and titanium oxide depart from stoichiometry in air at sintering temperatures, but alumina does not; the first two sinter at two-thirds of the melting point absolute, the latter at about 0·85 of the melting point.

LIKHTMAN[56] attributed the formation of small interparticle bridges on heating copper at very low temperatures to mobile copper atoms formed as the surface oxide is reduced, or to evaporation and condensation of the metal. He observed that in sealed-off evacuated apparatus in which copper was heated to 300° or 400°C, visible films of copper formed on liquid nitrogen traps.

The rate of sintering is affected by powder particle size. Fine powders usually sinter at rather lower temperatures than coarse ones, although, as the finer powder may give the less dense pressing, it may still be the more porous metal after a given sintering treatment. Uniformly sized powder does not sinter as rapidly as a graded mixture; results for electrolytic iron powders pressed at $4{\cdot}6 \times 10^6$ g/cm^2 (33 tsi) and sintered for 4 hours at 825°C in hydrogen illustrating these points are given by LIBSCH, VOLTERRA and WULFF[55]:

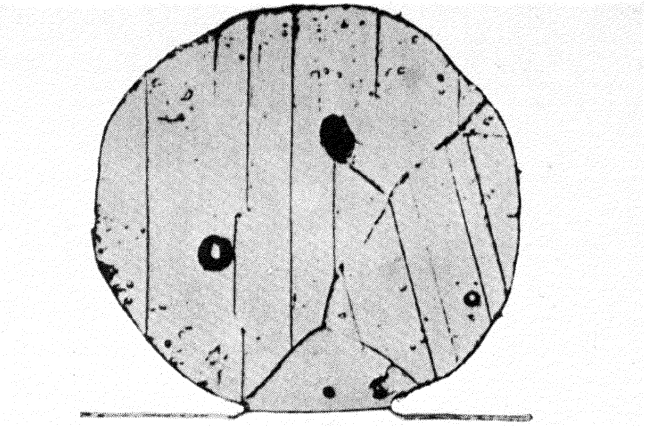


Fig. 4. Copper sphere sintered to a plane copper surface by heating in hydrogen for one hour at 900°C. Magnification × 770 (KUCZYNSKI[51])

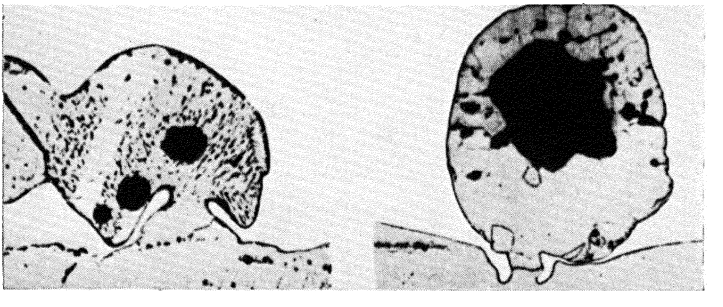


Fig. 5. Sections through: Left—a copper particle sintered to a nickel plane by 24 hours heating at 950°C in hydrogen; magnification × 770. Right—a nickel particle sintered to copper by 48 hours heating at 1020°C in hydrogen; magnification × 220. Black areas are pores, not rounded in nickel (KUCZYNSKI[52])

Powder Size (Sieve Mesh)	*Tensile Strength* (10^6 g/cm^2)	*Density* (g/cm^3)
(*A*) 100–200	2·01	7·20
(*B*) 200–325	1·86	7·06
(*C*) 325–500	2·11	6·90
4 parts (*A*) + 1 part each of (*B*) and (*C*)	2·18	7·27

Gold of such fine particle size as 10^{-6} cm sinters at 150°C as shown by changes in its resistivity and in its infra-red reflectance which, it is shown, are related (HARRIS, JEFFRIES and SIEGEL[32]).

These very low sintering temperatures were found by using particles of gold freshly made by condensation of the vapour; if the material were first "stabilized" by temperatures below 80°C it sintered less readily. It was suggested that the particles are in an unstable state when freshly prepared and that stability follows a diffusion of vacant sites to the surface; an activation energy for this stabilization process of 23,500 $\pm$ 1300 cal/g atom between 50° and 70°C was found.

KALISCHER[44] gave expansion-contraction curves for annealed and unannealed nickel powder compacts when heating in hydrogen at 13·3°C per minute. The unannealed powder began to contract at a lower temperature and shrank more than the annealed powder.

WARREN and LIBSCH[79] examined the sintering of 50/50 iron/cobalt mixtures. A specimen heated rapidly to 1400°C and sintered had more pores and larger pores than a similar specimen first heated between 800° and 1400°C before sintering. DELISLE[18] obtained some indication that loose iron powder sintered less readily if annealed than if strained.

Similarly, pressing a powder causes sintering to occur more readily and leads to a lower porosity after any given sintering treatment. Typical results are those for molybdenum obtained by GRUBE and SCHLECHT.[31] After sintering for 3 hours in hydrogen they obtained the following densities:

Pressure (kg/cm^2)	0	1000	2000	7000
Density (g/cm^3) after 1000°C	2·6	4·6	5·6	7·6
Density (g/cm^3) after 1700°C	6·5	9·4	9·5	10·0

Slightly greater porosity found in some metals after pressing at very high pressure is probably due to expansion by enclosed gas. The effects of pressure on several metals are reported and surveyed by BALKE.[6]

The application of pressure during sintering has a very marked effect. LAURENT and EUDIER[54] measured the widths of the junctions between 0·06 cm diameter copper spheres and plane surfaces after heating in Kuczynski-type experiments. On heating for 100 minutes at 1050°C in

cracked ammonia a junction with a diameter of 0·014 cm was formed; if a load of 10 gm pressed the sphere on to the plane the same heating gave a diameter of 0·036 cm. Loads reaching up to this value are produced by the pressures used in technical hot pressing practice.

Copper was pressed cold under a pressure of $8{\cdot}8 \times 10^5$ g/cm^2 (12,500 psi) to make compacts which were then sintered in hydrogen while under loads of up to $1{\cdot}16 \times 10^4$ g/cm^2 by RHINES and CANNON.[64] No flattening of pores could be discovered after sintering at 1000°C. Density increased more rapidly on loading, and in such a way that the logarithm of the time required to reach any chosen density was related linearly to the stress used. Typical results for a density equal to 95 per cent of the bulk density of copper are:

Time (Hours) .	38·0	27·2	14·5	4·0
Stress (g/cm^2) .	0	700	2100	4900
$\log(t_0/t)$/stress .	—	0·0002	0·0002	0·0002

Here t is the time to reach the density with the specimen under stress and t_0 the time when it is under no load. The function in the third row is constant for a linear relationship.

Any given load produces a constant proportional increase in rate for all times of sintering. A "densification parameter" $\beta = (q_t - q_p)/q_t$, where q_t and q_p are the porosity of specimens given the same sintering treatment unloaded and loaded, respectively, is independent of time for each load. Loading within the range studied does not, then, introduce a fresh mechanism into the processes of sintering. The effect of stress on sintering found in this work is, as RHINES and CANNON point out, similar to that of stress on creep rate.

It is now generally assumed that reduction of surface area and so of surface energy provides the driving energy for sintering. The energy so provided is small; early speculations have considered other possibilities but it seems that these can make only incidental contributions to the cause of sintering. The surface energy of copper powder with particles of 0·005 cm radius would be about 2×10^{-2} cal/g (0·1 joules/g). If copper powder were compacted by hot pressing, a pressure of about 10^4 g/cm^2 (200 lb/in.^{2}) would be required to supply about the same energy to the specimen.

The energy introduced into the powder particles by cold pressing and made available on heating has also been suggested as a factor in sintering. BALSHIN[7] supposed that recrystallization occurs by a re-organization following cold work and that individual atoms diffuse to nuclei which grow at the junctions between particles. The fact that the greater part of sintering usually occurs at times and temperatures beyond those needed for full annealing and that much densification

takes place after this anneal leaves much action by recrystallization improbable.

SAUERWALD,[66] in an early speculation, suggested that at some characteristic temperature which is about three-quarters of the absolute melting point and not affected by pressing, an interchange of atoms occurs between particles and leads to grain growth. Raising the temperature should then aid adhesion by increasing atom mobility and by smoothing away surface irregularities.

SMITH,[74] in another early speculation, suggested that sintering might depend upon changes in structure from amorphous to crystalline or from one crystal form to another, these presumably being supposed to free atoms for individual movement. SMITH'S own determinations of the lowest temperatures at which sintering of unpressed powders of different metals could be detected by prodding with a wire hardly support the view; he detected sintering in iron at 250°C, in copper at 250°C, silver at 180°C and molybdenum at 800°C, but none in manganese or tin.

Surface tension at a pore might induce local flow or surface diffusion, but density changes require flow away from the pores at positions where any direct effects due to surface tension are "diluted" away. Considerable interest now centres on values of the surface tensions of solids and of the energies of grain boundary areas. Recent work has been reviewed by FISHER and DUNN.[24] In one method of determining the surface energy of solids, weights are suspended from foils or wires which are heated, and the weight at which no change in length occurs is determined. This is the force required to balance the effect of surface and grain boundary tensions. UDIN[77] has shown that the grain boundary effect in such experiments is not, as earlier workers supposed, trifling, but that it can be dealt with in wires of "bamboo-like" structure, that is, having grain boundaries normal to their axis and occupying the full cross section. If W is the load giving no change in length, UDIN puts

$$W = \pi R\gamma - N\pi R^2\gamma_B \qquad \ldots \quad (1)$$

where R is the radius of the wire, N the number of grains per unit length and γ_B the grain boundary energy. He believes that $\gamma_B = \gamma/3$ and thus deduces from the experimental results of UDIN, SHALER and WULFF[78] that copper has a surface tension of 1650 ± 100 dynes/cm and a temperature coefficient of $-0{\cdot}55$ dynes/cm/°C. By the same procedure values of 1140 ± 90 dynes/cm for silver[26] and between 875° and 932°C and of 1400 ± 65 dynes/cm for gold[12] over the range 1017° to 1042°C have been given. The atmosphere appears to have a real effect; the above data were obtained in helium; one run in air gave a value for gold of 1207 ± 75 dynes/cm.

By another method BAILEY and WATKINS[5] obtained values for copper.

They prepared grooves along the grain boundaries of copper by thermal etching at 800° to 900°C and measured the angles at the bottom of these grooves between their sides. This angle was measured on specimens which had been heated in argon (ϕ_1), and on others heated in argon and lead vapour (ϕ_2), and the contact angle (ϕ_3) between copper and molten lead was also measured. By balancing surface tension forces one obtains the relation

$$\gamma_{Cu/Cu} = (\gamma_{Cu/Pb} + \gamma_{Pb/vapour} \cdot \cos \phi_3) \frac{\cos (\phi_2/2)}{\cos (\phi_1/2)} \quad . \quad . \quad . \quad . \quad (2)$$

The quantities on the right of this are known or measured in the experiments. Results again show that the atmosphere has an effect:

800°C in hydrogen	1750 dynes/cm
900°C in hydrogen	2000 dynes/cm
900°C in argon	1600 dynes/cm

Results from these two methods are in general agreement. RYMER and BUTLER[65] calculated a quite different value of 500 dynes/cm for gold in order to explain the anomalous surface lattice spacings measured by electron diffraction.

Considerations of surface energy show that a pore should close up if the surface energy of the material is more than half of that of the grain boundary which would be formed; for copper, it appears to be two or three times as great.

Sintering may be considered as a simple rate process. If the logarithms of the times to reach a certain stage in the process, that is for some property to reach some special value, are plotted against the reciprocal of the temperature a straight line is obtained if the Arrhenius equation is applicable. For such a rate process a definite activation energy exists. HÜTTIG[39] measured the electrical resistivity of electrolytic copper powder after various times at different temperatures in hydrogen and obtained fair linear plots of this type, taking the times required to reach selected resistivities. He obtained an activation energy of 56,000 cal/g atom. DUWEZ[21] treated porosity data for copper in the same way and obtained an activation energy of 80,000 cal/g atom for sintering in hydrogen, but for the same copper sintered in vacuum he did not obtain the required straight line plots of log time against $1/T$. This would seem to mean that more than one activation process is involved in the sintering. The activation energy is shown by DUWEZ's curves to fall as the sintering temperature is increased.

FRENKEL[25] suggested that sintering solids flow as Newtonian viscous liquids, that is, so that the strain rate is linearly related to stress, surface tension being the driving force. The viscous flow is supposed to arise from the diffusion of individual vacant atom sites and is thus distinct

from plastic flow which is due to the orderly slipping of atomic layers in a crystal. He compares the sintering of two spheres with the merging of two drops of liquid. Equating the decrease of surface energy released to the energy dissipated as a result of viscous flow, he deduces that the area of contact, πr^2, should increase at the steady rate $\frac{3}{2}\frac{R\gamma}{\eta}t$. FRENKEL also found that gas pressure in a pore will stop its shrinkage when this equals the pressure associated with surface tension pressure; this will be at a minimum for a pore of radius $\sqrt{\frac{PX^3}{2\gamma}}$, X being the original pore radius and P the gas pressure. He showed how an activation energy may be derived as already described, and believes that it is the activation energy of self diffusion. SHALER and WULFF[72] used FRENKEL'S model and calculated an activation energy of 85,000 cal/g for copper; this value is rather higher than those obtained directly for diffusion (60,000 cal/g atom). For the viscosity coefficient of a metal η FRENKEL gave the expression $\eta = RT/Da$, where T is the temperature, a the lattice constant and D the self-diffusion coefficient. Taking published values for D this gave for copper at 1000°C $\eta = 3{\cdot}1 \times 10^3$ poises, a viscosity roughly comparable with Golden Syrup at room temperature!

NABARRO[60] has criticized FRENKEL'S treatment in that the movements of lattice vacancies do not provide a mechanism for viscous flow. Thus if a vacancy in a row of atoms moves from left to right the atoms in that row all move a distance of one vacancy from right to left relative to the rows of atoms above and below, but these two rows remain unmoved relative to each other, which means that there has been no movement of one large part of the crystal relative to the other. FRENKEL'S treatment appears to require such a movement. NABARRO showed that, to be viscous, a crystal must contain mosaic structure or other imperfection in addition to vacant sites and its coefficient of viscosity will be about 10^8 times that given by FRENKEL'S expression; this brings the value for copper at 1000°C to the order of that of pitch at 0°C.

ALEXANDER, KUCZYNSKI and DAWSON[4] considered, from FRENKEL'S point of view, data for the creep of gold wires (0·001 in. diameter) at 970°C. They plotted strain against time for various loads and taking the strain rate from the most linear part of the curves plotted this against stress, obtaining a straight line. This meant that under these limited circumstances flow was viscous, and its coefficient of viscosity came out as $2{\cdot}36 \times 10^{12}$ poises; it also gives a surface tension for gold of 1280 dynes/cm^2. The linear parts of the strain against time curves were given by the data for longer times; the fact that the early parts are steeper and depart from linearity implies that during sintering flow is

not at first purely viscous and that the strain rate is higher than for the apparently viscous conditions.

Using the same data and plotting initial strain rate against stress gave a curve consisting of two straight lines intersecting at what should be a critical shear stress; it has a value of 7 kg/mm^2 (the measured value at room temperature is 9·2 kg/mm^2). The part of the curve below this stress corresponds to viscous flow and a viscosity of $1{\cdot}4 \times 10^{12}$ poises. Above the critical stress initial strain rate is again linear with

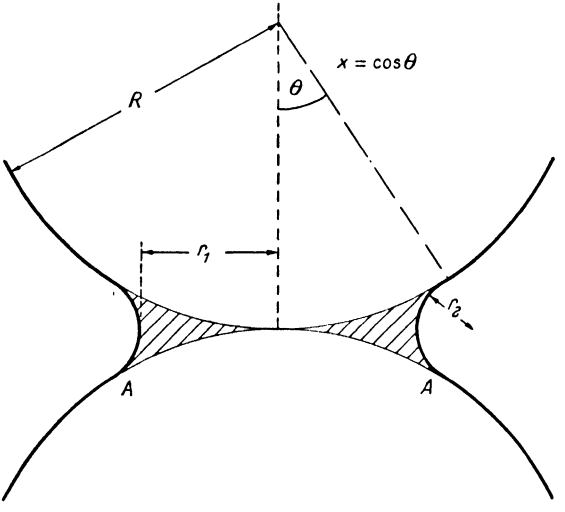


Fig. 6. CLARK and WHITE'S model of the sintering of two spherical particles; the shaded section is through the "lens" forming the junction[14]

stress but it increases more quickly and the coefficient of viscosity is $1{\cdot}5 \times 10^{12}$ poises. It is suggested that at low stresses viscous flow occurs by vacancy migration but for higher stresses a formula obtained by NABARRO[60] and connecting strain rate with the size of a mosaic block, or other region over which the lattice is coherent, gives a strain rate in agreement with the experiment, assuming mosaic blocks to be 10^{-4} cm across.

SHALER'S[71] measurements of the widths of junctions between particles should, when plotted against the sintering time, give a straight line through the origin if FRENKEL'S theory were correct. In fact the curves bend steeply towards the time axis during the first hours of sintering and, although they may become linear over long times (over 20 hours), the linear parts do not extrapolate to the origin. SHALER thought the form of these curves to be evidence for two mechanisms; he suggested an initial rapid flow by creep or slip, with only the slower viscous flow operating at longer times.

CLARK and WHITE[14] have examined the possibility of viscous flow

of a layer of material of thickness h to form a lens between two spherical particles. Their model is shown in Fig. 6. It is assumed that these particles remain spherical and that flow or diffusion maintains this form while matter transfers to the junctions and the volumes of the particles decreases. Clearly this can be true only of the early part of sintering and even then this is not the picture that metallography has given of the changes in metal compacts. CLARK and WHITE made experiments with refractory oxides which perhaps come nearer to their model.

Surface tension (γ) will impose a positive hydrostatic pressure on material in the spherical particles and, because of their concave surface, a negative hydrostatic pressure within a lens. The pressure difference acting from sphere to lens will be

$$P = \frac{\gamma}{R}\left(2 + \frac{x}{1-x} - \frac{1}{\sqrt{1-x^2}}\right) \quad . \ . \ . \ . \quad (3)$$

It is supposed that flow under this pressure is in shear and occurs viscously by activated jumps of moving units. It is assumed that this flow occurs primarily in a thin surface film of the almost cylindrical lens and that most of the pressure drop at the slow rate of flow occurring will be over a short length l on either side of the plane AA (Fig. 6) to which all the moving material must converge.

Flow into one lens of volume v can then be written as

$$\frac{dv}{dt} = 2\int_{r_0}^{r_1} 2\pi r \,.\, u_r \,.\, dr = \frac{\pi P r^4}{\eta l}\left[\frac{4}{3}\left(\frac{h}{r_1}\right)^3\right].$$

Here u_r is the velocity of flow where radius r_1 has the value r and r_0 is the radius of the stationary part of the lens, that is $r_0 = r_1 - h$. From the equation CLARK and WHITE obtain the change of volume shrinkage $\bar{V}$ with time as

$$\frac{d\bar{V}}{dt} = B\left(\frac{1/\rho + q}{1/\rho + V}\right) \cdot \frac{n\sqrt{1-x^2}}{2}\left[2 + \frac{x}{1-x} - \frac{1}{\sqrt{1-x^2}}\right] \quad . \ . \ . \quad (4)$$

where V is the total volume of the lenses in 1 gram of specimen, q_0 the initial pore volume in 1 gram, n the number of contacts with neighbours made by each particle and $B = \gamma h^3/\rho\eta l R^3$. Without varying temperatures the most significant term is η which, for an activated flow process, varies so that $B \propto e^{-E/RT}$, where E is an energy of activation for the flow process.

CLARK and WHITE tested their ideas about a sintering mechanism by obtaining experimental shrinkage-time curves and comparing them with calculated ones.

So far the assumption has been latent in this model that lenses remain distinct from each other. In fact they will merge at some value of x

which will depend upon the packing arrangement of the particles; if each particle had eight neighbours this will be when x is about 0·8. Actually CLARK and WHITE found that curves derived from their equations agreed with experimental data when taken beyond this expected limit to the range of their applicability.

In this way CLARK and WHITE obtained curves giving good fit with experiments made on a glass at temperatures of 598°, 626°, 648° and 676°C. This material had a density ρ, of 2·54 g/cm^3, q_0 the pore volume per gram was 0·21 cm^3, the surface tension γ was about 400 dynes/cm and at 676°C the viscosity was about 10^7 poises; n was assumed to be 8. Experimental results obtained with alumina or with beryllium oxide did not agree with curves calculated on this basis of simple viscous flow. For these materials a Bingham type of flow with a critical shear stress τ_c was considered; this gave rates of strain proportional to the stress due to surface tensions when reduced by subtracting the quantity τ_c. This gave a modified shrinkage curve:

$$\frac{d\bar{V}}{dt} = B\left(\frac{1/\rho + q}{1/\rho + V}\right) \cdot \frac{n\sqrt{1-x^2}}{2}$$

$$\left[2 + \frac{x}{1-x} - \frac{1}{\sqrt{1-x^2}} - \left\{\frac{3R_0 l\tau_c}{2h\gamma}\right\} \frac{1}{\sqrt[3]{1+V\rho}}\right] \frac{1}{(1+V\rho)^2} \cdot \quad (5)$$

the term in curly brackets being a dimensionless number.

Curves were calculated from this which agreed well with experiments made on α- and γ-alumina and a beryllium oxide.

The quantity $B = \gamma h^3/\rho\eta l R^3$ varies with temperature and in this the change in η is the most dominant factor. This varies with temperature according to the Arrhenius activation energy relationship $B \propto e^{-E/RT}$, where E is the activation energy for the flow process. It follows that log B should give a straight line when plotted against $1/T$. For alumina CLARK and WHITE found this true and it also held roughly for magnesia. For glass, however, the relationship obtained gave a curve which implied a rise in the activation energy with falling temperature. The activation energies for viscous flow of glass do vary in this way. Values of the activation energy obtained were 87,000 cal/mole for alumina and, most probably, 43,000 cal/mole for magnesia.

It appears probable from the metallographic work on the "sintering" together of wires, or that of spherical particles to planes, that the particle radii all vary little throughout a large part of the changes during which contact radii grow. This seems generally true until, in a powder, the pores would be sealed off, and this commonly occurs when the porosity is about 15 per cent of the total volume. In deriving their

shrinkage time relationships CLARK and WHITE assumed the particle radii to decrease uniformly. Their further assumption that the thickness, h, of the moving layer remains constant throughout sintering would appear open to question in the cases of solids giving Bingham flow. The assumption would imply that as the lens grows the thickness over which stress exceeds the yield value is not dependent upon the radii of the lens, r_1 and r_2.

MACKENZIE and SHUTTLEWORTH[57] have considered viscous or plastic flow of a mass containing uniformly distributed sealed-off spherical pores. Surface tension on the pores, all of radius X, is considered equivalent to an external pressure of $+ 2\gamma/X$ applied to a solid which is assumed to be incompressible, so that the effect is uniform through any size or shape of sintering material. The effective pressure in copper would be that of an applied load of one atmosphere. An equation of state, that is an equation linking strain rate with stress at constant temperatures, was sought, and again viscous flow and Bingham flow were examined. In this work no atomic mechanism was considered, but by equating energy from reduction of surface area with energy dissipated in flow, equations for shrinkage with sintering time were again developed.

MACKENZIE and SHUTTLEWORTH do, however, argue that flow must arise from a macroscopic mechanism in which atoms move in assemblies and not from any process in which they move as individuals, which, they say, would inevitably be too slow to explain the observed rates. To illustrate this they show that if the excess concentration of vacant lattice sites in copper at 1000°C due to pores of 10^{-4} cm radius in a specimen giving an average distance of 10^{-1} cm from the surface, then shrinkage due to volume diffusion would be at a rate of about 10^{-6}/hour. If sintering were by volume diffusion, however, it should occur from the surface inwards as diffusion of vacant sites to the surface would not occur from any pore while other pores screened the surface from it. While such an effect has been recorded[61] there is not sufficient evidence to show that it is a general phenomenon of sintering; plastic mounting of a specimen for metallographic examination is itself able to develop what appears as this effect.

The first step in this derivation of shrinkage-time relationships is to consider the shrinkage of one pore of radius X surrounded by a spherical shell (radius X_2) of non-porous matter, the whole placed in a mass of uniform matter (that is, the rest of the sintering specimen). The stress in this mass being everywhere a continuous hydrostatic pressure and the overall density of this pore with its shell being that of the whole mass, the radial strain rate is

$$\frac{du}{dX} = \frac{2u_1X_1^2}{X^3}$$

where u is the radial velocity at the surface of the pore. Derivation of this takes from the theory of viscosity the relationship $dG/dt = 3 \cdot (du/dX)^2 \cdot \eta$, where G is the energy. If there were only one isolated pore in a specimen the outer radius of the shell would be infinite. By putting a limiting thickness to the shell, that is giving X_2 a finite value, the interaction of other pores with one under consideration is taken into account. This treatment is strict for small values of X_1/X_2, but it is not clear how accurate it is when this quantity is large, that is when the pores are close together. This means that MACKENZIE and SHUTTLEWORTH'S expressions should hold more rigidly during the later stage of sintering.

A solid with Newtonian viscosity has strain rate proportional to stress and viscosity η independent of strain, so that the radial velocity is $u = \frac{\gamma}{2\eta} \cdot \frac{1}{\rho}$. For an isolated pore $\rho = 1$ and the factor $1/\rho$ represents the effect of all other pores on the rate of sintering of one pore. In a powder compact ρ is generally more than 0·5, so that a pore in a compact will not close up at rates greater than twice that of an isolated pore.

The rate of density change of a Newtonian solid with time during an isothermal sintering run is given in terms of the number of pores, m, in a unit volume, the surface tension and viscosity by:

$$\frac{\gamma m^{1/3}}{\eta}(t - t_0) = \frac{2}{3}\left(\frac{3}{4\pi}\right)^{1/3} \int_0^{\rho} \frac{d\rho}{(1-\rho)^{2/3}\rho^{1/3}} \quad . \quad . \quad . \quad . \quad (6)$$

For a Bingham solid of critical shear stress τ_c the viscosity at any moment is dependent on the strain, and the shrinkage of a pore, $-dX_1/dt$ (radial velocity of pore surface inwards) was found to be

$$-\frac{dX_1}{dt} = \frac{\gamma}{2\eta_\infty \rho}\left[1 - \frac{3\sqrt{2\tau_c X_1}}{2\gamma} \ln \frac{X_2}{X_1}\right]$$

where η_∞ is the ultimate value to which η tends.

With such materials a pore will not close unless $2\gamma/X_1 > \sqrt{2\tau_c} \ln \frac{1}{1-\rho}$; this means that there is a critical size of pore such that material containing only larger pores will not increase in density.

The rate of increase in density is given by the equation:

$$\frac{d\rho}{dt} = \frac{3}{2}\left(\frac{4\pi}{3}\right)^{1/3} \cdot \frac{\gamma m^{1/3}}{\eta_\infty} \cdot (1-\rho)^{2/3} \cdot \rho^{1/3}\left[1 - \alpha\left\{(1/\rho - 1)^{1/3} \ln\left(\frac{1}{1-\rho}\right)\right\}\right] \quad . \quad . \quad . \quad . \quad (7)$$

where $\alpha = \sqrt{2}\left(\frac{3}{4\pi}\right)^{1/3} \frac{\tau_c}{2\gamma m^{1/3}}$.

From this equation values of density can be calculated and plotted for any selected value of α against a reduced time $\gamma m^{1/3}(t - t_0)/\eta_\infty$. It is clear from the equation defining α that this quantity is zero for the case of Newtonian viscous flow, that is when τ_c is zero, and that it increases with τ_c; this quantity α also depends upon the number of pores per unit volume, so that $d\rho/dt$ depends upon this quantity. From equation (7) $d\rho/dt$ can be seen to be positive, that is, sintering with shrinkage t_0 occurs, only when the expression in the square bracket is positive. This is generally true if α is small, but when $\alpha > 0{\cdot}89$ the value of $d\rho/dt$ falls to zero as the square bracket term itself falls to zero, or when $(1/\rho - 1)^{1/3} \ln\left(\frac{1}{1-\rho}\right) = 1/\alpha$. If pressing before sintering had given a density greater than 0·94 of that of the massive material sintering could go on to the theoretical density.

The term $(1/\rho - 1)^{1/3} \ln\left(\frac{1}{1-\rho}\right)$ tends rapidly to zero as ρ tends to unity, which means that when pores are very small or very far apart they will cease to close.

MACKENZIE and SHUTTLEWORTH calculated from strain experiments on copper wires made by UDIN, SHALER and WULFF[78] that the critical shear stress of copper is greater than $2{\cdot}0 \times 10^5$ dynes/cm^2 at 1000°C, and from sintering experiments[73] that it is less than $2{\cdot}6 \times 10^5$ dynes/cm^2 at 850°C. At room temperature the value is 10^7 dynes/cm^2.

KUCZYNSKI[51] examined theoretically what the rate of growth of the contact between a sphere and a plane surface on heating should be in the cases of mechanisms based on viscous or plastic flow, on volume diffusion, on surface diffusion and on evaporation and condensation. He found by his experiments what power of the contact radius was proportional to "sintering" time and compared these results with predictions obtained from his considerations of mechanisms.

For an evaporation and condensation mechanism he takes the pressures of vapour p_1 and p_2 over surfaces of radii R and r_2 as given by the Kelvin equation,

$$\ln \frac{p_2}{p_1} = \frac{-2\gamma a^3}{kT}\left(\frac{1}{r_2} + \frac{1}{R}\right)$$

a being the interatomic distance in the solid. The rate of movement of matter at one temperature should be proportional to the pressure difference $(p_1 - p_2)$, and as R is effectively constant, p_1 does not alter, and so the rate of movement of matter is inversely proportional to r_2, that is the rate of movement $M = k'/r_2$, where k' depends on temperature and R. Expressing r_2 in terms of r_1, the quantity M was shown to be proportional to $1/r$,[2] where M is now the rate of condensation on the

junction of area $\pi r_1^3/R$. Putting $M \propto \frac{R}{\pi r_1^3} \cdot \frac{dv}{dt}$ KUCZYNSKI found that the cube of the junction radius is proportional to the time, $r_1^3 \propto t$.

Volume diffusion was considered in terms of the movement of vacant lattice sites. Near the junction the radius of curvature is small; an increase in volume here would reduce surface energy and could occur if the concentration of vacant sites increased. This concentration depends upon the curvature of nearby surfaces so that there will be a concentration gradient of vacant sites between the junction "lenses" and the rest of the sintering particles and vacant sites will diffuse away from the junctions or atoms will come into them. Kelvin's relationship may be written as $\Delta C = \frac{2\gamma a^3}{r_2 kT} \cdot e^{-E'/RT}$, where C is the concentration of vacant sites and E' is the energy required to produce one vacancy. If D_v is the coefficient of volume diffusion, that of diffusion of holes is $D' = D_v e^{E'/RT}$ and Fick's equation is $\frac{\pi r_1^3}{R} \cdot \frac{\Delta C}{r_2} \cdot D' = \frac{dV}{dt}$; combining these gave

$$\frac{r_1^5}{t} = \frac{40\gamma a^3}{kT} \cdot R^2 D_v \quad \ldots \ldots \quad (8)$$

The case for surface diffusion has been examined by CABRERA[13] and SCHWED[70] as well as by KUCZYNSKI. They have reached different relationships between junction radius and time. The equilibrium concentration of adsorbed atoms or of vacant sites at any point will depend upon the curvature of the surface there. There will thus be different concentrations at the junction and at surfaces of larger radii and this will be such that atoms flow to the junction. The equilibrium concentrations of adsorbed atoms or vacant sites on any surface may be kept up by condensation from vapour or from growth dislocations in the crystal surfaces.

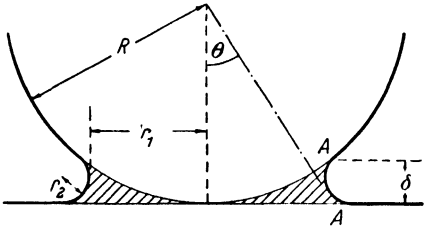


Fig. 7. Section through a spherical particle sintered to a plane surface as in KUCZYNSKI'S experiments[51]

CABRERA considered the surface as of two parts with uniform concentrations of adsorbed atoms over each and a discontinuity at the junction. According to SCHWED, KUCZYNSKI considered only a vacancy concentration gradient in the neck. SCHWED argued that these models are not adequate. Setting up an equation for diffusion past the junction A in Fig. 7, and avoiding these restrictions but assuming the flow to be balanced at any point by interchange of atoms between normal surface positions and sites which steps in the surface enfold

where they bend, he finds that for small values of the radius r_2 the growth law is $r_1^3 \propto t$; where r_2 is larger, the relation is $r_1^5 \propto t$. Here r_2 is to be considered small or large if it is much smaller or greater than a distance $(D_s t_s)^{1/2}/\pi$, where t_s is the relaxation time for condensation of atoms on the surface. It is not easy to estimate this distance with any certainty. Putting $D_s = 10^{-5}$ cm^2/sec and $r_1 = 10^{-4}$ cm would imply a relaxation time of 10^{-2} seconds as standing between the two sintering laws.

KUCZYNSKI originally proposed a law $r_1^7 \propto t$ for surface diffusion and CABRERA, amending the argument, gave $r_1^5 \propto t$, which would not allow distinction between surface and volume diffusion.

These predicted relationships, all of the form $(r_1)^K \propto t$, may be summarized as a table.

Sintering Mechanism	*K*
Viscous or plastic flow	2
Evaporation—condensation	3
Volume diffusion	5
Surface diffusion (SCHWED)	3 *if* r_1 *is small* 5 *if* r_1 *is large*

According to SCHWED, atom or site replacement by equilibrium reaction alone gives a flow which leads to no net junction growth as atoms arrive or leave at equal rates by this means.

KUCZYNSKI examined experimentally the behaviour of copper, silver[51] and glass.[50] He sintered round particles on to flat surfaces and then sectioned these and measured the radii of the junctions. This must be a very difficult experiment. One of these specimens is shown in Fig. 4. For particles of copper (4×10^{-3} cm radius) or of silver (2×10^{-2} cm radius) he obtained values of log (r/R), which gave linear plots against time, t; the slopes varied between 4·6 to 5·4. This corresponds with his deduced value of 5 for a volume diffusion mechanism.

With very much smaller copper particles (5×10^{-4} cm diameter) KUCZYNSKI'S experiments gave a value of K of about 6·5, which lies closer to his prediction for a surface diffusion mechanism, but in the light of CABRERA and SCHWED'S work this result is not easy to explain, and in fact the result for very small copper particles is inconsistent with the theory.

In his experiments with a glass KUCZYNSKI[50] used beads of c. $2{\cdot}5 \times 10^{-2}$ cm radius and obtained excellent straight-line plots of $(r_1/R)^2$ against t. (The data tabulated in his paper contain some misprints.) This is his relationship for a viscous flow mechanism of sintering. KUCZYNSKI'S work is thus a very vivid indication that sintering can be dominated by quite different processes and that the nature of the material is an important factor in determining which are important.

MACKENZIE and SHUTTLEWORTH suggest that in KUCZYNSKI's experiments shear stress will be zero except just at the interface and consequently any plastic movement of the bulk of the material is impossible. The elimination of plastic or viscous flow would make it possible for some other, and perhaps normally subsidiary, mechanism to dominate the process.

There are very few data on the mechanical properties of metals under small stresses at sintering temperatures. ALEXANDER, DAWSON and KLING[3] have obtained creep curves on 0·0028 cm diameter gold wires under small stresses at 920° to 1020°C. At such low stresses the minimum strain (creep) rate was proportional to the stress; thus at 1020°C strain rate was linearly related to the stress up to about 10^{-3} dynes/cm^2, at which stress the strain rate was about 8×10^{-6} per second. This result means that the flow observed under these conditions was viscous; if viscous flow is possible under such minute stresses it is surprising that KUCZYNSKI'S experiments suggested diffusion as the mechanism of sintering under small stresses. The explanation has been put forward[4] that in sintering experiments of this type the flow occurs in the small neck of sharp radius $r_2 = 10^{-4}$ cm, which is about the dimension of a mosaic block and that, as KUCZYNSKI'S experiments showed, this flow is by volume diffusion. In an aggregate of mosaic blocks, such as the gold wires, the creep experiment suggests that volume diffusion of vacancies within the blocks could lead to viscous flow of the specimens. Thus the same mechanism could appear as a diffusion process in sintering and as a viscous flow of the wires.

MURRAY and WILLIAMS[58] have considered the application of the MACKENZIE and SHUTTLEWORTH equation, (7), to hot pressing, that is when a pressure is applied at sintering temperature to a mass containing closed pores. This pressure would require a term to be added to the pressure due to surface tensions. They believe that for hot pressed metals $(d\rho/dt)\rho\ T$ is proportional to L^U, where L is pressure and U has a value normally between 3 and 2, but which falls as T rises. This power relationship could arise if a creep process were involved in sintering, creep rate being proportional to a power of the stress causing it. Pressure is not known to influence diffusion constants and so should not affect sintering due to diffusion. However, pressure alone will compact a metal and its effect would be added to that of any other process.

Diffusion appears to be much faster along grain boundaries than through grains, and as there is no indication that traces of impurity affect bulk diffusion markedly, the difference cannot be due to congregation at boundaries of foreign atoms. HOFFMANN and TURNBULL[38] diffused radioactive silver into the normal metal and by then taking auto-radiographs showed the active material to be concentrated at grain boundaries; its concentration fell away with depth in the silver.

In a rather similar study of diffusion of active silver into copper ACHTER and SMOLUCHOWSKI[1] found the rate of penetration to increase as the difference of orientation between the grains on either side of the boundary increased. Values of lattice (D_v), boundary (D_b) and surface (D_s) diffusion coefficients for the diffusion of silver and thorium into tungsten are collected into the table:

$$D = D_0(^{-E/RT})$$

	D_v	D_b	D_s
Silver into tungsten—			
D_0 (cm²/g) . .	0·895	0·025	0·16
E (cal/g atom) . .	45,950	20,200	10,300
Thorium into tungsten—			
D_0 (cm²/g) . .	1·0	0·74	0·47
E (cal/g atom) . .	120,000	90,000	66,400

The width of a band of silver through which silver can pass by volume diffusion at the same rate as it does by boundary diffusion along a boundary is:

Temperature .	200	400	600	800	960°C
Width . .	45	5×10^{-2}	$1{\cdot}3 \times 10^{-3}$	$1{\cdot}3 \times 10^{-4}$	$3{\cdot}6 \times 10^{-5}$ cm

This width must be the minimum grain diameter if bulk (lattice) diffusion is to dominate the diffusion transport of material.

GREENOUGH[30] studied the extension of silver wires under small stresses. His measured strain rate of $-7{\cdot}61 \times 10^{-9}$ per second at 920°C agrees with that of $-2{\cdot}66 \times 10^{-9}$ per second calculated by HERRING. This calculation assumes that mosaic boundaries cannot function as sinks for diffusion currents and that diffusional viscosity can only account for creep at very high temperatures and low stresses. HERRING supposes that these conditions apply in sintering and that grains may slide over each other tangentially as they grow normally to their boundaries.

In single crystal specimens of silver GREENOUGH[30] found no measurable strain, which indicates that plastic flow within grains makes no contribution to creep of the wires. His results for crystals 0·1 cm diameter tested for $282\frac{1}{2}$ hours at 910°C were:

Specimen	*Stress* dynes/cm²	*Strain*
Single Crystal 1. .	131,000	+ 0·005
Single Crystal 2. .	114,000	+ 0·001
Polycrystalline . .	126,000	+ 0·058

HERRING[36] has examined the effect of the linear dimensions of a sintering system for different sintering mechanisms. Thus, if two compacts contain particles, in one case of radius R_1, in the other of R_2 where $R_2 = \lambda R_1$ and if after intervals Δt_1 and Δt_2 the contacts have radii $(r_1)_1$ and $(r_1)_2$ where $(r_1)_2$ equals $\lambda(r_1)$, HERRING sought to relate the times Δt_1, Δt_2 with the mechanisms of sintering.

When considering a viscous flow process in a crystalline material it is not correct to take the expression $P = 2\gamma/R$ as the pressure inside a particle of radius R, for the stress will vary over the surface and shearing stresses will occur in the interior. For viscous flow HERRING finds $\Delta t_2 = \lambda \Delta t_1$. For evaporation and condensation when sintering particles of radius 10^{-5} cm or more $\Delta t_2 = \lambda^2 \Delta t_1$. Volume diffusion would give $\Delta t_2 = \lambda^3 \Delta t_1$. With surface diffusion $\Delta t_2 = \lambda^4 \Delta t_1$, if it is assumed that rate of migration is proportional to the gradient of a chemical potential. Experiments have not yet been made on this line. Great care will be required to avoid surface contamination of a specimen and to see that one crystal orientation does not dominate one sample.

To explain the increase in average pore size of pressed copper specimens which they observed on sintering RHINES, BIRCHENALL and HUGHES[63] consider volume diffusion. They found that the concentration of small pores decreased sharply with sintering time (Fig. 3) indicating that the shrinkage of small pores is very high; on the other hand, the largest pores, those of volume more than 10^{-9} cm^3, increased in volume. There is nothing in any slip mechanism of sintering to explain the growth of pores. The authors suppose it to be due to volume diffusion, vacant lattice site concentrations being higher in the vicinity of small pores than of larger ones, so setting up concentration gradients and diffusion. This implies that rapid maintenance of equilibrium concentrations occurs by the development or condensation of vacancies. Diffusion of this type between pores and the external surface of a specimen is suggested as the cause of overall shrinkage.

By assuming that the thermal energy of a crystal may be assigned to sites according to the Boltzmann distribution one may calculate the fraction, f, of sites which at any temperature have more energy than that required to form a vacant site. This energy was taken as 39·700 cal/g atom. This calculation gives the following values of f:

Temperature (°C)	700	800	900	1000
f	$1{\cdot}2 \times 10^{-9}$	$8{\cdot}0 \times 10^{-9}$	$3{\cdot}9 \times 10^{-8}$	$1{\cdot}5 \times 10^{-7}$

The Kelvin equation is taken as giving concentrations of vacant sites near curved surfaces. Using these values, RHINES, BIRCHENALL and HUGHES estimate the size of pore which will neither grow nor shrink in a rather idealized model. In this each pore is supposed to have nine

neighbours; it gains vacant sites from smaller pores and loses them to larger ones or to the free surfaces. The proportion of sites leaving one size of pore which reaches pores of any other size depends upon their area of cross section.

The results agree quite well with observation for long sintering times at high temperatures. It is suggested that quite large pores will never be removed by heat alone. Another effect forecast is that a non-wetted inclusion above a minimum size ought to provide a surface for vacant sites to discharge at, and so come to occupy a pore larger than itself. The shape and size of a specimen should also affect sintering rate; it seems improbable that this is so.

RHINES, BIRCHENALL and HUGHES think that evaporation and condensation could provide a rapid enough change of pore shape; at 1000°C they estimate a rate of condensation at sharp junction angles of about 10^{-11} g/sec; in a typical case of sintering about 10^{-9} g of material would flow into the junction during the rounding of pores.

It becomes gradually more clear that under properly controlled circumstances the mechanism of sintering may be dominated by one or other of different processes, particularly by viscous flow or by diffusion. There is still uncertainty about what factors are most important in selecting a mechanism, and so this cannot yet be set up from first principles and conditions for sintering then forecast. Moreover, many simple points of the course of sintering are yet unsettled; whether, for instance, pores do in general become spherical rapidly or only do so much more slowly than the shrinkage changes. The general course of sintering is still uncertain; it is known that much porosity is eliminated before pores are isolated, that small pores are removed before larger ones which may even grow; that the rounding of pores to spherical shape is of secondary importance; that grain boundaries running into pores may be important in their shrinkage; that conditions which favour vacant site formation make sintering quicker but that transformation temperatures are themselves of small importance; it is not clear whether porosity can be wholly removed by sintering alone. It would be valuable to have such points decided more certainly.

The author is indebted to his colleague Mr. F. O. JONES for frequent valuable discussions in connection with sintering; he wishes to thank Dr. T. E. ALLIBONE, F.R.S., for permission to publish this article.

NOTATION

A Point of junction between the original surface of a sintering particle and of the surface of material which has moved during sintering (Figs. 6 and 7).

a Interatomic distance.

B Proportionality constant in equation (4).

C Concentration of vacant lattice sites.
D_v, D_s and D_b Diffusion coefficients for volume, surface and grain-boundary diffusions.
D_o Value of a diffusion coefficient at 0°Abs.
D_H Diffusion coefficient for lattice holes.
E Activation energy.
E' Energy to produce a vacant lattice site.
f Fraction of lattice sites which are vacant.
G Energy.
h Thickness of a flowing layer of matter.
K Power of the junction lens radius, r_1, which should be proportional to sintering time, t, for any given mechanism.
k Boltzman constant.
L Compacting pressure.
l Short length at the junction between particles over which the drop in internal pressure occurs.
M Rate of movement of matter.
m Number of pores per unit volume.
N Number of grains per unit length of a wire in which each grain fills the whole cross section of the wire.
n Number of contacts made by a particle with its neighbours.
P Pressure inside sintering particles.
P Gas pressure.
p_1, p_2 Vapour pressures over surfaces of different radii.
q Total pore volume per gramme of material.
R Particle or wire radius.
R Gas constant.
r Radius.
r_1 Lens radius (Figs. 6 and 7).
r_2 Radius of the surface of a lens (Figs. 6 and 7).
T Absolute temperature.
t Time.
t_s Relaxation time for condensation of atoms on a surface.
U Index for hot-pressing relationship.
u Radial velocity at the surface of a pore.
V Lens volume in one gram of material.
$\bar{V}$ Shrinkage in volume when one gram of material is sintered.
v Volume of a single Clark and White 'lens.'
W Weight under which a loaded wire neither extends nor contracts when heated.
w Weight.
X Radius of a pore.
x $\cos\theta$.
α The expression $\sqrt{2}\left(\frac{3}{4\pi}\right)^{1/3} \cdot \frac{\tau_c}{2\gamma m^{1/3}}$ in equation (7).
β Densification parameter.
γ Surface energy.
γ_B Grain boundary energy.
δ Height of a lens (Fig. 7).
η Coefficient of viscosity.
η_∞ Ultimate coefficient of viscosity of a Bingham solid.
θ Half the angle subtended at the centre of a particle by the junction with another particle.

λ Herring's scale factor.
ρ Density.
τ_c Critical shear stress of a Bingham solid.
ϕ_1, ϕ_2, ϕ_3 Angles.

References

[1] Achiter, M. R. and Smoluchowski, R. *J. Appl. Phys.* 22 (1951) 1260
[2] Alexander, B. H. and Baluffi, R. *J. Metals* 188 (1950) 1219
[3] —— Dawson, M. H. and Kling, H. P. *J. Appl. Phys.* 22 (1951) 439
[4] —— Kuczynski, G. C. and Dawson, M. H. *The Physics of Powder Metallurgy*, Ed. by Kingston, W. E. 1951, Ch. 11
[5] Bailey, G. L. J. and Watkins, H. C. *Proc. Phys. Soc.* (B) 63 (1950) 350
[6] Balke, C. W. *Symposium on Powder Metallurgy, ASTM*, (1943) 11
[7] Balshin, M. Y. *Vestnik Metalloprom* 16 (1936) 91
[8] Beilby, G. *Aggregation and Flow of Solids* London (1921) pp. 69–78
[9] Bier, C. J. and O'Keefe, J. F. *Metals Tech.* 11 (1944) 5
[10] Birchenall, C. E. and Mehl, R. F. *J. Metals* 188 (1950) 144
[11] Bradley, R. S. *Phil. Mag.* 13 (1932) 853
[12] Buttner, F. H., Udin, H. and Wulff, J. *J. Metals* 3 (1951) 1209
[13] Cabrera, N. *J. Metals* 2 (1950) 667
[14] Clark, P. and White, J. *Trans. Brit. Cer. Soc.* 49 (1950) 305
[15] Cook, N and Pugh, S. F. "Symposium on Powder Metallurgy," *I.S.I. Special Report* No. 38 (1947) 162
[16] Dedrick, J. H. *The Physics of Powder Metallurgy*, Ed. by Kingston, W. E. New York (1951) p. 334
[17] —— and Gerds, A. *J. Appl. Phys.* 20 (1949) 1042
[18] Delisle, L. *Trans. Electrochem. Soc.* 85 (1944) 135
[19] Drapeau, P. "Powder Metallurgy," *A.S.M.* 1942, Ed. by Wulff, J. Ch. 31.
[20] Dushman, S. Dunbar, L. W. and Huthsteiner, H. *J. Appl. Phys.* 144, **15**, 108
[21] Duwez, P. "Atom Movements," *A.S.M.* (1950) 192
[22] —— *The Physics of Powder Metallurgy*, Ed. by Kingston, W. E. (1951) 107
[23] —— and Martens, H. *Metal Trans.* 185 (1949) 571
[24] Fisher, J. C. and Dunn, C. G. *G.E. Report* RL–463
[25] Frenkel, J. *J. Physics* 9 (1945) 385
[26] Funk, E. R., Udin, H. and Wulff, J. *J. Metals* 3 (1951) 1206
[27] Geach, G. A. and Jones, F. O. *Research* 2 (1949) 493
[28] Goetzel, C. G. *Treatise on Powder Metallurgy*, Interscience Publishers Inc., 1949 and 1950
[29] Greenough, A. P. *Nature* 166 (1950) 904
[30] —— *R.A.E., Technical Note No. MET.* 151, October 1951
[31] Grube, G and Schlecht, H *Z. Electrochem* 44 (1938) 367
[32] Harris, L., Jeffries, D. and Siegel, B. M. *J. Chem. Phys.* 18 (1950) 261
[33] Hausner, H. H. *Powder Metallurgy Bull.* 3 (1948) 4
[34] —— and Dedrick, J. H. *The Physics of Powder Metallurgy*, Ed. by Kingston, W. E. (1951) 334
[35] —— and Pinto, N. P. *A.S.M. Preprint* 38 1950
[36] Herring, C. *The Physics of Powder Metallurgy* Ed. by Kingston, W. E. (1951) Ch. 8
[37] —— Private communication
[38] Hoffmann, R. E. and Turnbull, D. *J. Appl. Phys.* 22 (1951) 634
[39] Hüttig, G. F. *Kolloid Z.* 98 (1942) 6

[40] HÜTTIG, G. F. *Z. Elektrochem* 54 (1950) 89
[41] —— and HENNIG, W. *Z. anorg. allg. Chem.*, 251 (1943) 260
[42] JOHNSON, W. A. *Trans Amer. Inst. Met. Engrs.* 145 (1941) 107
[43] JORDAN, C. B. and LUWEZ, P. *J. Metals* 1 (1949) 96
[44] KALISCHER, P. R. *Symposium on Powder Metallurgy, ASTM* (1943) 31
[45] KAMM, R., STEINBERG, M. A. and WULFF, J. *Trans Inst. Met. Min. Engrs.* 180 (1949) 694
[46] KIEFFER, R. and HOTOP, W. *Pulvermetallurgie and Sinterwerkstoffe, Springer-Verlag* (1943)
[47] —— —— *Kolloid Z.* 104 (1943) 208
[48] KIKUCHI, R. *Sci. Rep. Tojoku Imp. Univ. First Ser.* 26 (1937) 125
[49] KOPELMAN, B. *The Physics of Powder Metallurgy* New York (1951) 308
[50] KUCZYNSKI, G. C. *J. Appl. Phys*, 20 (1949) 1160
[51] —— *J. Metals* 187 (1949) 169
[52] —— *J. Metals* 185 (1949) 813
[53] —— and ZAVARINE, I. N. *Mikroskopie* 4 (1949) 193
[54] LAURENT, P. and EUDIER, M *Rev. de Met.* 48 (1951) 271
[55] LIBSCH, J., VOLTERRA, R. and WULFF, J. *Powder Metallurgy, A.S.M.* 1942 Ed. by WULFF, J. Ch. 35
[56] LIKHTMAN, V. I. *Dok. Acad. Nauk., U.S.S.R.* 71 (1950) 323
[57] MACKENZIE, J. K. and SHUTTLEWORTH, R. *Proc. Phys. Soc.* (B) 62 (1949) 833
[58] MURRAY, P. and WILLIAMS, J. Private communication
[59] MYERS, R. H. *Metallurgia* 38 (1948) 307
[60] NABARRO, F. R. N. "Report on a Conference on the Strength of Solids," *Physical Society* (1948) 75
[61] NICKERSON, R. A. and PARKER, E. R. *Trans. A.S.M.* 42 (1950) 376
[62] POSTLETHWAITE, A. W. and SHALER, A. J. *The Physics of Powder Metallurgy* New York (1951)Ch. 10 189
[63] RHINES, F. N. BIRCHENALL, C. E. and HUGHES, L. A. *J. Metals* 188 (1950) 378
[64] —— and CANNON, H. S. *J. Metals* 3 (1951) 529
[65] RYMER, T. R. and BUTLER, C. C. *Proc. Phys. Soc.* 59 (1947) 541
[66] SAUERWALD, F. *Z. anorg. allg. Chem* 132 (1922) 277
[67] —— *Metal. u. Erz.* 21 (1924) 117
[68] —— *Lehrbuch der Metallkunde* Berlin (1929) 21
[69] SCHLECHT, L., SCHUBARDT, W. and DUFTSCHMID, F. *Z. Elektrochem* 37 (1931) 485
[70] SCHWED, P. *J. Metals* 3 (1951) 245
[71] SHALER, A. J. *J. Metals* 185 (1949) 796
[72] —— and WULFF, J. *Phys. Rev.* 72 (1947) 79
[73] —— —— *Ind. Eng. Chem* 40 (1948) 838
[74] SMITH, R. C. *J.C.S.* 119 (1921) 2088
[75] SMITHELLS, C. J. *Tungsten* 3rd Ed. London (1952) 124
[76] STEINITZ, R. *Trans. Inst. Met. Min. Engrs.* 180 (1949) 694
[77] UDIN, H. *J. Metals* 3 (1951) 63
[78] —— SHALER, A. J. and WULFF, J. *J. Metals* 185 (1949) 186
[79] WARREN, D. and LIBSCH, J. F. *J. Metals* 3 (1951) 776
[80] WOLF, W. *Arch. Metallkunde* 1 (1947) 361

5

THEORY OF DISLOCATIONS

A. H. Cottrell

Introduction

THE fact that a second review of progress has become necessary, so soon after the first[1] (which will be referred to as T.D.1), shows that the theory of dislocations is being opened up rapidly. Some of this progress could have been anticipated from the ideas that were taking shape at the time the first review was being written but an important contribution has come from newer ideas, some of which have caused a radical change in outlook in certain parts of the theory. Judging from present-day ferment on topics such as the growth of crystals, work hardening, and the nature of slip bands, this pace of progress will be held for at least the next five years.

When the first review was written the broad position was as follows. Arguments about the impossibility of the simultaneous movement of all the atoms in a slip plane and about the low yield strengths of crystals had convinced most people that slip takes place by the movement of dislocations. Simple types of dislocations, the edge and the screw, had been described and the newly discovered phenomenon of polygonization had strengthened belief in at least the first of these. The elastic properties of dislocations were well understood and a detailed mathematical theory of their stress fields, strain energies, flexibility, forces of inter action, and mobility in simple lattices, already existed. The dynamical properties of dislocations were mainly unknown, although it had been recognized that dislocations could not move faster than sound, and that near this speed they would have enough kinetic energy to create new dislocations from themselves. This dynamic creation of dislocations was thought to be the basic process in the growth of slip bands. The source of the original dislocations was still a mystery, however.

Age-hardening had already been discussed in detail from the point of view of dislocations and a start had been made towards explaining the yield point of iron in terms of the anchoring of dislocations by segregated impurity atoms. The theory had also been applied to the analysis of the structures of grain boundaries, but not to the strengthening effect which they have. Several attacks on the work hardening problem had been made, but these had only penetrated far enough to show how formidable it really is. So long as it remained unsolved there could be little progress, apart from preliminary skirmishes, with the

closely related problems of recrystallization, creep, fatigue, and ductile fracture.

The most exciting of the new developments has been in crystal growth. Difficulties in the classical theory led to the idea that the growth of real crystals is catalysed at places where dislocation lines meet their surfaces. Many crystals contain dislocations because they could not have grown otherwise. Spectacular evidence supporting these ideas has come with the actual observation of the "growth spirals" predicted by the theory; it is now possible to locate, and even "see," dislocations in certain crystals.

An important new idea is the "Frank–Read source." This shows that certain geometrical arrangements of dislocation lines are able to create from themselves large numbers of new dislocations by a process not involving kinetic energy. This idea of static creation is not only opportune, because doubts have been raised about whether dislocations can move fast enough for dynamic creation; it is also something long needed for understanding slip bands, deformation twins, and work hardening. Another important idea is the creation and absorption of vacant atomic sites at dislocations. Things that may depend on this are the Kirkendall effect in diffusion, polygonization, work hardening, and annealing.

Various advances have been made in the elastic theory of dislocations, including the distribution of dislocations in slip bands, the energies of small-angle crystal boundaries, the effects of interatomic forces on the widths and mobilities of dislocations, and numerous effects concerning the interactions of dislocations with each other, with grain boundaries, and with other kinds of elastic singularities. The forms of dislocations in various crystal lattices have begun to be examined, and this has led to the recognition of various "imperfect" dislocations that line the boundaries of sheets of stacking fault in the lattices. These imperfect dislocations are believed to be important in determining slip planes, in deformation twinning, and in work hardening.

The theory of the yield point and strain ageing has been elaborated and experimental support gained for some of its predictions. Problems concerning the obstruction of slip at grain boundaries have been clarified and the attack on work hardening has continued. Observations on cross-slip, multiple slip, deformation bands, and the fine structure of slip bands, have suggested new ways of looking at the work hardening problem. In particular, interactions of dislocations other than the classical elastic ones have attracted some attention, for example, parallel dislocations on intersecting slip planes may sometimes run together to form immobile dislocations. Although the work hardening problem is not yet solved, or the problems of recrystallization, creep etc. that are related to it, several new advances have been made and

we can at least now say that dislocations seem to have that versatility in their properties that the complexities of these processes demand.

A satisfying recent trend in the theory as a whole is that it is becoming more quantitative, i.e. making quantitative predictions that can be tested experimentally. Two examples of this are the dependence on orientation of the energy of small-angle boundaries, and the time law of strain-ageing in iron.

The Growth of Imperfect Crystals from Vapours and Solutions

The driving force that makes a crystal grow from its vapour is determined by the ratio of the pressure of the vapour to the vapour pressure of the solid. Growth from a liquid solution similarly depends on the degree of supersaturation, while growth from a melt depends on the degree of undercooling.

The process by which a perfect crystal should grow has been analysed in detail in the classical theory of Gibbs, Volmer, Becker and Döring, and others (reviewed in references 2, 3, 4). Although the lattice is perfect, the crystal may have imperfect faces which consist of terraces of steps, each step being the edge of a single sheet of atoms or molecules which covers part of the face. Growth can occur easily on these imperfect faces, even when the degree of supersaturation is small, because atoms can deposit in the re-entrant corners provided by the steps and become firmly attached to the crystal by bonding with two or more of its atoms. This state of affairs cannot continue indefinitely, however. As the steps advance by the repeated attachment of new atoms they leave behind themselves smooth close-packed crystal faces on which there are no projections against which later atoms could lodge themselves. The crystal cannot grow further unless atoms condense on these smooth faces and assemble together to make a nucleus for the next layer. This is difficult because isolated atoms or small groups of them can be thrown off easily by thermal agitation, and the chance of a sufficient number of them coming together to make a permanent attachment to the crystal is small if the vapour is only mildly supersaturated. This conclusion can be verified by mathematical analysis of the nucleation process[2, 3, 4] and it has been shown in this way that perfect crystals could only grow at the rates observed with real crystals if the degree of supersaturation exceeded about 50 per cent.

Most experiments contradict this result. For example, Volmer and Schültze,[5] working on the growth of inorganic and organic crystals from vapours, showed that the rate of growth is practically as fast as the rate at which molecules strike the surface from the vapour, even when the supersaturation is only about 1 per cent. Moreover, the rate of growth is directly proportional to the degree of

supersaturation, whereas the theory predicts a spectacular exponential dependence.

FRANK[3, 4] realized that neither the theory nor the experiments were wrong, but that they applied to different crystals. The crystals which grow so easily in practice are *imperfect*. In particular they contain dislocations. The importance of dislocations in crystal growth can be seen as follows. What is needed for continual and rapid growth is an imperfection of the crystal faces which ensures that growth steps are

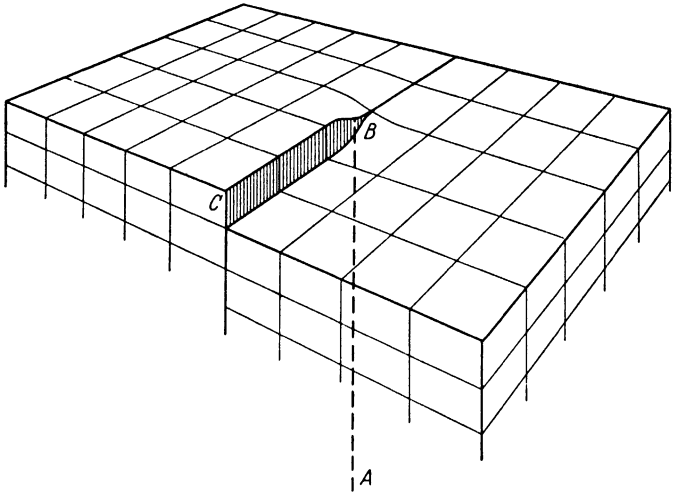


Fig. 1. A screw dislocation AB which ends at a point B on a crystal face

always on them, no matter how much they grow. As FRANK pointed out, this condition is ensured if screw dislocations end at these faces (or edge dislocations if these approach the surface obliquely so that there is a component of slip standing out of the surface). Fig. 1 shows a crystal face on which there is a *point of emergence* B for a screw dislocation AB in the crystal. From the definition of a screw dislocation (T.D.1, page 79) the direction of slip is along AB, so that a step BC is formed on the face running outwards from the point of emergence. Growth can occur by condensation along the step, and a new layer is added as the step rotates about its point of emergence.

The essential point is that this step never disappears. Because of its special topological property of ending in the middle of the face, which is a consequence of the geometry of the dislocation, the step is always present no matter how much the crystal continues to grow. It does not retain the simple form shown in the diagram, of course. Since that part of the step near to the point of emergence requires the addition of only a few atoms to make one revolution, whereas outlying parts need

large numbers of atoms, the centre rotates faster than the outside, and the step winds up into a "growth spiral," as in Fig. 2.

FRANK'S theory also considers growth from several points of emergence on the same face. The rate of growth is, over a wide range, not very sensitive to the density of dislocations which emerge on the face. An important case is that where two dislocations of opposite signs

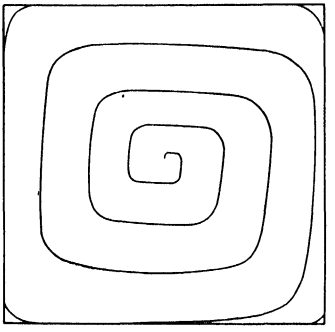
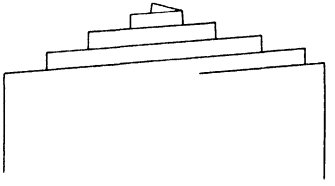
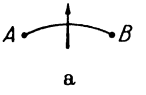
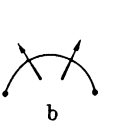
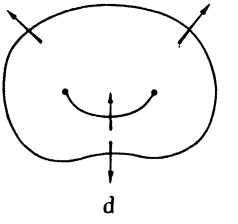


Fig. 2. A single growth spiral









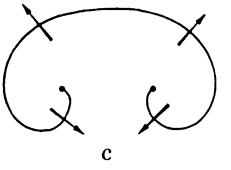


Fig. 3. Stages in the formation of a growth ring from a step joined to two points of emergence, A and B

emerge and are joined by a common growth step. This step generates not a continuous spiral but a series of concentric growth rings, as in Fig. 3.

Striking confirmation of this theory has been provided by the actual observation of growth spirals and rings, showing all the features expected from the theory, on the faces of various crystals. Fig. 4 shows the original example by GRIFFIN[6] on a crystal of beryl. This has been the fore-runner of many others observed on various inorganic and organic crystals.[7] By using electron microscopy, multiple-beam interferometry, and other optical methods, it has been proved that the steps are equal in height to the lattice constant of the crystal, or are small multiples of that height.

Because of their small lattice constants, and for other reasons, metals are not the easiest materials for showing growth steps. However, by depositing large molecules on them, the steps can be made visible, and this method has been used by FORTY[8] to show growth spirals in a magnesium crystal grown from a vapour. Fig. 5 is an example; the

height of the step in this case is probably equal to the c axis of the crystal.

FRANK'S theory explains the *growth* of imperfect crystals, not their nucleation, and separate considerations are needed to decide how the dislocations originate in the imperfect nuclei. FRANK suggests various possibilities for this, e.g. (*a*) nucleation in regions of high supersaturation, (*b*) plastic buckling of very thin and fragile nuclei, (*c*) nucleation on an imperfect substrate, (*d*) buckling under large strains created by a gradient of composition, and (*e*) the collapse of sheet-like cavities formed by the aggregation of vacant atomic sites. From a practical standpoint the origin of these imperfect nuclei is not important; what matters is that they grow in preference to perfect ones.

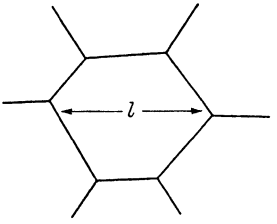


Fig. 6. Network structure made from dislocation lines

THE STATIC CREATION OF DISLOCATIONS

In 1950, at a conference in Pittsburgh, FRANK and READ[9] pointed out that slip bands might grow by a dislocation mechanism similar in some ways to that for the growth of crystals. The dislocations concerned in crystal growth need not lie in slip planes. In fact, as FRANK has suggested, they are likely to form a three-dimensional network structure, as shown in Fig. 6. In such cases we can expect dislocation lines which for part of their length lie in a slip plane, and then at some point turn out of this plane along an intersecting direction, or join other dislocations

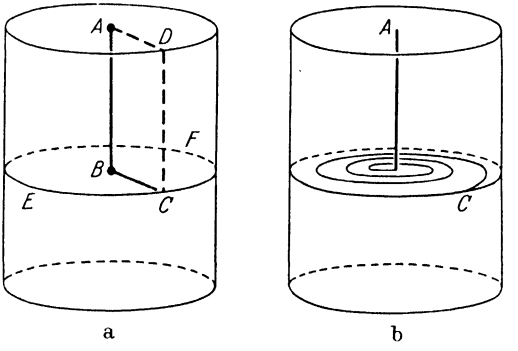


Fig. 7. The formation of a slip band by the rotation of a dislocation line BC about an internal point of emergence B

lying in other planes. A simple example of one of these Frank–Read sources is shown in Fig. 7. Here an edge dislocation ABC forms the internal boundary of a quarter-plane $ABCD$ of atoms. Suppose that the dislocation line BC can move freely in its slip plane CEF and that the part AB is for some reason unable to move. Then B is an internal

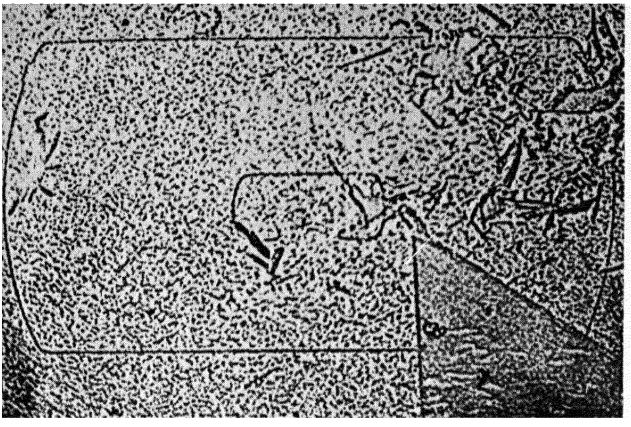


Fig. 4. Growth steps on a prism face of beryl (after GRIFFIN) × 600

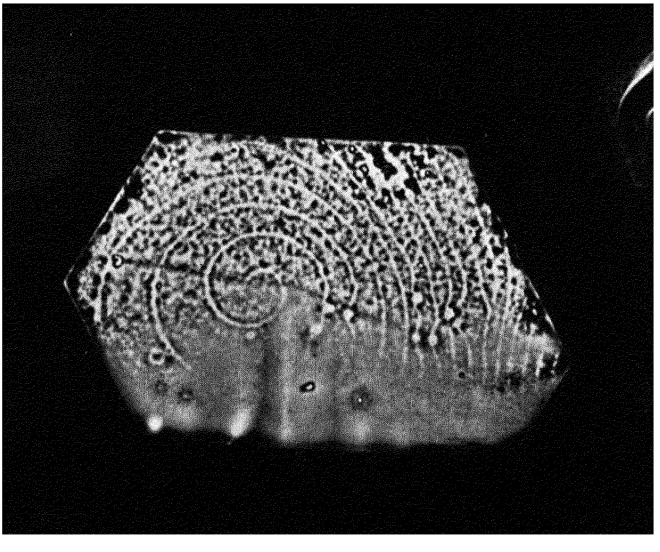


Fig. 5. Growth steps on a magnesium crystal (after FORTY) × 325

point of emergence and the dislocation BC must remain attached to it as it glides in its slip plane. The dislocation rotates about this point and winds up into a rotating spiral, as shown, which produces a translational slip of one atom spacing each time it makes one revolution.

Pieces of dislocation with two internal points of emergence, such as those in Fig. 6, can generate slip bands by making dislocation rings. Fig. 3 illustrates this process if the lines in these diagrams are thought of as dislocations. When the scale l of the network structure is small compared with dimensions of the crystal most of the Frank–Read sources will be of this type. Only those within a distance l from the surface will be of the single-ended type shown in Fig. 7.

The stress needed to start a source working is important. Consider a double-ended source, such as those in Fig. 6. The dislocation has a line tension (T.D.1, page 117). This causes it to contract in length until it is a straight line joining its two points of emergence. Before the source can be set in action this line has first to be bent into a loop of radius $l/2$ (Fig. 3). The applied stress σ needed to do this (T.D.1, equations 28 and 29) is about

$$\mu b/l, \qquad \ldots . \quad (1)$$

where μ is the shear modulus and b is the strength of the dislocation, i.e. the length of the slip displacement it produces.* Since b is normally one atomic spacing this stress is about the yield stress of a single crystal when l is 10^4 atomic spacings.

FISHER[10] has pointed out that single-ended sources of the same length begin to work at only one-half of the above stress. This is because the other end of such a dislocation is not attached to a fixed point, but to some point along the line where the slip plane meets the surface of the crystal. Provided that the dislocation meets this line perpendicularly, in order to preserve an equilibrium of forces here, this external point of emergence is free to move along the line. As a result it is only necessary to bend the dislocation into a loop of radius l, not $l/2$, to set it working. FISHER[10] has used this idea to explain an observation of BROWN and HONEYCOMBE[11] that, on a purely electropolished surface of aluminium, the slip appears in the form of innumerable fine markings, instead of the much coarser slip bands some 10^{-3}–10^{-4} cm apart, which are typical of strained crystals with mechanically worked surfaces. FISHER argues that when the surface is unworked many more Frank–Read sources are active at the surface than in the interior, since for a given length of source the surface ones become active at only one-half of the stress needed for the interior ones. On the other hand, when the surface has

* To conform to recent practice, we now use μ and b for those quantities symbolized by G and λ, respectively, in T.D.1.

been mechanically worked these surface sources are less active because of the cold-working strains there.

There can be little doubt that in practice a slip band can be made very quickly, in a fraction of a second. Can a single Frank–Read source make a slip band as fast as this? The critical thing is the radius of curvature of the central loop of the rotating dislocation spiral, which is about $\mu b/\sigma$. This central loop makes about $v/\{2\pi(\mu b/\sigma)\}$ revolutions per second, where v is the linear velocity of the dislocation line in the slip plane. The time needed to make a typical slip band, e.g. 1000 revolutions, is of order $10^3(2\pi\mu b/v\sigma)$ seconds. For v we shall assume the value 10^4 cm sec^{-1}, which is about 0·1 that of sound (see next section). Taking also $\sigma/\mu = 10^{-4}$ and $b = 3 \times 10^{-8}$ cm, the time to make the slip band is less than 0·0002 sec.

The Dynamic Properties of Dislocations

The above theory of the static creation of dislocations raises the question of the present status of the earlier dynamic theory (T.D.1, page 100), and this leads to the more general question of the dynamics of moving dislocations.

The original conclusion of Frenkel and Kontorova (T.D.1, page 99) that the energy of a dislocation rises to infinity with its speed according to the "relativistic" formula

$$U = U_0/(1 - v^2/c^2)^{1/2}, \qquad \ldots . \quad (2)$$

where U_0 is the rest energy of the dislocation and c is the speed of sound in the medium, has now been confirmed by the analysis of dislocations in three-dimensional bodies.[12, 13, 14] A simple case is that of a screw dislocation lying along the z axis and moving along the x axis at a steady speed v. When $v = 0$ the elastic equation of equilibrium governing the strain field of the dislocation is

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0, \qquad \ldots . \quad (3)$$

where w is the component of displacement along the z axis, the other two components being zero. For a simple screw dislocation of strength b this has the solution

$$w = b(\theta/2\pi) = (b/2\pi)\ \tan^{-1}(y/x), \qquad \ldots . \quad (4)$$

where θ is the angle measured about the screw axis from some arbitrary reference line.

When this dislocation is moving the equation of equilibrium is found by equating the force acting on any volume element due to the stress

field to the force of inertia due to the acceleration of the matter in this volume element. Equation (3) then becomes

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 w}{\partial t^2}, \qquad \ldots . \quad (5)$$

where t is the time and

$$c = (\mu/\rho)^{1/2} \qquad \ldots . \quad (6)$$

is the speed of transverse sound or shear waves in the medium, ρ being its density. We need a solution that reduces to equation (4) when $v = 0$, and moves along the x axis with a constant form and constant speed v. Equation (5) is a wave equation and it is known from general experience with these in electrodynamics that steady propagational solutions are formed by making a *Lorentz transformation*. In the present case this means substituting the variable

$$x_1 = (x - vt)/(1 - v^2/c^2)^{1/2} \qquad \ldots . \quad (7)$$

in the equation. Noting that for a steady propagational solution the relation $\partial^2/\partial t^2 = v^2(\partial^2/\partial x^2)$ holds, the equation then becomes

$$\frac{\partial^2 w}{\partial x_1^2} + \frac{\partial^2 w}{\partial y^2} = 0, \qquad \ldots . \quad (8)$$

which is identical in form with equation (3). Evidently the required solution is simply

$$w = (b/2\pi)\tan^{-1}(y/x_1) \qquad \ldots . \quad (9)$$

The field of the moving dislocation is the same as that of the stationary one except that x_1 replaces x. The factor $(x - vt)$ in x_1 accounts for the steady propagation of the dislocation while the Lorentz contraction factor $(1 - v^2/c^2)^{1/2}$ shows that the range of the stress field along the axis of motion is reduced. The energy of the moving dislocation can be found from this result. We add the density of elastic energy $\frac{1}{2}\mu[(\partial w/\partial x)^2 + (\partial w/\partial y)^2]$ to the density of kinetic energy $\frac{1}{2}\rho(\partial w/\partial t)^2$ and a short calculation for $\partial w/\partial x$, $\partial w/\partial y$, and $\partial w/\partial t$ then leads to equation (2).

This equation shows that a dislocation cannot reach the speed of sound because its energy would become infinite. An important question is whether it can reach a speed about 80 per cent that of sound, for then its behaviour would become significantly different from that of a slow dislocation; for example, it would have enough excess energy to create new dislocations. The practical speed limit is reached when the energy gained by the moving dislocation from the applied stress is balanced by that lost from the dislocation by various dissipative processes. Some of these processes have been considered by SEITZ,[15] ESHELBY,[16] LEIBFRIED,[17] and NABARRO.[18] The main ones are as follows:

Thermoelastic Damping

In an elastic solid a change in hydrostatic pressure causes a change in temperature. When an edge dislocation moves the hydrostatic stress field caused by the dislocation also moves and so the temperature at any fixed point is continually changed. Heat flows down the temperature gradients set up in this moving field. This is a thermodynamically irreversible process and, therefore, dissipates the mechanical energy of the system. ESHELBY[16] has shown that thermoelastic damping is unlikely to prevent the dislocation from reaching "relativistic" speeds.

Radiation Damping

This occurs when elastic waves are transmitted outwards from the moving dislocation. Because of the atomic structure of the crystal and the breakdown of Hooke's law at large strains (T.D.1, page 84), the precise arrangement at the centre of a dislocation alters periodically as it moves from one lattice row to the next. The frequency with which it alters is v/b, where v is the speed and b the lattice spacing in the direction of motion. Elastic pulses are transmitted out to the surrounding crystal with this frequency and dissipate some of the energy of the dislocation. The mathematical analysis of these pulsations is difficult, and their effect is correspondingly uncertain, but NABARRO[19a] believes that the dissipation is large enough to prevent relativistic speeds being reached unless the applied stress is at least one order of magnitude greater than the yield stress of soft crystals. SEITZ[15] has reached the same conclusion by a different line of argument.

Scattering of Sound Waves

When a sound wave runs into a dislocation both may be disturbed.[18] The inhomogeneous strain field at the centre of the dislocation refracts the sound wave through an angle. In turn the sound wave imposes an oscillating stress field which may make the dislocation oscillate in its slip plane and radiate out elastic waves. The effect of these is that the impinging sound wave exerts a force on the dislocation. When the dislocation is moving it runs into more sound waves on the leading side than on the trailing one and so is subjected to more forces pushing it back than forward. By assuming that a moving dislocation scatters sound waves in the same way that a line of atoms would, LEIBFRIED[17] concludes that the damping due to the oscillation of the dislocation is so severe that fast speeds are quite impossible at the yield stresses of soft crystals. However, NABARRO[18b] has pointed out that Leibfried's argument does not allow for the fact that the scattering power is proportional to the wavelength of the impinging waves. Those waves approaching the dislocation from behind have an increased effective

wavelength, due to the Doppler effect, and are, therefore, scattered more completely than those approaching from the front. In fact, this difference in scattering power exactly cancels the difference in the number of sound waves met by the dislocation on the two sides, so that Leibfried's effect does not exist, at least to the first order in v/c.

We see that, although no firm conclusion has been reached, opinion is generally against the idea that dislocations can reach relativistic speeds, at low stresses at least. This opinion is supported by an experiment of LEIBFRIED,[20] who studied the extension of loaded aluminium crystals with an extremely sensitive extensometer and could not detect any of the sharp sudden increments of strain that would be expected if slip bands were forming from dislocations moving at relativistic speeds.

Nevertheless, various suggestions have been made recently for dislocation processes involving relativistic speeds. OROWAN[21] has suggested that the unidirectional contraction in the stress field of a fast dislocation will cause it to turn out of its original slip plane, and he has used this to explain a spread of slip from one plane to another. HOLDEN[22] has suggested that, at the upper yield point in iron, a dislocation that has broken free from its solute atoms (T.D.1, page 97) reaches high speeds and uses its kinetic energy to free nearby anchored dislocations.

MOTT[23] and FISHER, HART and PRY[24] have used the idea of dynamic creation of dislocations at Frank–Read sources to explain the characteristic size of slip bands. They both consider double-ended sources and suggest that the stress $\mu b/l$ is necessary only to start the source working. The first loop of a sequence formed at the source must grow from a standing start and needs this stress, but all loops formed immediately afterwards in the sequence have a running start and do not need it. The dislocation line bridging the source has kinetic energy as it moves through its position of minimum length and, if dissipation is small, this energy is enough to carry the dislocation forward to the critical semicircular shape beyond which it can expand into the slip plane once more. If this is true it means that once a source starts working it continues to do so until the stress round it has been reduced temporarily by an amount $\mu b/l$. This fall in stress is considered to occur as a result of the "back-stress" set up from the dislocations themselves created by the source, either when they become piled up against some obstacle (such as a deformation band) in the crystal,[23] or while they are still advancing out into the slip plane.[24] In both cases several hundred dislocations have to be created before the stress at the source falls sufficiently to stop it working, and this is held to account for the slip "avalanches" that the observations of HEIDENREICH and SHOCKLEY,[25] and BROWN[26] seem to require.*

* See, however, the recent work of H. WILSDORF and D. KUHLMANN-WILSDORF, *Z. angew. Phys.*, 1952, in press.

The Characteristic Dislocations of Crystals

The most important feature of a dislocation is its *Burgers vector*. This is the vector defining the movement of the material on one side of a slip plane, relative to that on the other side, when the dislocation glides through this plane. A more sophisticated definition has been given by Frank.[27] We shall now consider what conditions have to be satisfied by the Burgers vectors of dislocations in various crystal structures. Dislocations with such vectors will be called the *characteristic* dislocations of the crystals.

The essential condition is that the slip displacement which the Burgers vector represents must be such as to translate the atoms from initial positions to final positions where they are mechanically stable,

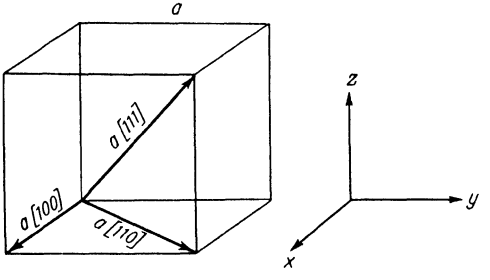


Fig. 8. Unit vectors in the simple cubic lattice

since mechanically unstable configurations cannot exist. There are two kinds of such allowed translations, those that preserve the lattice (identity translations) and those that generate a new configuration, such as a twinned lattice. Dislocations are said to be *perfect* or *imperfect*, respectively, according as their Burgers vectors give identity translations or new configurations. Imperfect dislocations lie along the boundaries of *stacking faults* in the lattice.

Dislocations with large Burgers vectors, even though these give allowed translations, may be unstable for another reason. The elastic energy of a dislocation is proportional to the square of its Burgers vector, so that the energy is reduced when large dislocations split up into smaller ones. For example, if the lattice constant is b a perfect dislocation with strength $2b$ lowers its energy from $4b^2$ to $2b^2$ if it splits into two widely separated dislocations each of strength b. Frank[28] has generalized this into a simple rule. In any combination or dissociation of dislocations the elastic energy is lowered when the sum of the squares of the Burgers vectors is smaller for the product dislocations than for the reacting ones.

The smallest perfect dislocations are called *unit* dislocations. In a simple cubic lattice of spacing a the unit vectors are $a[100]$, $a[110]$, and

$a[111]$, as shown in Fig. 8. In this notation a vector $a/n[hkl]$ is one with components ah/n, ak/n, and al/n, along the principal crystal axes. Dislocations with vectors $a[100]$ have the lowest energies, but the dissociations

$$a[111] \rightarrow a[110] + a[001] \rightarrow a[100] + a[010] + a[001] \quad . \quad . \quad . \quad . \quad (10)$$

lead to no reduction in elastic energy, according to Frank's rule. However, they raise the configurational entropy, since there are more ways of distributing many dislocations of small strengths than few dislocations of large strengths through the crystal.

A unit dislocation has lowest energy when its Burgers vector is parallel to a direction of closest packing in the lattice, which explains why this is the preferred direction of slip. The choice of slip plane is

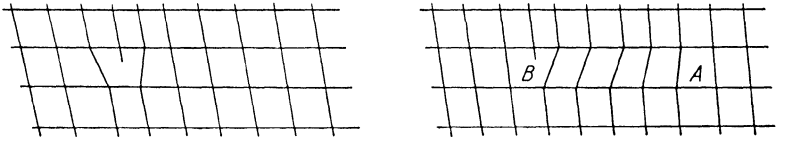


Fig. 9. Dissociation of a unit dislocation (a) into two partial dislocations (b)

less easily made, both in theory and practice. According to NABARRO[19b] this depends on two factors, both of which favour planes of wide spacings: (a) in a close-packed lattice a unit dislocation will dissociate into *partial* dislocations (see below) when it lies in a close-packed plane; this leads to a lowering of strain energy and so dislocations should collect preferentially in these planes; (b) the Peierls–Nabarro lattice force (T.D.1, page 90) resisting the movement of a dislocation is smallest when the Burgers vector is as small as possible and the spacing of the slip planes holding the dislocation is as big as possible. CHALMERS and MARTIUS[29] define a parameter β, which is the strength of the dislocation divided by the spacing of the slip planes, and point out that most observed slip systems can be predicted from the rule that the preferred one is that with the smallest β.

If the lattice is suitable a unit dislocation may dissociate into *partial* dislocations, which are imperfect dislocations with Burgers vectors less than one atom spacing in length. This can happen when the crystal has more than one mechanically stable configuration. Fig. 9 illustrates this diagrammatically. Here a unit edge dislocation dissociates into two partial dislocations, A and B, which are joined by a strip of faulted material where the configuration is altered. The vector sum of the Burgers vectors of the partial dislocations is equal to the Burgers vector of the original dislocation. The question of how far the

partial dislocations separate depends on the energy of the stacking fault between them. This energy is proportional to the area of slip plane covered by the fault and hence to the distance between the dislocations at its boundary. Thus if ε is the energy per unit area of the fault, the force per unit length pulling the dislocations together is also ε. The repulsive force due to the elastic interaction of the dislocations is, by a generalization of equation (9) in T.D.1, $\mu(\boldsymbol{b}_A \cdot \boldsymbol{b}_B)/2\pi kr$, where $(\boldsymbol{b}_A \cdot \boldsymbol{b}_B)$ is the scalar product of the Burgers vectors and r is the distance between the dislocations; $k = 1$ for screw dislocations, and $(1 - \nu)$ for edges, where ν is Poisson's ratio. The dislocations are in stable equilibrium at the distance

$$r = \mu(\boldsymbol{b}_A \cdot \boldsymbol{b}_B)/2\pi k\varepsilon \quad . \quad . \quad . \quad . \quad (11)$$

where these two forces balance. In general this is only a few atom spacings. The two partial dislocations thus do not separate very far and to a distant observer appear like a single dislocation; they are referred to jointly as an *extended* dislocation. In crystals in which unit dislocations are able to extend themselves the plane of slip is more clearly defined than in others, because an extended dislocation has a unique plane, the plane of its stacking fault, in which it can move easily.

We shall now consider the characteristic dislocations of some particular crystal structures.

Face Centred Cubic Crystals

The shortest lattice vector is $a/2[110]$, which joins an atom at a cube corner to a neighbour at a face centre, and defines the observed slip direction. Unit dislocations in {111} planes can lower their energies by splitting into two partial dislocations, the *half-dislocations* of HEIDENREICH and SHOCKLEY.[25] Consider slip on a (111) plane, as in Fig. 10. The vector $\boldsymbol{b}_1 = a/2[10\bar{1}]$ defines an observed slip direction. THOMPSON and MILLINGTON[30] and MATHEWSON[31] pointed out that, if atoms are regarded as balls rolling over one another, then one in a position of type B would roll most easily to a nearby position of type C. On an atomic scale of magnitude the atoms are expected to follow a zig-zag path $B \rightarrow C \rightarrow B$ *etc*, slipping alternately along $[2\bar{1}\bar{1}]$ and $[11\bar{2}]$, with a mean slip direction $[10\bar{1}]$.

A slip from B to C changes the order of stacking of the (111) layers. A perfect sequence . . . *ABCABCABC* . . . becomes the faulted sequence . . . *ABCACABCA* . . . which contains a stacking fault . . . *CACA* . . . consisting of a thin layer of hexagonal close packing. Such a fault preserves the typical close packing of nearest neighbours and so should have a low energy of formation. This is confirmed by the observations that these faults are formed freely in cobalt[32] and lithium[33] during phase transformations, and in certain copper-silicon

alloys during cold working.[33] Moreover, the surface energy of {111} boundaries of annealing twins in copper, where the stacking sequence is very similar to that of a fault, is only 21 erg cm^{-2} according to FULLMAN.[34]

HEIDENREICH and SHOCKLEY[25] proposed the dissociation reaction

$$\frac{a}{2}[10\bar{1}] \rightarrow \frac{a}{6}[2\bar{1}\bar{1}] + \frac{a}{6}[11\bar{2}] \qquad \ldots . \quad (12)$$

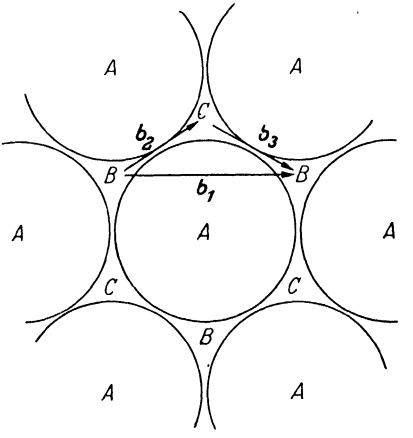


Fig. 10. Slip directions in a (111) plane

for a unit dislocation in (111). This reaction is formally possible since the sum of the components of the Burgers vectors is the same for both reactant and product dislocations. Thus the right-hand side sums to

$$\frac{a}{6}[2 + 1,\ \bar{1} + 1,\ \bar{1} + \bar{2}] = \frac{a}{6}[30\bar{3}] = \frac{a}{2}[10\bar{1}]. \qquad \ldots . \quad (13)$$

It is also energetically possible since by Frank's rule the energy is lowered from $a^2/2$ to $a^2/3$ when it occurs. In Fig. 10 the reaction is represented by $\boldsymbol{b}_1 \rightarrow \boldsymbol{b}_2 + \boldsymbol{b}_3$. The original unit dislocation splits into two partials which move apart in the (111) plane and generate a stacking fault in the region between them. The equilibrium spacing is given by equation (11) and if we take $\varepsilon = 20$ dyne cm^{-1}, $(\boldsymbol{b}_A \cdot \boldsymbol{b}_B) = a^2/12$, $\mu = 4 \times 10^{11}$ dyne cm^{-2}, and $a^2 = 10^{-15}$ cm^2, this spacing is 10 atoms.

A shear stress applied along $[10\bar{1}]$ drives both dislocations in the same direction, since they both contribute to a $[10\bar{1}]$ slip, but a stress applied along $[1\bar{2}1]$ could either separate them or close them together. The stress to separate them completely is large, about $10^{-2}\mu$, as is also the stress to force them together. Except in the presence of such stresses they always preserve their equilibrium spacing and move as a unit in the (111) plane. They cannot glide out of this plane, because

their stacking fault prevents this. If they should be forced together, to make a unit $a/2[10\bar{1}]$ dislocation, and this happens to be a pure screw it may then dissociate into two other partials, $a/6[21\bar{1}]$ and $a/6[1\bar{1}\bar{2}]$, lying in the $(1\bar{1}1)$ plane. This reaction, which would cause cross-slip,[35, 36, 37] might occur at places of high stress concentration or at points of emergence for the $a/2[10\bar{1}]$ dislocation, but apart from this deviations from slip on single (111) planes will be rare.

FRANK[27] has suggested another type of partial dislocation, the *sessile* dislocation, in the face centred cubic lattice. Suppose in Fig. 11 that we are looking sideways at the edges of (111) planes. There is a fault in the centre where part of the central A layer is missing. The neighbouring B and C layers collapse towards each other in this region and fit together in correct close-packed order in the faulted sequence $ABCBCA$.

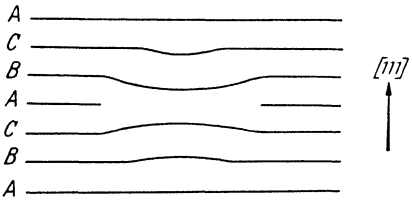


Fig. 11. A sessile dislocation

At the rim of this region an edge dislocation is formed with a Burgers vector $a/3[111]$. Its direction of glide is [111] but it cannot move this way because this would involve joining pieces of B layer to A layer etc which would leave badly disarranged atoms on the surface delineated by its track. The dislocation is therefore unable to glide, as its name implies. Sessile dislocations can be made either by the aggregation of vacant sites into sheets on {111} planes, forming disc-shaped holes in these planes whose sides then collapse, as in Fig. 11,[38, 39] or by the dissociation of unit dislocations; for example, the reaction

$$\frac{a}{2}[110] \rightarrow \frac{a}{3}[111] + \frac{a}{6}[11\bar{2}] \qquad \ldots . \quad (14)$$

produces a sessile dislocation and a half-dislocation, both lying in (111).

Since face centred cubic crystals can glide on intersecting {111} planes reactions may occur between dislocations moving on these planes. One of the most interesting is that studied by LOMER.[40] He considered two unit dislocations, both parallel to $[1\bar{1}0]$. The first moves in (111) and has a Burgers vector $a/2[10\bar{1}]$, while the second moves in $(11\bar{1})$ and has a vector $a/2[011]$. By Frank's rule they attract each other and so run down their respective slip planes until they meet and coalesce at the $[1\bar{1}0]$ line where these planes cross. The reaction is

$$\frac{a}{2}[10\bar{1}] + \frac{a}{2}[011] \rightarrow \frac{a}{2}[110]. \qquad \ldots . \quad (15)$$

This new dislocation is an edge and cannot glide easily because its slip plane, i.e. the plane containing its line and its Burgers vector, is (001) and not a common plane for slip.

COTTRELL[41] has pointed out that this dislocation may become completely immobile by dissociating,

$$\frac{a}{2}[110] \rightarrow \frac{a}{6}[11\bar{2}] + \frac{a}{6}[112] + \frac{a}{6}[110]. \quad . \quad . \quad . \quad . \quad (16)$$

The first two of these partial dislocations separate along their slip planes, (111) and (11$\bar{1}$) respectively, and produce stacking faults in these planes connecting them with the immobile $a/6[110]$ dislocation in the site of the reaction, the line where these faults join. The mobile dislocations separate by a few atoms until a position is reached where the elastic repulsion on them is just balanced by the surface tension of their stacking faults. This triangular group of parallel dislocation lines obviously cannot glide.

Close Packed Hexagonal Crystals

The unit vector in the basal plane is of length a, the lattice constant in this plane, and follows a $\langle 2\bar{1}\bar{1}0 \rangle$ direction of closest packing, which is also the observed slip direction. Dislocations with such vectors can follow any path through the crystal but prefer to lie in basal planes where they can dissociate into half-dislocations, as in the face centred cubic lattice. Basal slip is common in this structure. The shortest lattice vector out of the basal plane is [0001] and of length c, the lattice constant along the hexagonal axis. Dislocations with this vector are not easily moved but are important in crystal growth. The growth spiral shown in Fig. 5 starts from such a dislocation. Dislocations with Burgers vectors along the hexagonal axis may also be important as *pole* dislocations in deformation twinning[42, 43] (see below).

Body Centred Cubic Crystals

The shortest lattice vector is $a/2[111]$ which joins an atom at a cube corner to the one in the cube centre, and defines the observed slip direction.

This lattice can be twinned by a shear of $1/\sqrt{2}$ or $\sqrt{2}$ in a $\langle 111 \rangle$ direction on a $\{112\}$ plane, and twins of this kind can be produced by deformation in iron. This raises the possibility of a unit dislocation in a (112) plane dissociating as follows

$$\frac{a}{2}[11\bar{1}] \rightarrow \frac{a}{6}[11\bar{1}] + \frac{a}{3}[11\bar{1}], \quad . \quad . \quad . \quad . \quad (17)$$

to make two partial dislocations separated by a stacking fault consisting of a monolayer of twinned crystal. This dissociation is less certain than those discussed above because the original packing of nearest neighbours is not preserved on the twin interface. These partial dislocations

have parallel Burgers vectors and cannot be separated by a homogeneous stress. Also, they are both of pure screw type when they lie along $[11\bar{1}]$, and since $[11\bar{1}]$ is the line along which (112), $(\bar{1}21)$ and $(2\bar{1}1)$, intersect, they can glide out of their original slip plane when in this orientation. Thus the limitation to a single slip plane is less severe than in the other crystal structures considered.

A sessile dislocation may perhaps be formed by a dissociation of the type

$$\frac{a}{2}[111] \rightarrow \frac{a}{3}[112] + \frac{a}{6}[11\bar{1}] \qquad \ldots . \quad (18)$$

when the reacting dislocation lies in (112). The dislocation $a/3[112]$ is pure edge and sessile. It forms the boundary of a stacking fault made by inserting two fault planes in the (112) stacking sequence. The sequence is . . . *ABCDEFA* . . ., and the fault is of the kind . . . *ABCDCDEFA* . . . .

The Growth of Deformation Twins

The idea that deformation twinning takes place by the continuous growth on an atomic scale of twinned material has often been suggested, and the general arguments for a dislocation mechanism of twinning are the same as for slip: it is unbelievable that the atoms all move simultaneously, and the stress for twinning is far below the theoretical shear strength of a perfect lattice. Examples of deformation twinning are known which consist either of simple shears on the twinning plane (body centred cubic iron) or of shears accompanied by localized re-arrangements ("shuffles") of the atoms (hexagonal metal crystals). The shears can be produced directly by partial dislocations, but not the shuffles; it is presumed that these follow the shears automatically, since the structure will generally be mechanically unstable until the twinning process is completed.

The main problem is to understand how twinning develops homogeneously through successive planes. This seems to require either a twinning dislocation on every plane, or the motion of a single dislocation successively from plane to plane. Cottrell and Bilby[44] have recently suggested a mechanism, based on that of Frank and Read for slip, whereby successive twinning planes are swept consecutively by a single twinning dislocation. Fig. 12 illustrates it. Here *OA*, *OB*, and *OC* are dislocation lines and *CDE* is the twinning plane. The twinning dislocation *OC* has its Burgers vector in this plane, and this vector is such that the dislocation produces the correct shear for twinning between the atoms bounding this plane as it sweeps through the plane, rotating about its point of emergence *O*. The dislocations *OA* and *OB* are immobile and each has a component in its Burgers vector along the normal to the twinning plane; this component of the vector makes the

dislocations OA and OB partly or wholly of screw character, with a pitch equal to the spacing of the twinning layers. Because of this, these twinning layers are not a number of separate planes stacked on one another, but the leaves of a single continuous spiral surface (a helicoid or Riemann surface), winding round the *pole* dislocation AOB like a spiral staircase. It follows that, as the twinning dislocation rotates round the pole dislocation, it not only produces a monolayer sheet of twinned crystal but also climbs up the "staircase" to the next layer, where it repeats the process. In this way a thick layer of twin can be built up. The time needed to make a twin by this process is similar to that to make a slip band by the Frank–Read mechanism, and it is known experimentally that deformation twins can form very quickly, in a few microseconds.[43, 45]

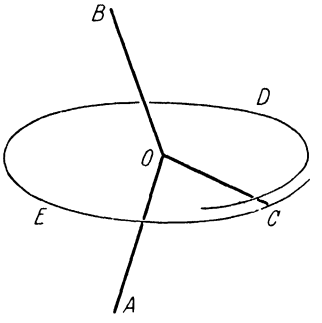


Fig. 12. The pole mechanism of deformation twinning

Cottrell and Bilby have applied this theory to the body centred cubic lattice. Suppose that the line AOB in Fig. 12 was originally a unit dislocation with Burgers vector $a/2[111]$ and that the part OB lies in a (112) plane. Under a suitable stress the reaction 18 can occur at OB, leaving OB itself as a sessile dislocation $a/3[112]$ and producing the twinning dislocation OC, $a/6[11\bar{1}]$. If OC turns into a twinning plane CDE which is $(\bar{1}21)$ or $(2\bar{1}1)$ it can rotate in this about the pole dislocation and make a twin.

A similar analysis can be made for the face centred cubic lattice, starting from the reaction 14, but in this case only a twin monolayer can be formed, not a thick twin, because the OB part of the pole dislocation, $a/3[111]$, has to lie in the (111) twinning plane and the $a/6[11\bar{2}]$ twinning dislocation runs into it at the end of one revolution in this plane. This process may be important in the production of single stacking faults in this lattice, a problem which has also been considered from the dislocation standpoint by Christian.[46] Pole mechanisms for the growth of deformation twins in hexagonal metals have been suggested recently,[42, 43] and in these the pole dislocation has a Burgers vector along the hexagonal axis. It is probable that shear transformations of the martensitic type occur by the movement of dislocations,[47] and Frank[48] has recently discussed the crystallography of the austenite-martensite transformation from this standpoint. The problem is complicated by the fact that the deformation is not a single shear.[49]

Interactions of Dislocations

Three kinds of interactions involving dislocations have now been recognized, two of which are elastic in nature while the third is

geometrical. The first kind of elastic interaction is the ordinary long-range one discussed in T.D.1. We call this the "Hookean" interaction because it depends on stress fields round dislocations which in turn depend on Hooke's law being obeyed at small strains. Recent work on Hookean interactions has been concerned either with confirming existing results by methods of greater rigour and power,[50–54] or with developing new results, mainly about arrays of dislocations in slip planes or in crystal boundaries. Non-Hookean interactions originate in deviations of the atomic law of force in the slip planes from Hooke's law, and are important when considering the stress necessary to drive a dislocation through a grain boundary, or a sheet of stacking fault, or any similar inhomogeneity in the material. Geometrical interactions occur when intersecting dislocations in different glide systems cut through one another. The "jogs" that are then formed in these dislocations have some interesting properties; in particular jogs in screw dislocations create vacant atomic sites or interstitial atoms when these dislocations are moved. We shall discuss some of the more interesting of these various effects.

Dislocations in Crystal Boundaries

The idea that grain boundaries between crystals inclined at small angles to each other are composed of dislocations has already been discussed in T.D.1. It has attracted interest because the properties of these small-angle boundaries can be measured experimentally and so form a testing ground for the theory. Shockley and Read[55] have been mainly responsible for developing this line of approach.

There are three properties capable of experimental study:

(*a*) The relation between the spacing of dislocations in the boundary and the angle between the grains.

(*b*) The mobility of the boundary under stress.

(*c*) The energy of the boundary, and its variation with the angle between the grains.

We shall discuss these from the standpoint of the classical Burgers tilt boundary shown in Fig. 13. Here two crystals are joined along the plane $x = 0$ after being rotated by equal and opposite amounts about the y axis. The angle of rotation is

$$\theta = 2 \tan^{-1} (b/2h) \simeq b/h, \qquad \ldots . \quad (19)$$

where h is the spacing of edge dislocations of strength b in the boundary. When this angle is sufficiently small ($< 10^{-3}$ radians) the spacing between the dislocations is large enough to be resolvable in a microscope. Shockley and Read pointed out that the dislocation structure of such boundaries has in fact already been resolved by Lacombe[56] in a study of "veining" in aluminium. Lacombe found that a network of

faint lines existed in a crystal which consisted of rows of separate etch pits, and that there was a small difference in orientation between the regions separated by these "veins." SHOCKLEY and READ suggested that the etch pits were formed at the points of emergence of the dislocations in the boundaries. The order of magnitude of the angle θ across such a boundary was 10^{-3} radians or less, while the spacing of the etch pits was not more than about 3×10^{-4} cm. Further work along these lines should enable a rigorous check of equation (19) to be made.

The dislocations of the simple boundary of Fig. 13 are geometrically free to glide as a unit along their slip planes, although this property does not extend to some of the more complicated boundaries containing two or more kinds of dislocations.[55] Moreover, the force from an applied stress on any one of the dislocations in this simple boundary is exactly the same as it would be (T.D.1, equation (1)) if this dislocation were alone present in the crystal, and so the boundary should move when a small shear stress is applied to the crystal. Boundaries of this simple and mobile kind are frequently formed during the bending and cutting of hexagonal crystals,[57] where they are called *bend planes*, and direct visual observation of their movement under stress has now been obtained.[58]

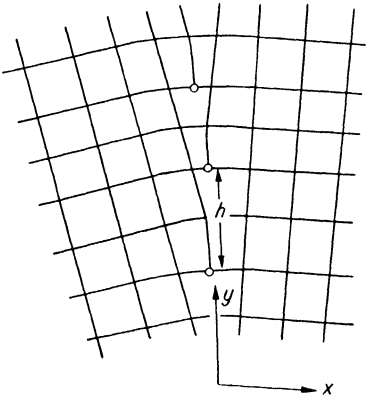


Fig. 13. The simple Burgers boundary

The reason why the boundary of Fig. 13 is mechanically stable was given in T.D.1. When the horizontal distance between two parallel edge dislocations of the same sign is smaller than the (vertical) distance between their slip planes, they are attracted by their Hookean interaction along the slip direction and pull into alignment, one above the other. The strain energy is smaller for the arrangement of dislocations in Fig. 13 than for other arrangements. The boundary in this diagram is a surface across which each crystal exerts forces on the other. These forces are associated with local indentations along the lines where a step on one face is pressed into the other face. These steps occur at intervals h up the face, so that the forces alternate in sign along the y axis with a period h. Elasticity theory shows that periodic forces of this kind, with zero resultant and moment, produce stress fields that penetrate inwards from the surface hardly beyond a distance of order h.

Now the strain energy of a dislocation is logarithmically proportional to the radius of its stress field (T.D.1, equation (7)). It follows that the strain energy of a dislocation in the boundary is proportional to log h,

and this is smaller than could be obtained by any other distribution of the dislocations in the crystal. Dislocations occur at intervals h along the boundary so that the elastic energy of unit area of the boundary is proportional to $(1/h) \log h$, i.e. to $-\theta \log \theta$. The energy contributed by the cores of the dislocations is proportional to the number of these, i.e. to θ. Hence the surface energy of the boundary should increase with angle according as the function $\theta(A - \log \theta)$, where A is a constant.[55]

A general formula for the stress field of the boundary can be found by adding together the fields of the dislocations in it.[55, 59] This gives the shear stress acting on the slip plane of one of the dislocations in the slip direction as

$$\sigma_{xy} = \frac{\mu b}{2\pi(1-\nu)} \frac{\pi^2 x}{h^2} \frac{1}{\sinh^2 (\pi x/h)}, \qquad \ldots \quad (20)$$

where x is the distance in the slip direction from the boundary. We can use this to find the strain energy of the boundary. Suppose that we start with a single crystal and make the boundary by (a) cutting from one side to the centre along every slip plane that is to contain a dislocation, and (b) applying to the two faces of each of these cuts the force needed to displace one face rigidly past the other by the distance b. In this process all dislocations in the boundary are to be created simultaneously so that the applied forces are the same on all cuts. The total work done by these forces is the energy of the boundary. We know the magnitude of the force at every point on the cut faces from equation (20), and also at every stage during the creation of the boundary because, from Hooke's law, the displacement of the cut surfaces increases linearly with the force on them. Thus the strain energy per dislocation is $\frac{1}{2} \times$ force $\times$ displacement, integrated along the slip plane, i.e.

$$\tfrac{1}{2} b \int_{x_0}^{x_\infty} \sigma_{xy} dx = \frac{\mu b^2}{4\pi(1-\nu)} [\log (\pi x_0/h) + \log 2 - 1],$$

where x_0 is the radius of the core of the dislocation, x_∞ is the length of the cut plane, and where equation (20) has been used for σ_{xy}. Introducing θ for b/h, the energy of unit area of boundary is

$$\gamma = \frac{\mu b}{4\pi(1-\nu)} \theta(A - \log \theta),$$

where

$$A = 1 + \log (b/2\pi x_0). \qquad \ldots \quad (21)$$

This result is a special case of a more general one derived by SHOCKLEY and READ for boundaries containing more than one kind of dislocation. The contribution of the core energies can be included in the formula by suitably altering the magnitude of A, and hence x_0. Using the dislocation model of PEIERLS and NABARRO (T.D.1, page 90), VAN DER MERWE[60] has shown that A is typically about 0·5.

The measurement of relative grain boundary energies in solids has been made possible by the methods of SMITH[61] and SHUTTLEWORTH, KING and CHALMERS,[62] which determine the relative surface tensions of boundaries from the angles at which three of them meet along a line. These have been used to test the $\theta(A - \log \theta)$ law by DUNN and LIONETTI[63] and AUST and CHALMERS.[64] Fig. 14 shows the results of AUST and CHALMERS on lead compared with such a curve, using the value $A = 0{\cdot}2$.

It is surprising that there is agreement not only at low angles, where it might be expected, but also at large angles where the theory is no

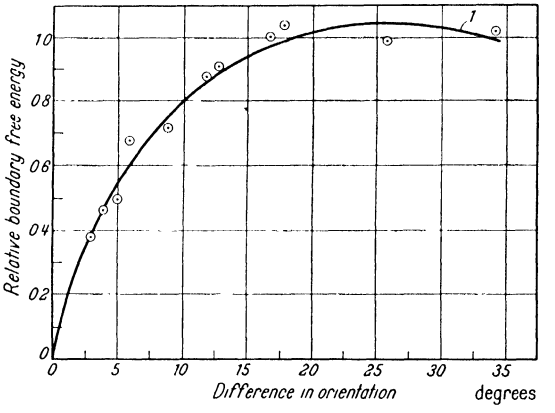


Fig. 14. Variation of boundary free energy with difference in orientation between crystals of lead; (1) theoretical curve with $A = 0{\cdot}2$; ○ experimental values of AUST and CHALMERS

longer reliable. LOMER and NYE[65] have shown with the bubble raft (T.D.1) that the nature of the boundary changes rapidly, when θ exceeds a few degrees, from one made of dislocations to one in which dislocations cannot be recognized. In particular the dislocations, which are wide (T.D.1, page 91) when θ is small, contract when θ approaches the range where the spacing between dislocations is equal to their width. This contraction changes them into defects resembling in some ways rows of vacant sites.

NABARRO'S[19] explanation of the contraction is as follows. It is possible to regard an ordinary dislocation as a narrow assemblage in the slip plane of many dislocations of infinitesimal strengths.[66] The width β of the dislocation is governed by the elastic repulsions of these infinitesimal dislocations and by the misfit energy in the strip of slip plane containing it. Now a wall of dislocations such as that in Fig. 13 can be regarded as a set of parallel and closely spaced walls, the component dislocations of which are the infinitesimal dislocations which define β. This becomes important when the spacing h is smaller than

β, for then β is determined by the elastic repulsion of walls of infinitesimal dislocations. As equation (20) shows, the stress field of a wall of dislocations has a shorter range than that of a single one. The walls of infinitesimal dislocations therefore move closer together as θ is increased into the range where $h \simeq b$, which is the effect that LOMER and NYE observe.

DISLOCATIONS IN SLIP BANDS

The slip displacement at one end of a long slip plane can differ substantially from that at the other end, as may be seen from the fact that the arrangement of slip bands on one side of a worked crystal is often quite different in detail from that on the other side. This difference in displacement is taken up by a set of dislocations spaced through the slip plane. Various aspects of such arrays of dislocations are interesting. On the purely formal side, it is possible to regard the whole set of dislocations as a single dislocation of a more general kind, with a Burgers vector that is not constant along the slip plane.[67–70] Another approach is that of NYE,[71] who, by a photoelastic study of internal stresses in deformed silver chloride crystals, has shown very clearly that some of these stresses are caused by rows of edge dislocations, all of the same sign, contained in bent slip planes.

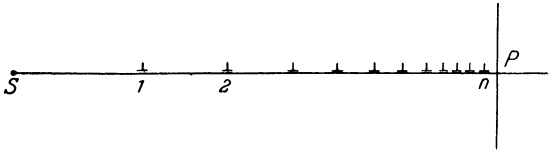


Fig. 15. A row of edge dislocations pressed against an obstacle

In T.D.1, page 105, it was shown that a row of n edge dislocations, all of the same sign, which are pushed against an obstacle in the slip plane exert a strong pressure against it (Fig. 15); in particular the leading dislocation in the sequence is driven against it by a stress n times the applied stress. It was also pointed out that the dislocations are spaced more closely together at the obstacle end of the sequence, and that because they are all of the same sign they bend the slip plane. The point of interest about this arrangement of dislocations is that it represents the spearhead of a slip band held up by an obstruction. The pressure exerted by the dislocations amounts in effect to a concentration of shear stress at the end of the slip band, and this concentration may enable the slip band to push through the obstacle into the crystal beyond it.[72]

This stress concentration depends upon L, the length of the slip plane from the source S to the obstacle P, since this is a factor on which the number of dislocations piled up depends. We can estimate n

roughly as follows. Suppose that the source becomes active when an arbitrarily small stress acts on it, and that this stress is also sufficient to make dislocations glide in the plane. If a stress σ is applied to the crystal the source feeds dislocations into the slip plane until the back stress from these cancels the applied stress. Roughly, then, the source stops working when the applied stress is relaxed to zero inside an approximately circular region of radius L centred on the source. The elastic strain in this region is correspondingly reduced from the value σ/μ to zero, and this is accomplished by a slip displacement $2L(\sigma/\mu)$ at the source. To produce this displacement by dislocations of strength b the number required is

$$n = 2L\sigma/\mu b. \quad \ldots . \quad (22)$$

A more accurate theory has been given by ESHELBY, FRANK and NABARRO,[73] which enables the positions of the dislocations to be found. These positions, denoted by x_i ($i = 1, 2, \ldots n$), are defined as the values of the distance x along the slip plane at which the polynomial

$$f(x) = (x - x_1)(x - x_2) \ldots (x - x_n) \quad \ldots . \quad (23)$$

is zero. The equilibrium of the dislocations can be stated in the form of a differential equation for $f(x)$ the solution of which is a polynomial equal to zero at n different values of x. In the problem of Fig. 15 it turns out that the distribution of the n dislocations along the slip plane is identical with that of the radial distribution of nodes in the ns state of the hydrogen atom.

For large values of n the following results are obtained:[73]

(a) The number of dislocations packed into a length L of the slip plane is

$$n = \pi L\sigma k/\mu b, \quad \ldots . \quad (24)$$

where k is unity for a screw dislocation, $(1 - \nu)$ for an edge and ν is Poisson's ratio. This compares with our equation (22) above.

(b) The distance between the two leading dislocations is

$$d = 1{\cdot}84(\mu b/2\pi kn\sigma). \quad \ldots . \quad (25)$$

(c) The shear stress on the glide plane at a distance x ahead of the leading dislocation is practically the same as that due to this dislocation alone in the region where $x \ll d$. At distances where $d \ll x \ll L$ the stress concentration factor is approximately

$$1 + (L/x)^{1/2}. \quad \ldots . \quad (26)$$

This stress concentration can be large, and the dislocations then produce the same stress concentration as does a freely slipping crack of the same length.[74]

(d) The total slip produced by the sequence is the same as if a single dislocation of strength nb were moved a distance $3L/4$.

FRANK[75] has suggested that this high concentration of stress at the end of a piled up row of dislocations may be sufficient in a polycrystalline specimen to force the creation of dislocations on the other side of a grain boundary which is holding up the row. In the next section we shall turn to consider the nature of the force preventing the leading dislocation from entering the grain boundary.

Note added in proof: D. KUHLMANN-WILSDORF, J. H. VAN DER MERWE, and H. WILSDORF, [*Phil. Mag.* 43 (1952) 632] have recently obtained experimental evidence supporting the theory of dislocations in slip bands. With the aid of an electron microscope they measured the spacing, length, and amount of glide, of individual fine slip lines formed in lightly strained aluminium. The number of dislocations in an elementary slip line was calculated from these measurements, and it was shown that the observed spacing of the lines is comparable with the theoretical critical distance below which rows of dislocations in parallel slip planes cannot pass each other.

OBSTRUCTION OF DISLOCATIONS AT GRAIN BOUNDARIES

As a dislocation approaches a grain boundary it may be subjected to three forces, one long-range and two short-range ones. The long-range force occurs where there is a change of elastic constants beyond the boundary, for example, where the boundary is a free surface, or where the boundary is between two phases of an alloy, or between two crystals in an elastically anisotropic material. The force occurs because, as the dislocation approaches the boundary, more of its strain field is imposed on the material beyond, and the energy to deform this is different from that to deform the material on the near side of the boundary. The only case that has been studied in detail is that of a free surface. KOEHLER[76] showed that a dislocation is attracted towards this by an "image" force inversely proportional to the distance from the surface.

When applying this result in practice it should be remembered that many surfaces are covered with films, e.g. oxide coats on metals. The presence of an adherent film turns the boundary of the crystal into an interface between phases, and this will generally repel dislocations approaching it from inside the crystal. There is much experimental evidence for the fact that oxide coats make slip more difficult in soft metal crystals.[77] A striking indication of the pressure of piled up dislocations against such oxide coats has been given by BARRETT;[78] he showed that if the oxide coat were removed from a zinc crystal that had recently been twisted plastically, the crystal would spontaneously twist a little further as the piled-up dislocations were allowed to run out.

The two short-range forces acting on a dislocation approaching a boundary are that due to the stress field of the dislocations in a boundary between different crystals, and one due to the fact that the law of force between atoms in the slip plane changes at the boundary. This latter is an example of a non-Hookean interaction and we shall use it to

illustrate the general theory of such interactions, developed by JASWON and FOREMAN.[79]

Suppose in Fig. 16 that a dislocation in the slip plane PQ is approaching a boundary AB beyond which the law of force between the atoms in the slip plane is different. In fact QR need not be a real slip plane at all, but merely a geometrical continuation into the region beyond the boundary. To make a definite picture, suppose that AB is the boundary between two elastically isotropic crystals that differ only in orientation from one another. Then for small strains the law of force between the atoms across the slip plane is the same in PQ and QR, and is Hooke's law $\sigma_{xy} = \mu\phi/d$, where σ_{xy} is the shear stress, ϕ is the shear displacement

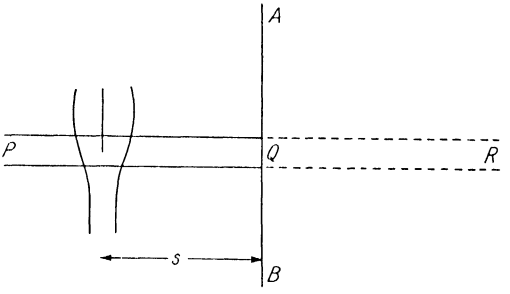


Fig. 16. A dislocation approaching a boundary

and d is the spacing of the planes. For large strains the force is larger in QR than in PQ, since QR is not physically suitable for slip.

At some distance from it the dislocation produces strains in the boundary that are small enough for Hooke's law to be obeyed. There is no Hookean distinction between PQ and QR, so that no force is caused by the change in the nature of the slip plane until the dislocation is close enough for Hooke's law to be no longer obeyed at the boundary. This non-Hookean force is thus a short-range one.

JASWON and FOREMAN give a theorem for calculating s, the distance to which a dislocation can be pushed towards the boundary by an applied stress σ. Later, NABARRO[19b] showed that this could be derived rather simply by applying the principle of virtual work. Consider a piece of slip plane lying between x and $x + dx$ from the dislocation, and denote the energy in this piece by $g(\phi)dx$ when $x < s$, and by $h(\phi)dx$ when $x > s$. The shear displacement ϕ depends on x, and the laws of force in the two regions of the slip plane are $\sigma_{xy} = -dg/d\phi$ and $\sigma_{xy} = -dh/d\phi$, respectively. Let the entire field of displacement representing the dislocation be moved forward rigidly by a small distance δx. The strain energy in the (Hookean) regions outside the slip plane is not altered, while the energy in the slip plane is altered by the same amount as if the boundary AB were moved from s to $s - \delta x$.

This energy change is $\{h(\phi_s) - g(\phi_s)\}\delta x$, where ϕ_s is the value of ϕ at s. The work done by the applied stress in moving the dislocation is $\sigma b\delta x$ per unit length. Hence the condition for equilibrium is

$$\sigma b = h(\phi_s) - g(\phi_s). \quad \ldots \quad (27)$$

This is an equation for s and to solve it we have first to choose suitable functions for $g(\phi)$ and $h(\phi)$. JASWON and FOREMAN use the Peierls sinusoidal law (T.D.1, page 90) for g and Hooke's law for h. Next we have to determine ϕ_s as a function of s; the simplest procedure here is to take an average of the values for ϕ obtained by assuming in turn that $g(\phi)$ or $h(\phi)$ is valid throughout the whole slip plane.

Using the above laws of force JASWON and FOREMAN show that the stress needed to hold the dislocation at s is given approximately by

$$\sigma = (\mu/6\pi^2)(\beta/s)^4, \quad \ldots \quad (28)$$

where $\beta = b/2(1 - \nu)$. The short-range character of this interaction, varying as s^{-4}, is to be noted. As a result, the stress needed to hold a dislocation at 10 atom spacings from the boundary is only $10^{-5}\mu$, whereas that needed to hold it at 1 atom spacing is nearly as large as the theoretical shear strength of the lattice, $\mu/30$.

MECHANICAL HYSTERESIS IN POLYCRYSTALS

One of the effects of the obstruction of dislocations at grain boundaries is that a polycrystal should show mechanical hysteresis.[80] Other obstacles such as mosaic boundaries, oxide films, and stacking faults, may act in the same way and much of the following discussion can also be applied to them.

For simplicity we suppose that the applied stress is only sufficient to produce slip in a few favourably oriented grains, although the theory can also be applied to larger deformations if attention is concentrated on the excess of slip in one grain over that in its neighbours. When the proportion of plastic grains is small it is fair to assume that each of these behaves as if it were mounted in an elastic matrix. Under these circumstances the maximum slip it can undergo is that which relaxes the stress in it from the initial value σ to some lower one σ_i, which is the critical stress below which dislocations are no longer mobile. If the linear dimension of the grain is L the total slip displacement in it is of order $L(\sigma - \sigma_i)/\mu$, and if the proportion of such grains in the specimen is α the overall strain produced by slip in them is of the order

$$\alpha(\sigma - \sigma_i)/\mu, \quad \ldots \quad (29)$$

when $\alpha \ll 1$.

The stress σ_i takes account of the various obstacles, such as foreign atoms and precipitated particles, that a dislocation has to overcome

when gliding through its slip plane. We can regard this stress as acting like a frictional force which prevents a dislocation from moving in either direction unless the applied stress is greater than it.

This model has been used by NABARRO in the following way to discuss mechanical hysteresis.[80] Suppose that the applied stress is increased from zero towards a value σ which exceeds σ_i in a few of the grains. At first the deformation is purely elastic. When the stress reaches σ_i dislocations begin to move and pile up at the boundaries of the plastic grains, their mutual repulsion producing a back stress σ_d. At the applied stress σ dislocations move up to the boundaries until the back stress acting on each of them reaches the value given by $\sigma_d = \sigma - \sigma_i$. Plastic flow then stops. On unloading no plastic flow can occur until the applied stress is reduced by $2\sigma_i$, so that the unloading curve starts off as a straight line. When the applied stress is reduced to $\sigma - 2\sigma_i$ the dislocations begin to move backwards into the centre of the grain, giving a reversed plastic flow. This becomes possible because the actual stress acting in this direction on each one of them is now $\sigma_d - (\sigma - 2\sigma_i)$ which, since $\sigma - \sigma_i = \sigma_d$, is equal to σ_i.

When $\sigma > 2\sigma_i$ the applied stress $\sigma - 2\sigma_i$ is positive, so that reversed plastic flow begins before the specimen has been fully unloaded. On the other hand, when $\sigma < 2\sigma_i$ the specimen unloads completely without undergoing reversed plastic flow. Two effects can then be observed. If the applied stress is taken through zero and increased in the reverse direction, the elastic limit at which flow begins in this direction is $2\sigma_i - \sigma$, which is numerically smaller than the original elastic limit, σ_i. This is the *Bauschinger effect.* Instead of loading in the reverse direction, the specimen can be rested in the unloaded state. The stress σ_d ($= \sigma - \sigma_i$) is then not strong enough to overcome the frictional stress σ_i due to the obstacles in the slip plane, but even so the dislocations may occasionally surmount these obstacles with the aid of thermal agitation. This is an example of the *elastic after-effect.*

KUHLMANN[81] has shown that this kind of model can be used to explain why the residual strain in the after-effect changes logarithmically with time. The stress σ_d depends on the number n of dislocations piled against the boundary. The activation energy for the return of a dislocation is a function of $\sigma_i - \sigma_d$ and hence of n; for small changes in n it can be written in the form $W - \beta n$, where W and β are constants. The rate of change of n due to the movement of dislocations back into the grains is $dn/dt = -A \exp\{-(W - \beta n)/kT\}$, where A is a constant. This integrates to give

$$n = n_0 - (kT/\beta) \log (1 + t/t_0), \quad \ldots . \quad (30)$$

where n_0 is the initial value of n and t_0 is defined by

$$n_0 = -(kT/\beta) \log (\beta A t_0/kT). \quad \ldots . \quad (31)$$

The residual strain is proportional to n and hence depends logarithmically on time.

Dislocation Jogs and the Creation of Vacancies

We turn now to the third kind of interaction involving dislocations, the geometrical interaction which occurs when two non-parallel dislocation lines cut through each other. Heidenreich and Shockley[25] pointed out that this would be a source of hardening in crystals.

Consider two straight dislocation lines A and B, as in Fig. 17 (a). Suppose that dislocation A glides in its plane and cuts through B, as

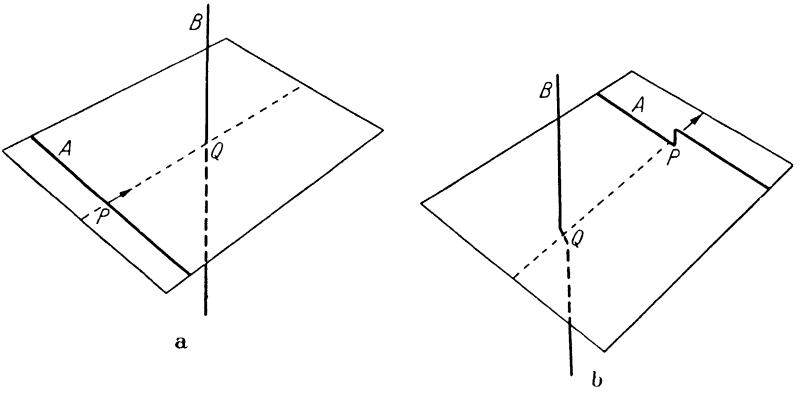


Fig. 17. The formation of dislocation jogs

in diagram (b). Then on each dislocation a *jog* will be formed, at P and Q. For instance, the part of B above the slip plane of A is moved, being carried along in the top half-crystal, relative to the part below the plane when the dislocation A passes through. The direction and magnitude of this displacement are equal to the Burgers vector of dislocation A. Since the Burgers vector must be constant along a single dislocation line, and since this line cannot end within the crystal (T.D.1, page 80), the displaced ends of dislocation B are necessarily connected by a short piece of dislocation, the jog, which lies in the slip plane of A and has the same Burgers vector as B. To summarize, a jog is a dislocation with the same Burgers vector as the dislocation line which contains it, while the length and direction of the line of the jog is Burgers vector of the dislocation that has cut through.

A jog will usually be one atom long. Now the energy of a dislocation is about μb^3 per atom plane (T.D.1, page 116), so that the energy of a jog is $\alpha\mu b^3$, where $\alpha \simeq 1$. To calculate the stress needed to make a jog consider the situation shown in Fig. 18, where a dislocation line is about to cut through a row of equidistant dislocations spaced l apart.

If this dislocation moves forward a distance of, say, b, i.e. one atom

spacing, it cuts through all the intersecting dislocations and makes $1/l$ jogs in unit length. The dislocation energy is increased by $\alpha\mu b^3/l$, but the work gained from the applied stress σ is σb^2. Equating these, the stress needed to make the jogs is

$$\alpha\mu b/l, \qquad \ldots . \quad (32)$$

so that if the density of intersecting dislocations is ρ, the stress needed to drive a dislocation through them is

$$\alpha\mu b\rho^{1/2}. \qquad \ldots . \quad (33)$$

Taking $\rho = 10^{12}$ cm^{-2} for a heavily worked material (T.D.1, page 104), and $b = 3 \times 10^{-8}$ cm, this is $\alpha\mu/50$. With $\alpha = 1$ (which may be an overestimate) this stress is about ten times greater than that predicted

Fig. 18. A dislocation cutting through a row of dislocations

from Taylor's theory of work hardening (T.D.1, page 106). Taylor's theory also leads to the relation (33) for the stress, but with an α of the order of 0·1.

If the applied stress is less than $\alpha\mu b/l$ the dislocation can still go forward with the aid of thermal fluctuations. The activation energy to make a jog is

$$U = \alpha\mu b^3 - \sigma b^2 l. \qquad \ldots . \quad (34)$$

There are two points about this: (*a*) since $\alpha\mu b^3$ is only a few electron volts (T.D.1, page 117) the activation energy is always fairly small, even at low stresses, so that there should be an appreciable temperature dependence for the hardening caused by intersecting dislocations; (*b*) the activation energy varies linearly with stress, a relation that is needed to explain some creep results described later; NABARRO[82] has shown that hardly any other process involving dislocations can give this linear relation. The significance of these effects for work hardening and creep will be discussed later.

We must now consider whether a jog is an impediment to the movement of the dislocation containing it. Jogs in edge dislocations should not have much effect. The interesting case, first pointed out by THORNTON READ and later discussed by SEITZ[39, 83] and MOTT,[84] is that where two screw dislocations cross each other. If a screw dislocation contains a jog that is perpendicular to the screw axis then this jog is a piece of edge dislocation, since its line and its Burgers vector are perpendicular to each other. These two directions define a plane, the slip plane of the edge dislocation, in which it can glide easily. But this implies a movement of the jog along the screw axis and not a movement

with the screw dislocation itself. Any movement of the screw, with the jog, in the direction perpendicular to the slip plane of the jog, involves either lengthening or shortening the half-plane (T.D.1, page 86) belonging to the edge dislocation in the jog; this can only occur by the jog leaving behind itself a row of vacant atomic sites, or of atoms misplaced into interstitial positions, as it moves along with the screw. If the temperature is high enough for self-diffusion, or if the jog is made with sudden violence, these vacancies or interstitial atoms will be dispersed into the surrounding crystal.

The purely geometrical basis of this effect can be understood with the aid of Fig. 19. Suppose in diagram (*a*) that slip is to occur along the lines *AB* and *CD*. Normally one would slip before the other, as in

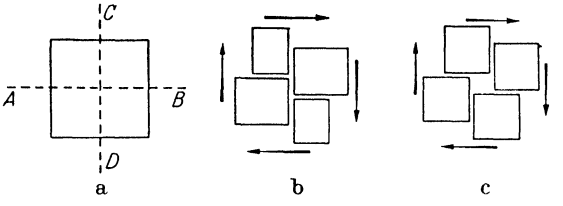


Fig. 19. Forming vacancies by slip on two planes

diagram (*b*), where *AB* has slipped before *CD*. However, if they both slip at the same time, and in the directions shown in diagram (*c*), a hole must be formed at the centre, where they cross. The essential condition to make the slip processes on these two planes overlap in time is that they occur by the movement of screw dislocations, one in each plane, in the crystal at the same time.

The generation of vacancies or interstitial atoms from a jog pulled along with a dislocation is of interest, and SEITZ[39, 83] has discussed this and some other processes for creating these defects. A jog in an edge dislocation may occasionally create a defect by a process illustrated in Fig. 20. Here *CDE* represents an edge dislocation with a jog at *D*, the half-plane being *ABCDE*, moving in the direction shown. The line *CD* can be regarded as an incomplete row of atoms attached to the lower edge of an otherwise perfect half-plane. The end of the row *D* is clearly a place where an atom is likely to become detached from the half-plane and left behind in an interstitial position as the dislocation moves on, or alternatively is picked up from the lattice and added to the row, in which case a vacant site is left behind. In close-packed metal lattices, where an interstitial atom has more energy than a vacancy, it is likely that the second of these processes will operate mostly, so that the incomplete row *CD* will gradually become longer and most of the lattice defects created will be vacancies.

SEITZ[39, 83] and MOTT[23] have summarized the experimental evidence

for the generation of lattice defects by cold work. The main items are as follows:

1. GYULAI and HARTLY[85] showed that the electrical conductivity of sodium chloride is increased about 100 times after 10 per cent plastic strain. This conductivity is believed to be due to vacancies and decreases on subsequent ageing at a rate consistent with this hypothesis.
2. MOLENAAR and AARTS[86] strained polycrystals of copper, silver, and aluminium, at the temperature of liquid air and measured

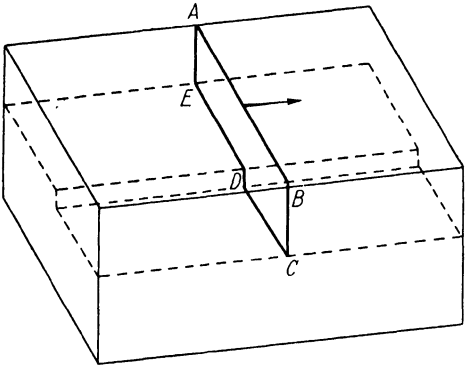


Fig. 20. A jog D in an edge dislocation CE

the increase of resistivity caused by this. Ageing at room temperature removed nearly one-half of this increase without altering the stress-strain curve. SEITZ believes that this recoverable part of the resistance is due to vacancies created during the cold working process, and that it disappears when these condense into aggregates or at jogs in edge dislocations. From the magnitude of the resistance change he estimates that 10 per cent plastic strain creates about 10^{18} vacancies per c.c.

3. DUGDALE[87] has shown that the resistivity of platinum is increased about 0·5 per cent by fast neutron bombardment, an effect which is almost certainly caused by the creation of vacancies and interstitial atoms. This recovers on annealing at the same rate as, and with the same activation energy (1·1 eV) as, the recoverable resistance caused by cold work.

THEORY OF THE YIELD POINT AND STRAIN AGEING

It was shown in T.D.1 that the stresses round a dislocation can be relieved by a suitable arrangement of nearby solute atoms, and that a solute atom with polar co-ordinates r, θ relative to a positive edge dislocation has an interaction energy

$$V = A \sin \theta / r \qquad \ldots . \quad (35)$$

with it. Here A is a constant that depends on the elastic properties of the material, on the degree of misfit of the solute atom, and on the strength of the dislocation. This formula breaks down at the centre of the dislocation, where Hooke's law does not hold, and it is expected that a large solute atom is most strongly bound at the position $\theta = 3\pi/2$ and $r(= r_0) \simeq 2 \times 10^{-8}$ cm which lies immediately below the half-plane of the dislocation. For carbon and nitrogen in α-iron one obtains[88] $A \simeq 3 \times 10^{-20}$ dyne cm^2 and, at the position of greatest binding, $V\ (= V_{\max}) \simeq 1\ eV$.

An approximate experimental value can be deduced from the work of DIJKSTRA.[89] Using an internal friction method developed by SNOEK[90] he measured the solubility of nitrogen in iron in the presence of numerous dislocations introduced by cold work. At 300°C this was 0·0025 weight per cent as compared with 0·012 weight per cent for unworked material, this latter being the solubility in the presence of iron nitride particles. According to DIJKSTRA the binding energy of nitrogen in iron nitride is 0·35 eV, so that $V_{\max}$ in dislocations is somewhat greater than this, say about 0·5 eV. Another method of estimating $V_{\max}$ which also gives 0·5 eV is discussed below.

This interaction can lead to two kinds of segregations of solute atoms round dislocations. The first is the dilute "atmosphere" or Maxwellian distribution, defined by

$$c = c_0 \exp(-V/kT), \qquad \ldots\ (36)$$

where c is the concentration in a region with a binding energy V and c_0 is the overall concentration. This expression is valid when c is small, so that the volume change produced by the segregation is small compared with that produced by the dislocation. As the temperature falls this atmosphere becomes more concentrated and below a critical temperature range it condenses into a line of solute atoms, lying along the dislocation line at the position of maximum binding.[88]

An estimate of the temperature T_0 below which condensation should occur (if diffusion is still possible) can be made as follows.[91] Suppose that there is no shortage of solute atoms for dislocations, i.e. that the number c_0N in unit volume of the crystal is large compared with ρ/b, the number of sites to be filled along the lines of maximum binding in the dislocations; here N is the total number of atoms in unit volume of crystal, ρ is the density of dislocations, and b is the atomic spacing. Since $\rho/bN \simeq \rho b^2$, which is about 10^{-3} when $\rho = 10^{12}$ cm^{-2} and 10^{-7} when $\rho = 10^8$ cm^{-2}, this condition is satisfied except in the most dilute solutions. Then under these circumstances equation (36) can be used to estimate T_0. At the point of maximum binding we have $V = V_{\max}$ and require that the temperature T_0 should be such as to give $c = 1$ at this point. Hence

$$T_0 = V_{\max}/k \log (1/c_0) \qquad (37)$$

MOTT[91] has used this to estimate V_{max}. He argues that the yield point of iron is not observed above about 700°K, and that this temperature should be identified with T_0 since a condensed atmosphere is needed for a strong yield point;[88] then by taking $c_0 \simeq 10^{-4}$ he concludes that $V_{max} \simeq 0{\cdot}5\ eV$ for carbon in iron.

The stress needed to pull a dislocation away from its atmosphere depends on whether this is condensed or not. With a condensed one the interaction energy per unit length, for a dislocation moved a distance x from its line of solute, is

$$V(x) = -Ar_0/(x^2 + r_0^2). \quad \ldots\ (38)$$

The corresponding force on the dislocation, $-dV/dx$, has a maximum at $x = r_0/\sqrt{3}$ equivalent to a (tensile) yield stress

$$\sigma_0 \simeq A/b^2r_0^2. \quad \ldots\ (39)$$

For carbon and nitrogen in iron this is 6×10^{10} dyne cm^{-2}. An extrapolation[88] of observed strengths to 0°K gives $1{\cdot}25 \times 10^{10}$ dyne cm^2. It is expected that the observed value should be lower, since yielding in iron is known to begin at places at high stress concentration, and to spread from these through the rest of the specimen in the form of Lüder's bands.

A rough estimate of the stress needed to pull a dislocation away from a Maxwellian distribution can be made as follows. The quantity A/kT $(= l)$ has the dimension of length and, since the density of the atmosphere is $\exp[-(l/r)\sin\theta]$, can be regarded crudely as the "radius" of the atmosphere. We expect that the force on the dislocation from its atmosphere is a maximum when the dislocation is displaced by a distance of about l from the centre. All the solute atoms in an area of order l^2 are then pulling it back. The force from one of these is A/r^2, and we shall take an average value $l/2$ for r. The number of solute atoms in the area is of order $l^2c_0N \exp(l/r) \simeq 7l^2c_0N$, per unit length of dislocation. The critical stress for the breakaway is then

$$\sigma_0 \simeq (1/b)(4A/l^2)(7l^2c_0N) \simeq 28Ac_0N/b. \quad \ldots\ (40)$$

A more accurate calculation[92] leads to the same result as this, but with a numerical factor of 17 in place of 28. Since N is of order $1/b^3$, or $1/br_0^2$, this critical stress is essentially the same as that of equation (39), but multiplied by a factor of order $10c_0$. The presence of c_0 means that in dilute solutions the yield stress for a Maxwellian atmosphere is far smaller than that for a condensed one. It is the absence of c_0 in equation (39) that makes the yield point for a condensed atmosphere so strong even in very dilute solutions. We suppose that the rather exceptional behaviour of iron has its origin here; that room temperature is well

below T_0 in this case but above T_0 for those metals which do not show the yield phenomenon.

A feature of the theory is that the yield point should depend strongly on temperature. There are three effects:

(*a*) As the temperature is increased through the range of T_0 the yield stress changes from the value given in equation (39) to that in equation (40).

(*b*) Because the atmosphere, and particularly the condensed one, is highly localized round the dislocation, only a small displacement of a small length of the dislocation line is needed to create a nucleus of

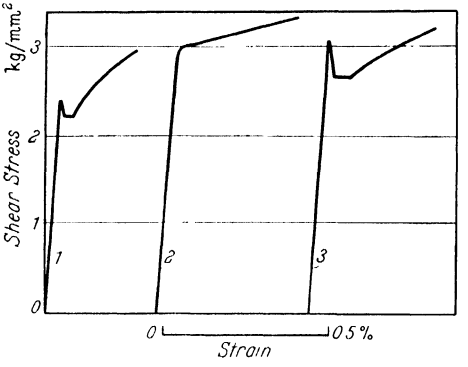


Fig. 21. Yielding in a single crystal of β-brass. (1) first test; (2) immediately reloaded; (3) after strain ageing for two hours at 250°C (after ARDLEY)

freed dislocation sufficient to enable the applied stress to drag away the rest of the dislocation; the energy needed for such a nucleus is small enough for thermal fluctuations to make an important contribution to it.[88]

(*c*) LOMER[93] has pointed out that all the available sites in a condensed atmosphere are not filled because the configurational entropy is increased if a few remain empty. These empty sites, which weaken the anchoring of the dislocation by the atmosphere, increase in number as the temperature is raised.

Fig. 21 shows an example of the yield phenomenon obtained in a single crystal of β-brass grown in nitrogen.[94] It shows the characteristic sharp fall from the upper to the lower yield point, the absence of this on immediate retesting when the material is *over-strained*, and the return of it on *strain ageing*. Several examples of materials showing the yield phenomenon are now known. Soft polycrystalline iron or mild steel, containing a little carbon and nitrogen, is the familiar one. Yield points have now been obtained in single crystals of iron containing carbon or nitrogen, in some cases during the first test made on them,[95, 96]

but in others only after strain-ageing treatments have been given.[22, 97] Examples amongst non-ferrous metals are polycrystalline molybdenum,[98] and single crystals of zinc[99, 100], cadmium[101, 102, 103] and α- and β-brass.[94] It is often not possible to demonstrate the existence of a yield point in a soft single crystal by means of the tensile test, because of difficulties in avoiding stress concentrations, but it can be proved by means of a bend test.[104]

The fact that, in iron, larger yield points are obtained in polycrystals than in single crystals, and in fine grained polycrystals than in coarse grained ones, has led some authorities to suggest that the yield point in iron is caused by films of hard material, probably rich in carbon, that encase the grains and prevent the soft metal inside them from flowing until, at the upper yield point, they break and release the plastic flow.[105] The difficulty with this view is to see how it can lead to an appreciable fall of stress at the yield point, and why in practice there is such a small amount of plastic deformation before the upper yield point is reached.[106, 107]

A qualitative explanation can be based on the dislocation theory,[107] as follows. There are three important stresses to be considered: (*a*) the observed upper yield stress σ_c, (*b*) the theoretical breakaway stress σ_0 (at the temperature of the experiment) for an anchored dislocation, (*c*) the applied stress σ_0/q at which dislocations begin to break away in regions of severe stress concentration q. Several experimental facts can be correlated[107] on the basis that $\sigma_0/q < \sigma_c < \sigma_0$. Yielding begins at a few places of high stress concentration, before the upper yield point is reached, and creates small regions in which plastic flow has begun while the rest of the specimen still remains elastic. The spread of these nuclei of yield is resisted by obstacles in the path of the freed dislocations in them, the most important obstacles being grain boundaries. The upper yield is the stress at which one of these nuclei first becomes able to spread into the rest of the material.

Starting from equation (35) for the interaction energy, a calculation can be made of the rate at which solute atoms segregate to a free dislocation during strain ageing. The expression obtained by COTTRELL and BILBY[88] for this is

$$n(t) = \alpha n_0 (ADt/kT)^{2/3}, \qquad \ldots \quad (41)$$

where $n(t)$ is the number of atoms segregated in time t to unit length of a dislocation in a region where the initial number of atoms in solution is n_0 per unit volume; D is the diffusion coefficient for these atoms and α is a number nearly equal to 3. The physical basis of this formula can be understood as follows. The gradient of the interaction energy provides a force attracting a solute atom to a dislocation and, migrating by thermal agitation, the atom acquires a drift velocity v proportional

to it. Since V varies as r^{-1}, where r is the distance from the dislocation, the gradient of V and the velocity v vary as r^{-2}. Thus atoms starting at a distance r from the dislocation take a time t proportional to r/v, i.e. to r^3, to reach it. The number that reach the dislocation in a time less than this is proportional to the initial number at all distances less than r, i.e. is proportional to r^2, and hence to $t^{2/3}$.

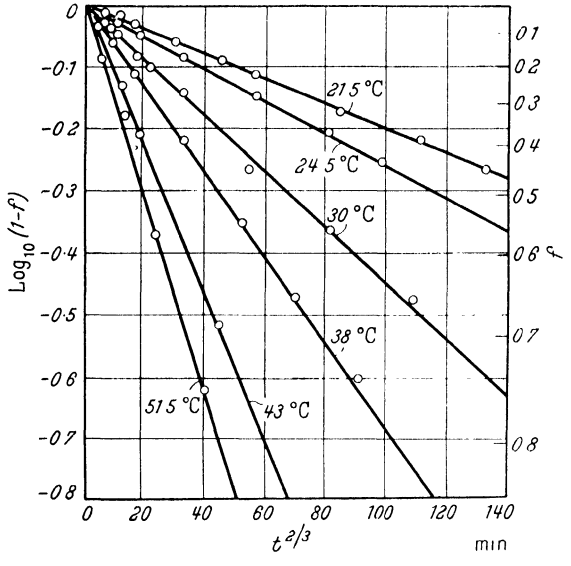


Fig. 22. The effect of temperature on the rate of strain ageing in iron containing carbon (after HARPER)

The actual rate of strain ageing should become slower during the later stages than equation (41) indicates because the dislocations become saturated with solute atoms. HARPER[108] made a simple generalization to include such effects by supposing that the rate of segregation decreases in proportion with the amount segregated. This gives for f, the fraction segregated,

$$f = 1 - \exp\,[-\alpha\rho(ADt/kT)^{2/3}], \qquad \ldots\ (42)$$

where ρ is the density of dislocations. Fig. 22, taken from HARPER'S internal friction measurements of the strain-ageing process, shows that, for the segregation of carbon in cold worked iron, $\log\,(1-f)$ varies linearly with $t^{2/3}$ as the theory requires. From the slopes of these lines the activation energy is found to be 20,000 cal/mol, in agreement with the value for diffusion of carbon in ferrite. By fitting equation (42) to the experimental results the density of dislocations can be found. The values HARPER obtained, e.g. $2{\cdot}5 \times 10^{11}$ cm^{-2} after 10 per cent strain,

agree with those of COTTRELL and CHURCHMAN,[109] who followed the strain-ageing process by an electrical resistance method.

These experiments show that the amount of carbon or nitrogen transferred to dislocations during strain ageing in moderately worked iron is of the order of 10^{-3} to 10^{-2} weight per cent. But the equilibrium solubility of these elements at room temperature is much smaller than this, of the order of 10^{-7} weight per cent for carbon and 10^{-5} weight per cent for nitrogen.[89, 110] This means that either the solution is supersaturated when the strain-ageing experiments are made or that some particles of carbide and nitride redissolve in order to feed the dislocations. Probably both of these occur in practice. It is known[89, 110] that precipitation in unworked iron becomes extremely slow when the amount left in solution is no more than 10^{-3} weight per cent, and that the metal remains supersaturated to this extent even after one month at room temperature. This is probably a result of the higher solution pressure of finely dispersed precipitates. A recent study[111] of these effects has shown that heat treatments designed to make it difficult for dislocations to obtain carbon and nitrogen (e.g. prolonged ageing before straining in order to make the precipitates coarse and reduce the supersaturation) can slow down the rate of strain ageing by an order of magnitude. It was found that the initial rate was closely proportional to the concentration of carbon and nitrogen in solution, as equation (41) requires.

WORK HARDENING

To construct a complete theory of work hardening three steps are necessary:

1. The adoption of some model to represent the structure of the cold worked crystal.

2. The calculation of the stress needed to enforce slip in such a structure.

3. The calculation of the strain needed to produce the structure. All three have to be taken if the stress-strain curve is to be calculated. The second step is the only one that is even moderately easy. Most theories have failed at the first one because the model adopted has not agreed with observed cold-worked structures. So far as the writer is aware, no one has ever accomplished the third step without the aid of an *ad hoc* assumption. Early theories, some of which were reviewed in T.D.1, achieved comprehensiveness by the adoption of a highly artificial model, but later ones have usually preferred to explain less in order to keep the model as realistic as possible.

The simple picture of slip is one of translation of thin parallel sheets relative to one another along a fixed direction. On such a picture it has always been a puzzle to understand why these sheets should not

slip right off one another once they have started to slide, i.e. why there should be any work hardening. The recent tendency is to uphold the primitive notion from which this difficulty springs; to believe that in this simple "laminar" slip there is in fact little or no work hardening and that the intense work hardening so often observed in practice is caused by more complex or "turbulent" slip processes involving inhomogeneous lattice rotations and slip on intersecting planes.

The main evidence for this view is the striking difference in the work hardening of crystals of hexagonal and cubic metals. Many shear stress-strain curves of crystals deformed in tension can be approximated in the plastic range by the relation

$$\gamma = (\sigma/h)^m, \qquad \ldots . \quad (43)$$

where σ is the increase in shear stress, above the initial yield value, needed to produce the shear strain γ, and where $m \simeq 1$ for hexagonal metals and 2 for cubic ones. The parameter h is the coefficient of work hardening; at room temperature it is about 4×10^8 dyne cm^{-2} for aluminium, 5×10^7 for magnesium, and $2{\cdot}5 \times 10^7$ for zinc. The difference between metals of cubic and hexagonal structures does not stop at this. With zinc and cadmium crystals the slip bands all appear early in the deformation, whereas with aluminium they increase in proportion with the applied stress. In a detailed study of tensile deformation HONEYCOMBE[112] showed that x-ray Laue asterisms are observed from aluminium crystals (except those oriented to deform in multiple slip) but not from cadmium ones. In these aluminium crystals neither the work hardening nor the asterisms could be eliminated except by annealing at temperatures where recrystallization occurred. The cadmium crystals, however, would not recrystallize and the small amount of work hardening that could be introduced in them by tensile straining was quickly removed by mild annealing.

Some of this evidence has been available for many years, but not until recently has it seemed feasible to draw the above conclusion from it. For it immediately leads to the following difficulty. Admittedly, a cubic crystal oriented to deform from the start by double slip may well behave differently from one of cadmium where double slip is not observed; but it is easy to obtain orientations where cubic crystals should deform at the start by single slip, and why do these behave differently from the hexagonal ones? TAYLOR[113] (see also references 35 and 114) showed that strain hardening in aluminium is more rapid when the crystal deforms from the start in double slip than in single slip, but this difference was only about 20 per cent and quite trivial when compared with the difference between cubic and hexagonal crystals.

Again, the recent tendency is to uphold the notion upon which the

difficulty rests; to say that when a cubic crystal is made to deform in pure single slip it does so like a hexagonal one and shows little hardening. However, just as it is difficult to prevent fast laminar flow in a fluid from breaking up into turbulent flow, so it is difficult, with a cubic crystal, to prevent laminar slip from breaking up into turbulent slip. It is not sufficient merely to choose an initial orientation suitable for single slip. Other steps are also necessary and if these are not taken the crystal generally does not behave in the simple way expected of it.

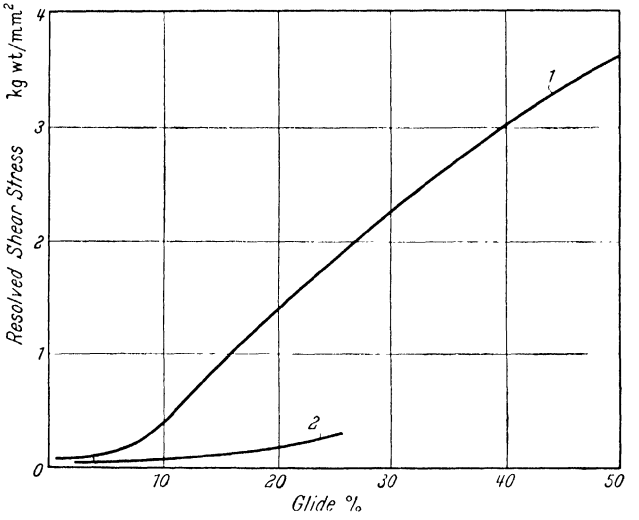


Fig. 23. Contrast between behaviour of a silver crystal (1) with a clean surface but exposed to air for 2 hours at room temperature, and (2) a similar crystal tested in argon. Both tested at room temperature (after ANDRADE and HENDERSON)

The main evidence for this view is the effect known as "easy glide"; under certain conditions the plastic part of the stress-strain curve of a face centred cubic crystal begins with a region where the hardening coefficient is abnormally small, or even zero, after which the curve rises steeply to give the intense work hardening typical of such crystals. The observations of this are as follows:

(*a*) ANDRADE and HENDERSON[115] observed such regions in crystals of pure gold and silver tested in tension, and concluded that easy glide is favoured by high purity, complete absence of surface contamination (e.g. oxide films), and a low temperature of deformation. Fig. 23 shows the remarkable effect that exposure to air has upon the work hardening of silver crystals.

(*b*) MADDIN, MATHEWSON and HIBBARD[35] showed that some single crystals of α-brass deform on a single set of slip planes with practically no hardening until the stage is reached where double slip begins, when

hardening becomes intense. Vacant atomic sites should be produced in large numbers, by the crossing of screw dislocations, once double slip begins and there is some indirect evidence for this in the observations of MASIMA and SACHS,[116] who showed that during single slip the change in electrical resistance and density was small but became substantial when double slip began.

(*c*) Crystals of aluminium stretched in tension show easy glide provided they are pure and have orientations which avoid double slip.[117, 118]

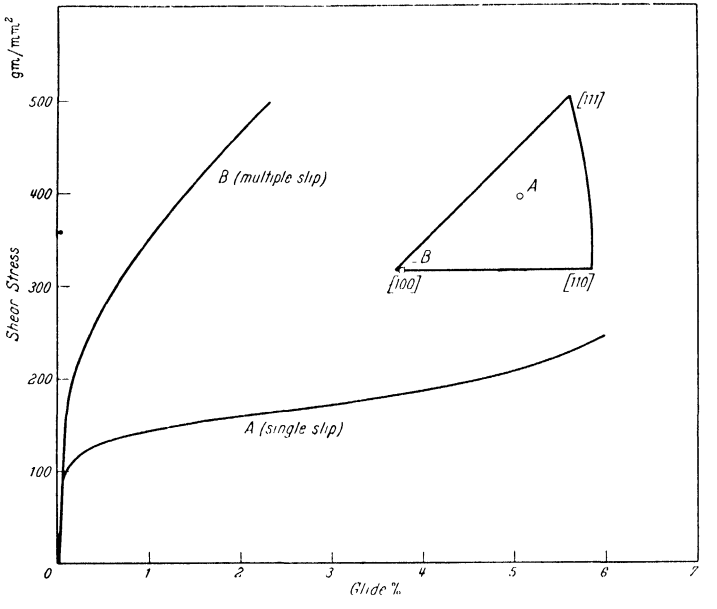


Fig. 24. Stress-strain curves on pure aluminium crystals (after LÜCKE and LANGE)

Fig. 24, taken from the work of LÜCKE and LANGE,[118] shows stress-strain curves on two crystals of pure aluminium (99·99 per cent), one of which (*A*) deformed initially by single slip and showed a region of easy glide in which hardening was very small, while the other (*B*) deformed by multiple slip and began hardening immediately. MASING and RAFFELSIEPER[117] showed that the region of easy glide became smaller the nearer the initial orientation was to one giving multiple slip, and that it was suppressed entirely in crystals of less pure aluminium (99·5 per cent); these gave roughly parabolic curves and intense work hardening for all initial orientations.

(*d*) In an attempt to avoid plastic bending at the grips of a specimen, RÖHM and KOCHENDÖRFER[119] strained long crystals of pure aluminium (99·99 per cent) in a special shearing apparatus. They found that this produced simple laminar slip which resembled slip in hexagonal metals

in that it gave a linear stress-strain relation, a small work hardening which could be removed easily by mild annealing, and no x-ray asterisms.

(*e*) PRATT[120] studied the deformation of sodium chloride crystals by means of special optical methods. Two kinds of stress-strain curves were obtained, steeply rising parabolic ones on crystals that deformed throughout on intersecting slip systems, and less steep and roughly linear ones on crystals in which, after a small initial period of double slip, only one slip system was active.

The impression gained from these results is that a cubic crystal oriented for single slip will start to deform in this ideal manner and give little work hardening, as would be expected. However, sooner or later, something happens that breaks up the flow into a more turbulent kind that gives intense work hardening. By taking certain precautions one may delay the onset of this second stage, or even postpone it, but if these are not taken it begins so soon that the "easy glide" characteristic of ideal single slip is not observed. In some cases the onset of turbulent flow is delayed until double slip clearly begins (α-brass) but this is not true of aluminium stretched in tension, where it usually begins within the first few per cent extension; one cannot be sure, of course, that small amounts of slip on "unpredicted" planes do not occur during this stage. However, the structural feature which appears to be connected with the change in the nature of the flow in this case is the *deformation* band or *kink* band. HONEYCOMBE[112] showed that these begin to appear after 1 per cent extension in crystals that deform by single slip, but not in those that deform by double slip.

Detailed studies of deformation bands have been made by CAHN,[37] HONEYCOMBE,[112] CHEN and MATHEWSON,[121] and LALOEUF and CRUSSARD.[122] Fig. 25 is an example of a band. It consists of a thin sheet extending across the crystal. At small strains the sheet is plane and normal to the active slip direction. It forms an obstacle to slip and many slip bands fail to penetrate it, the piled up dislocations in them producing the characteristic double bend of the lattice illustrated in Fig. 26. The intense stresses from these cause slip to occur on other slip systems in the band, and these intersecting slip bands become clearly visible after strains of about 20 per cent. The spacing between deformation bands is commonly about 10^{-2} to 10^{-1} cm and the change in orientation across a band is sufficient to account for the x-ray Laue asterisms observed from crystals containing bands. Crystals which deform entirely in multiple slip develop the bands either not at all or on too fine a scale to be seen.

The process by which the bands form is still obscure. The two main views are that they start from bend planes which run into the crystal from the edges, having been formed by the lattice bending caused by

grip constraints,[37, 112, 120, 121] or that they grow by the successive trapping of dislocations moving in slip planes near to one containing an obstacle against which dislocations have piled up.[23, 84] Deformation bands bear a geometrical resemblance to the *kinks* produced in hexagonal metal crystals by compression along an axis in the basal plane. FRANK[123] has suggested that macroscopic kinks can grow from short ones because the stress concentration at the end of an incomplete wall of edge dislocations can rise sufficiently to create new dislocations in the perfect lattice beyond it. A wall of edge dislocations which extends part of the way through a crystal is equivalent to a wedge, of atomic sharpness at

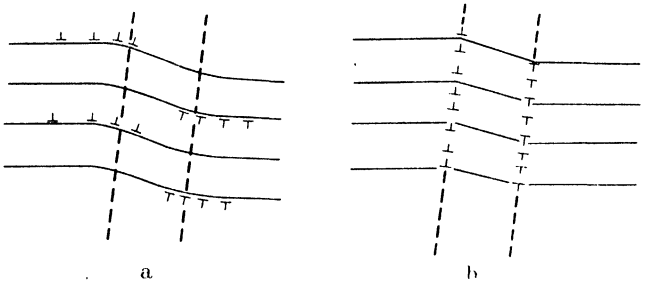


Fig. 26. Piling of dislocations against a deformation band (*a*) before and (*b*) after polygonization

its edge and composed of the half-planes of these dislocations, which is driven into the crystal.

The above discussion points to the conclusion that strong work hardening in cubic metals originates in the interactions among dislocations on intersecting slip bands. The arguments for this are as follows:

(*a*) This kind of hardening is only observed in crystals capable of slipping on intersecting planes.

(*b*) It is only observed in these crystals under circumstances that favour slip on intersecting systems, i.e. when the orientation is suitable for multiple slip, or when deformation bands are present.

(*c*) It does not seem possible to achieve the stability which the relative permanence of the cold worked state requires by means of purely Hookean interactions amongst slip dislocations in parallel planes. A dislocation in an array such as those considered by TAYLOR and KOEHLER (T.D.1, page 105) is at rest only so long as the net shear stress acting on it is zero. It is difficult to design arrays for which this stress is precisely zero on all dislocations in both the presence and absence of applied stress, and most arrays that one might expect to form turn out on analysis to be mechanically unstable unless some of their dislocations are permanently fixed in position and not merely allowed to rest because

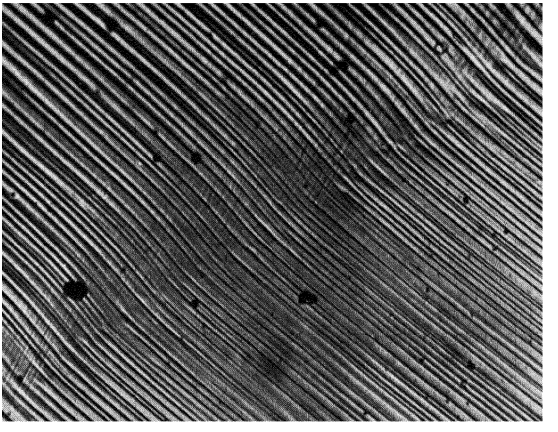


Fig. 25. A deformation band in aluminium: crystal extended 20 per cent (after CAHN) × 250

no shear stress acts on them. It is not necessary that all dislocations should be anchored, and the elastic after-effect described earlier shows that some of them are in fact mobile. What is required is that some dislocations in key positions should become incapable of moving and then act as obstacles confining the other dislocations, which are still mobile, to move within restricted areas of their slip planes.

Two mechanisms involving slip on intersecting planes which would provide such locking have been suggested. The first is that of LOMER and COTTRELL, described earlier, in which slip dislocations are turned into sessile ones by combination with slip dislocations on intersecting planes. The other mechanism, also described earlier, is that which leads to the formation of jogs. A dislocation may be able to glide freely in one slip plane but nevertheless will be an obstacle to dislocations moving on intersecting planes if, to be carried along with these, it has to glide out of its plane. A row of such dislocations, as shown in Fig. 18, holding up slip in an active plane provides a suitable obstacle for work hardening. Furthermore, as we have noted earlier, the hardening caused by such obstacles is sensitive to temperature, and this is observed in practice. The density of intersecting dislocations needed to account for the observed hardening in heavily strained crystals is in the range 10^{10} to 10^{12} cm^{-2}, and there is some independent evidence for such densities (T.D.1, page 104, also references 108 and 109).

Hardening due to the mutual interference of dislocations in different slip planes should, of course, be most severe in crystals oriented for multiple slip. However, it should not be thought that this kind of hardening is absent in a crystal which shows only one set of visible slip bands. What matters is the density of dislocations in intersecting slip planes and this is not the same as the total amount of slip on these planes. This point of view is brought out in a theory which is presented below. Even with an orientation suited for double slip there is a tendency to develop a patchwork grouping of regions in which slip is predominantly on one plane, as HONEYCOMBE[112] showed with his "bands of secondary slip." Such a clustering of parallel slip planes into packets is to be expected since the array of similar dislocations built up in a cluster is difficult to penetrate by dislocations approaching along a different family of slip planes.

The ideas outlined above provide the elements of a theory of hardening for turbulent slip, but not for laminar slip, the theory of which has still to be developed. Various suggestions have been put forward by KUHLMANN,[124] SEITZ,[83] OROWAN,[21] and MOTT[23, 84] which may prove important for this. An attempt to develop the theory of turbulent hardening quantitatively has been made by MOTT,[23] whose theory will now be described.

The theory is two-dimensional and based on the piling up of edge

dislocations against obstacles spaced at intervals of $2L$ along their slip planes. At each obstacle a group of n dislocations of the same sign is piled up and anchored by combination with a few dislocations on intersecting planes which have been forced into activity to relieve the high stresses round the group. Any slip on secondary planes other than these small and highly localized movements is not essential to the theory. Each group acts at large distances like a "super-dislocation" of strength nb, and its stress field reaches n times as far, at the same intensity, as that of a unit dislocation. This is held to account for the wide spacing between the slip bands, such as is observed in aluminium.[26] The picture of work hardening which emerges is similar to that of TAYLOR (T.D.1) but with the following main differences: (*a*) the dislocations are anchored, (*b*) the strength of each dislocation is n times larger, (*c*) the spacing of them is correspondingly larger, (*d*) the dislocations are not necessarily arranged in any regular array. Let the distance between the active slip planes be d. Then the density of the "super-dislocations" is $(Ld)^{-1}$, and the mean distance between them is $(Ld)^{1/2}$. Each exists in a field of internal stress due to its neighbours given approximately by

$$\sigma_i = \frac{\mu(nb)}{2\pi(Ld)^{1/2}}. \qquad \ldots . \quad (44)$$

The plastic strain is $\rho(nb)L$, or, substituting $\rho = (Ld)^{-1}$,

$$\gamma = nb/d. \qquad \ldots . \quad (45)$$

Eliminating d from these relations gives a parabolic relation between σ_i and γ,

$$\sigma_i = (\mu/2\pi)(n\gamma b/L)^{1/2}, \qquad \ldots . \quad (46)$$

similar to that of TAYLOR's theory (T.D.1, equation (23)).

To eliminate the somewhat arbitary parameter L in this relation, MOTT introduces his idea of the dynamic generation of loops at a Frank–Read source. The first time a source of length l is made to work, a stress $\Delta\sigma \simeq \mu b/l$ is needed, in addition to the frictional stresses due to impurities etc, but thereafter the source works spontaneously until the back-stress on it has cancelled this stress. The number n of dislocations needed to do this from a distance L is given by $\Delta\sigma \simeq n\mu b/2\pi L$, which implies, since $\Delta\sigma \simeq \mu b/l$, that $n/L \simeq 2\pi/l$. Substituting this in the stress-strain relation gives

$$\sigma_i = \mu(\gamma b/2\pi l)^{1/2}. \qquad \ldots . \quad (47)$$

MOTT emphasizes three things about this result:

(*a*) When l is constant it gives a parabolic relation between stress and strain, as is observed.

(*b*) When $l \simeq 10^{-4}$ cm, which is reasonable, the coefficient of work hardening is of the right order of magnitude.

(*c*) The stress-strain relation does not depend on L, the spacing of the obstacles. This is important since the intensity of turbulent work hardening is roughly the same in crystals of different orientations.

To account for the dependence of work hardening on temperature, MOTT takes up an idea due to BROWN,[26] and suggests that vacancies created in an active slip plane during plastic flow migrate into neighbouring planes, up to 100 atoms away, and soften these sufficiently to cause them to start slipping. This is also held to account for the stepped

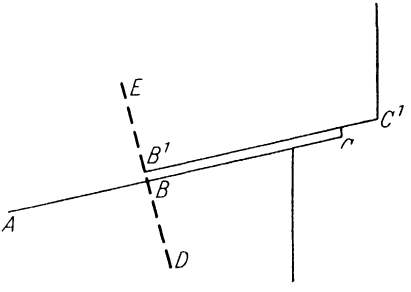


Fig. 27. Formation of steps on a slip band

structure of slip bands observed by HEIDENREICH and SHOCKLEY,[25] and BROWN.[26] However, alternative explanations are possible. We have seen that the hardening caused by intersecting dislocations is sensitive to temperature. As regards the formation of steps in slip bands, the suggestion has been put forward independently by FISHER[24] and the writer that these are caused by the displacement of outlying parts of active slip planes, relative to the central ones, as a result of slip on intersecting planes. This process is illustrated in Fig. 27. Here a slip band ABC emerges on the surface at C. At an earlier stage in the deformation $B'C'$ was the continuation of the line AB of the band, but this part has now been moved up by slip on an intersecting plane DE. Since $B'C'$ is now cut off from the Frank–Read source generating the band no more slip can take place on this step. This effect of the cessation of slip on existing steps, when new ones are formed, has been observed by BROWN,[125] who calls it "sterile slip."

ANNEALING

In T.D.1 a theory was given which explained why polygonization occurs in a plastically bent crystal, and why polygon boundaries are perpendicular to the active slip direction, from the elastic properties of edge dislocations. It did not explain how polygonization occurs, however. Two points are important here: (*a*) in the freshly bent crystal the

dislocations are grouped "horizontally" in slip planes, whereas after polygonization they are grouped "vertically" in walls; to produce this change they must climb out of their slip planes besides gliding along them; (*b*) polygonization only occurs rapidly at temperatures where self-diffusion is also rapid. As MOTT[84] has pointed out, this means that polygonization must occur by the addition or removal of matter at jogs on the edges of the half-planes by atomic migration. The rate at which a jog gives off a vacancy can be written as $c \exp(-W/kT)$, where W is the activation energy for self-diffusion and where the frequency factor c contains a term proportional to the force pushing the edge dislocation out of its slip plane.[84] If n/b is the number of jogs per unit length, where b is the atomic spacing along the edge, the velocity with which the dislocation climbs away from its slip plane is

$$nbc \exp(-W/kT). \quad \ldots . \quad (48)$$

As it climbs, the dislocation has to be replenished with jogs since the existing ones run to the end of the line. New jogs can be made by the action of thermal fluctuations, in which case the equilibrium density of jogs along the edge is given by the relation $n = \exp(-U/kT)$, where U is the work of formation of a jog; the activation energy for climbing is then increased to $U + W$, and polygonization cannot occur rapidly except at high temperatures. Alternatively, the dislocation may glide slowly under an applied stress in its slip direction, while it is climbing and so generate its own jogs by cutting through intersecting screw dislocations. In this case the activation energy for climbing is only W, not $U + W$, and polygonization occurs at lower temperatures. This is perhaps why cell-structures, very reminiscent of complex polygonization boundaries, are formed during slow creep in polycrystals at temperatures much lower (e.g. 250°C instead of 600°C) than are necessary to produce them by annealing after cold work.[126] The fact that these cell-structures have crystallographically irregular boundaries is not an argument against their being formed by polygonization, since the mutual constraints of neighbouring crystals force slip to occur by dislocations with more than one kind of Burgers vector; SHOCKLEY and READ[55] have suggested, and LOMER and NYE[65] have confirmed with the bubble raft, that the equilibrium orientation of a small-angle boundary is always perpendicular to the combined Burgers vector of the dislocations in it.

Apart from this progress in the theory of polygonization, the development of a structural theory of recovery has been held up while the nature of the cold-working process is being clarified. In the meantime, a formal theory of recovery has been developed by KUHLMANN[81, 124] and COTTRELL and AYTEKIN,[127] in which the basic idea is that the rate, $-\partial\sigma/\partial t$, at which the yield stress σ recovers on annealing is controlled

by an activation energy, $Q - \beta\sigma$, that depends linearly on the yield stress. An analysis similar to that used to derive equation (30) then shows that the yield stress decreases logarithmically with the annealing time, a relation which has been confirmed experimentally.[127, 128, 129]

Recrystallization and grain growth both occur by the migration of a grain boundary in a direction normal to itself. The driving force in the first process is the difference in energy between the cold worked and annealed grains which meet at the boundary; in the second one it is the surface tension of the boundary. The velocity with which a boundary moves depends, of course, on the driving force, as the theories of MOTT[130] and TURNBULL[131] show, but evidence has been accumulating recently that the enormous range in the mobilities of different grain boundaries is due, not so much to different driving forces, but to differences in orientation. It is convenient here to distinguish four types of boundary:

(1) The small-angle (e.g. $\theta < 1°$) simple Burgers boundary (Fig. 13), containing one family of parallel and uncontracted edge dislocations (see page 227).

(2) As (1), but with larger change of angle (e.g. $5° < \theta < 20°$) and contracted dislocations.

(3) The small-angle or medium-angle unsymmetrical boundary containing dislocations with more than one kind of Burgers vector.

(4) The large-angle (e.g. $\theta \simeq 30°$) incoherent boundary across which there is no continuity along lattice rows, and which may be imagined in terms of a monomolecular layer of "liquid" between the crystals.

The high mobility of type (1) boundaries has been demonstrated in the bend plane experiments of PARKER and WASHBURN,[58] in the bubble raft experiments of LOMER and NYE,[65] and in the observations of GUINIER and TENNEVIN,[132] and DUNN and DANIELS,[133] on polygonization. Type (2) boundaries, on the other hand, do not appear to be very mobile, as shown by the fact that the growth of polygonal domains becomes much slower when the angle between them increases.[133] LOMER and NYE[65] have argued that boundaries made from contracted dislocations should be immobile because the Peierls–Nabarro force (T.D.1, page 90) resisting the motion of a dislocation increases rapidly as the width of the dislocation is reduced.[66]

Boundaries of type (3) are often found in worked polycrystals[126, 134, 135] and there is much evidence for their immobility. TIEDEMA, MAY and BURGERS,[136] and BECK, SPERRY and HU[137] showed that small-angle boundaries between strained crystals and adjacent strain-free ones of almost identical orientation had very low mobilities; the presence of small unconsumed grains in single crystals grown by the strain-anneal method is due to this effect. Similarly, the domain structures observed by KELLAR, HIRSCH and THORP,[134] and by HEIDENREICH,[135] in cold worked polycrystalline aluminium do not coarsen on prolonged ageing

even though they form practically instantaneously during working and are on a very fine scale ($\simeq 2 \times 10^{-4}$ cm). The difficulty of moving such boundaries lies in the facts that the glide motions of their dislocations lie along different directions, and that the climbing motions of these dislocations involve diffusion through the regions of "good crystal" which lie between them.[55, 65]

The mobility of large-angle boundaries is shown by the well-known observation that, in recrystallization and grain growth, the new crystals have radically different orientations from those they replace. It seems that the development of annealing textures has its origin here. BECK,

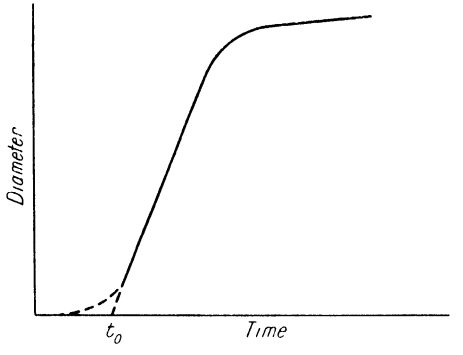


Fig. 28. Increase in the size of a growing crystal during recrystallization

SPERRY and HU,[137] for example, showed that the new crystals which grew fastest into a worked crystal were rotated, relative to its orientation, by 30 to 40°, about a [111] axis. A large-angle boundary should move rapidly by the transference of atoms, singly, from one crystal to the other, since each atom in such a boundary is next to a defect similar to a vacant atomic site. The movement bears a resemblance to that of diffusion in a grain boundary and, as LOMER and NYE[65] point out, this is consistent with the observation of ACHTER and SMOLUCHOWSKI[138] that grain boundary diffusion becomes much faster when the boundary angle is increased beyond 20°.

This dependence of mobility on orientation can be used to explain another effect in recrystallization. It is known that recrystallization starts slowly,[139, 140] and that measurements of the linear growth of a new grain give a curve of the kind shown in Fig. 28. Over a large range the growth increases linearly with time and the rate only falls away from this constant value when the new grains impinge on each other. However, if the linear part is extrapolated back linearly to zero diameter it meets the time axis at some positive value t_0, and this has led to the idea that recrystallization is preceded by an earlier stage when

nucleation of new crystals occurs. In the writer's opinion this extrapolation is misleading; there does not seem to be any certain evidence that distinctly different processes occur during "nucleation" and growth. A reasonable view is that of BURGERS and LOUWERSE,[141] and CAHN,[142] who suppose that the new grains are formed by the growth of initially small pieces of the old ones. Now the boundary dividing a growing piece from its surroundings is a small-angle one at first, and so its velocity of movement will be smaller than at a later stage of growth, where the growing crystal will have advanced into more remote regions of the crystal, regions which have radically different orientations from its own and which make large-angle boundaries with it.

On this view, the slow start of the recrystallization process is explained by a slowness in the velocity of boundary migration during the early stages of growth; thus the linear curve ought not to be extrapolated straight back to t_0 in Fig. 28, but curved back into the origin, as shown by the alternative extrapolation. This view also resolves a difficulty in BURGERS' and CAHN'S theories, which is to understand why re-crystallization begins in the most intensely curved regions of the crystal, in spite of the fact that they are the most heavily strained regions and the driving force for growth should be least favourable there. We suppose that growth starts from them because only there is the orientation so widely different from average that a grain growing out of one of them can obtain for itself a large-angle boundary; unless it can do this it cannot grow at a sufficient speed. This also helps to explain why crystals deformed in laminar flow do not recrystallize, on the grounds that the orientation in them does not vary widely enough to allow large-angle boundaries to be made; similarly it explains why recrystallization starts so readily from a scratch or similar sharp surface indentation. As CAHN[142] has already pointed out, the early stages of recrystallization should be recognizable as polygonization, and TIEDEMA[143] has recently obtained some direct evidence for this.

CREEP

The confirmation of some of the general predictions of the exhaustion theory of creep, described in T.D.1, has led to further investigations with the result that the theory has been both broadened in scope and simplified in treatment. We shall now sketch a version of it based on the following assumptions:

(1) The flowing specimen can be regarded as an aggregate of elements; in many but not all creep problems the elements are dislocation lines.

(2) Each element has an activation stress σ_a (= yield stress of the element σ_y − applied stress σ) which has to be applied before it flows.

(3) With this activation stress is associated an activation energy $U(\sigma_a, \sigma)$ which is an increasing function of σ_a.

(4) When an element is thermally activated by a temporary acquisition of the energy U, it makes a *jump*, i.e. contributes an increment v to the overall plastic strain.

(5) $U(\sigma_a, \sigma)$ is the same function of σ_a for all elements; in particular we shall make the assumption

$$U(\sigma_a) = A\sigma_a, \qquad \ldots . \quad (49)$$

which is consistent with activation energy function of equation (34). This function is appropriate when the hardening responsible for σ_a is caused by intersecting dislocations; with other kinds of hardening other functions (e.g. $U \propto \sigma_a^{3/2}$ for precipitation hardening; see T.D.1) may be better.

(6) The increment of strain v is the same for all jumps.

(7) The elements are independent of each other and cannot jump more than once.

Let $N(\sigma_a, t)d\sigma_a$ be the number of elements with activation stresses between σ_a and $\sigma_a + d\sigma_a$ at time t. Then the rate of creep at this time is given by

$$\dot{\gamma}(t) = v\int_0^\infty N(\sigma_a, t)f(\sigma_a)d\sigma_a, \qquad \ldots . \quad (50)$$

where

$$f(\sigma_a) = \nu \exp\{-U(\sigma_a)/kT\} \qquad \ldots . \quad (51)$$

is the frequency of jumping per element, and ν is a frequency factor. Since each element jumps only once, the numbers of unused elements are exhausted at a rate

$$\frac{\partial}{\partial t} N(\sigma_a, t) = -N(\sigma_a, t)f(\sigma_a), \qquad \ldots . \quad (52)$$

which integrates to give the time dependence of $N(\sigma_a, t)$ as

$$N(\sigma_a, t) = N(\sigma_a, 0) \exp\{-f(\sigma_a)t\}, \qquad \ldots . \quad (53)$$

where $N(\sigma_a, 0)$ is the initial distribution.

With a suitable assumption for $N(\sigma_a, 0)$, e.g.

$$\begin{aligned} N(\sigma_a, 0) &= N \ (0 \leqslant \sigma_a \leqslant \sigma_m) \\ &= 0 \ \ (\sigma_m < \sigma_a) \end{aligned} \qquad \ldots . \quad (54)$$

and with the energy function of equation (49), the integral in equation (50) for the creep rate is elementary. However we shall use another method which is only approximate but has greater generality and is more suggestive physically. With reasonable values for the physical parameters involved, the function $\exp\{-f(\sigma_a)t\}$ is practically a step function of σ_a when $\nu t \gg 1$. For small values of σ_a it is nearly zero, and for large ones nearly unity. Only in a very small range of σ_a, centred

about the value for which $f(\sigma_a)t = 1$, does it change from the one extreme to the other. Thus the exhaustion process can be regarded as the advance of a sharp step or thermal front along the σ_a axis. As the step sweeps across each part of the $N(\sigma_a, 0)$ curve all the elements in that part are activated and make their contribution to the creep. In a time t the step arrives at the activation stress σ_t defined by

$$U(\sigma_t) = kT \log \nu t, \qquad \ldots . \quad (55)$$

or, if equation (49) is used,

$$\sigma_t = (kT/A) \log \nu t. \quad (\nu t \gg 1.) \qquad \ldots . \quad (56)$$

The creep strain at time t is then

$$\gamma(t) = v\int_0^{\sigma_t} N(\sigma_a, 0) d\sigma_a, \qquad \ldots . \quad (57)$$

or, with $N(\sigma_a, 0) = N$,

$$\gamma(t) = (vNkT/A) \log \nu t. \quad (\nu t \gg 1.) \qquad \ldots . \quad (58)$$

At time t the stress

$$\sigma_y = \sigma_t + \sigma \qquad \ldots . \quad (59)$$

is the "yield stress" of the material in the sense that, if the applied stress could be raised to this value, flow could occur without the aid of thermal fluctuations. By making sudden small increments of applied stress during a creep experiment the presence and position of this thermal front can be examined.

Two investigations have been made recently to check the exhaustion theory. DAVIS and THOMPSON[144] used a precipitation-hardened copper-silver alloy and showed that their creep curves obeyed a time law of the kind $\gamma \propto (\log \nu t)^{2/3}$, which is consistent with the exhaustion theory if the activation energy function $U \propto \sigma_a^{3/2}$ appropriate to a precipitation-hardened alloy is used. By means of incremental loading tests they also demonstrated the existence of a critical stress, i.e. the existence of the thermal front, which could be reached by a suitable load increment (e.g. an increment of 1·4 kg mm^{-2} after creep for 24 hours at room temperature under a constant applied stress of 6 to 8 kg mm^{-2}) and at which "instantaneous" plastic flow was produced. This investigation also showed that the magnitude of the activation energy expected from the precipitation-hardening theory (see T.D.1) was of the right order, but that the observed value ($\simeq 1$ sec^{-1}) of ν was far smaller than that expected ($\simeq 10^8$ sec^{-1}) for the vibration of a dislocation against a precipitate. This discrepancy has still not been satisfactorily explained.

The other investigation was that of WYATT,[145, 146] who used copper hardened by the working received during the creep process itself. At low temperatures (i.e. < 300°K) his specimens obeyed the simple

logarithmic creep law of equation (58), and he proved that, under these conditions, a *mechanical equation* of *state*

$$\dot{\gamma} = f(\sigma, \gamma, T) \qquad \ldots . \quad (60)$$

was obeyed. At constant temperature T, for example, this means that the rate of strain $\dot{\gamma}$ at any given instant depends only on the *current* values of the applied stress σ and strain γ at this instant, and not on past values. Essentially the method of proof was to make sudden increments of applied stress during the creep experiment and to show that the portions of the creep curve which followed such increments could always be fitted exactly on earlier portions of the curve by sliding them backwards by a suitable amount. From the distance along the $\log t$ axis that it was necessary to slide the curves back, WYATT concluded that the activation stress increases linearly with $\log t$, which implies a linear relation, equation (49), between activation energy and activation stress. A further analysis of WYATT'S results[147] shows that the activation energy function in these experiments is of the form

$$U(\sigma_a, \sigma) = B\sigma_a/\sigma, \qquad \ldots . \quad (61)$$

which is almost exactly what would be expected from the theory of intersecting dislocations; this can be seen by substituting $B = \alpha\mu b^3$, $(\alpha\mu b/l) - \sigma = \sigma_a$, and approximating $\sigma + \sigma_a \simeq \sigma$, in equation (34). Moreover the value of B deduced from the experiments[147] is about 4 eV which is a very reasonable magnitude for $\alpha\mu b^3$. These experiments of WYATT therefore lend good support to the idea that work hardening is caused by intersecting dislocations having to cut through one another.

At higher temperatures and larger creep strains logarithmic creep usually gives way to a combination of ANDRADE'S $t^{1/3}$ creep and steady-state creep. WYATT has shown that a mechanical equation of state is not obeyed during $t^{1/3}$ creep, so that any change made in the exhaustion theory to enable it to deal with these other forms of creep would have to be a profound one. Such a change has been suggested recently,[147] based on the idea that once an element has jumped, it is not necessarily unable to jump again, but faces an energy barrier which is uncorrelated with the one it has jumped and is chosen at random from the whole range of barriers. This idea is suggested by the model of a dislocation meeting a random sequence of obstacles during its glide motion through a slip plane. It leads to a creep function which contains a steady-state term and has a transient term rather similar to, but not identical with, the $t^{1/3}$ function. Except at extremely small strains, where the creep function approximates to the logarithmic type, a mechanical equation of state is not obeyed.

A formal theory of steady-state creep, the so-called recovery theory, can be built up on the idea that the creep rate remains constant because

work hardening and thermal softening are balanced against each other during this kind of flow.[148, 149, 127] Thus OROWAN[149] defines a coefficient of work hardening, $h = \partial\sigma/\partial\gamma$, and a rate of recovery, $r = -\partial\sigma/\partial t$. During steady-state creep the yield stress remains constant,

$$d\sigma = \frac{\partial\sigma}{\partial\gamma}\,d\gamma + \frac{\partial\sigma}{\partial t}\,dt = 0,$$

so that the rate of steady flow is given by

$$\kappa = \dot{\gamma} = r/h. \qquad \ldots\ (62)$$

A consequence of the theory is that if the applied stress is suddenly reduced by a small amount, during steady flow, there should be a pause in the flow before the new steady rate is taken up, since the yield stress has first to fall by recovery by the amount that the applied stress was reduced. The existence of these induction periods has now been confirmed.[127, 150] COTTRELL and AYTEKIN,[127] for example, showed with zinc crystals that the observed time of induction can be accounted for approximately, with the aid of equation (62), in terms of the reduction in stress, the rate of steady flow, and the coefficient of work hardening. They also observed that the dependence of κ on applied stress and temperature could be represented by the relation

$$\kappa = C \exp\{-(Q - \beta\sigma)/kT\} \qquad \ldots\ (63)$$

where the parameters Q and β, when measured in a standard range of σ, had the same values as in recovery experiments on the same material. Since equation (62) allows us to write $-(1/h)(\partial\sigma/\partial t)$ for κ, equation (63) is essentially the differential equation for recovery which, as mentioned on page 253, integrates to give the logarithmic time law of softening.

The explanation of the final stage of the creep curve, in which the rate of flow increases progressively to fracture, remains an unsolved problem. Various suggestions have been put forward invoking effects due, for example, to changes in cross-section, opening of cracks in heavily worked slip planes and near grain boundaries, recrystallization, and metallurgical phase changes; accounts of these have been given in recent books.[151] However, there is no chance of making a detailed quantitative or even qualitative, theory of the late stages of creep until we have a more intimate picture of the various structural changes which occur during a creep experiment, which accounts for the great contemporary interest in metallographic studies of creep.[126, 127, 152]

Two features which have claimed most attention are the cell-structures formed within the grains, and the sliding of grains past one another along their boundaries. The nature of cell-structures and their relation to polygonization have already been discussed on page 252. Slip along grain boundaries has been studied directly by observing the

sliding of one crystal over another in a bicrystal,[153] and indirectly by means of internal friction measurements.[154, 155] The advantage of this second method is that the movements involved are very small so that extraneous effects, due to interlocking of crystals at places where the boundaries change their direction, are reduced. On the other hand, because it is an indirect method of study, the interpretation of its results is less certain. Theories have been developed, for example, in which the grain boundaries are considered to act either as surfaces along which slip occurs,[130, 149, 154] rather as if they consisted of monatomic layers of liquid, or as sources and sinks of vacancies, in which case creep is caused by the migration of these vacancies through the grains from one boundary to another.[156, 157, 158]

Direct observation of displacements at grain boundaries[159] and of plastic deformation[160] and cracks[159] in grains at the ends of such boundaries leave no doubt that grain boundary slip is a most important feature of creep in polycrystals at high temperatures. These boundary movements cannot occur without some sympathetic plastic deformation at the corners where grains interlock and the final rate of flow must depend on this deformation within the grains, even although it is started by slip along the boundaries. The nature of the interaction of slip on the boundaries with flow in the grains is the problem which must be solved if the accelerating stage of creep is to be properly understood, or if the theory of creep is to be of practical value in the design of creep-resistant alloys. Here, as is usual in these more complicated problems of crystal plasticity, one looks to metallographic examination to provide the vital clues to inspire the theory, and there are signs from the recent work[126, 152, 160] that they will not be long forthcoming.

References

[1] Cottrell, A. H. *Progress in Metal Physics I* (Ed: B. Chalmers), Chap. II, p. 77 (1949)

[2] Burton, W. K. and Cabrera, N *Disc. Faraday Soc.* 5 (1949) 33 40

[3] ——, Cabrera, N. and Frank, F. C. *Nature, Lond.* 163 (1949) 398; *Phil. Trans. Roy. Soc.* (A) 243 (1951) 299

[4] Frank, F. C. *Advances in Physics* 1 (1952) 91

[5] Volmer, M and Schültze, W. *Z. phys. Chem.* (A) 156 (1931) 1

[6] Griffin, L. J. *Phil. Mag.* 41 (1950) 196

[7] See, for example, Dawson, I. M. and Vand, V. *Nature, Lond.* 167 (1951) 476; *Proc. Roy. Soc.* (A) 206 (1951) 555; Verma, A. R. *Nature, Lond.* 167 (1951) 939; Amelinckz, S. *Nature, Lond.* 167 (1951) 939; Forty, A. J. *Phil. Mag.* 42 (1951) 670

[8] Forty, A. J. *Phil. Mag.* 43 (1952) 481; see also Amelinckz, S. *Phil. Mag* 43 (1952) 562

[9] Frank, F. C. and Read, W. T. *Pittsburgh Symposium on the Plastic Deformation of Crystalline Solids* Carnegie Institute of Technology and Office of Naval Research 1950, p. 44; *Phys. Rev.* 79 (1950) 722

[10] FISHER, J. C. *Private communication*; HOLLOMON, J. H. *Report of 9th Solvay Conference* (1952) p. 576
[11] BROWN, A. F. and HONEYCOMBE, R. W. K. *Phil. Mag.* 42 (1951) 1146
[12] FRANK, F. C. *Proc. Phys. Soc.* (A) 62 (1949) 131
[13] ESHELBY, J. D. *Proc. Phys. Soc.* (A) 62 (1949) 307
[14] LEIBFRIED, G. and DIETZE, H.-D. *Z. Phys.* 126 (1949) 790
[15] SEITZ, F. *Pittsburgh Symposium* (reference 9) p. 1
[16] ESHELBY, J. D. *Proc. Roy. Soc.* (A) 197 (1949) 396
[17] LEIBFRIED, G. *Z. Phys.* 127 (1950) 344
[18] NABARRO, F. R. N. (*a*) *Phil. Mag.* 41 (1950) 1270; (*b*) *Proc. Roy. Soc.* (A) 209 (1951) 279
[19] NABARRO, F. R. N. (*a*) *Private communication*; (*b*) *Advances in Physics* 1 (1952) 271
[20] LEIBFRIED, G. *Z. Phys.* 127 (1950) 580
[21] OROWAN, E. *Report of 9th Solvay Conference* (1952) p. 535
[22] HOLDEN, A. N. *J. appl. Phys.* 22 (1951) 1290
[23] MOTT, N. F. *Phil. Mag.* (1952). To be published
[24] FISHER, J. C., HART. E. W. and PRY, R. H. *Phys. Rev.* To be published
[25] HEIDENREICH, R. D. and SHOCKLEY, W. *Rep. Conf. on Strength of Solids* Physical Soc. London, 1948 p. 57
[26] BROWN, A. F. *J. Inst. Metals* 80 (1951) 115
[27] FRANK, F. C. *Phil. Mag.* 42 (1951) 809
[28] ——, *Physica* 15 (1949) 131
[29] CHALMERS, B. and MARTIUS, U. M. *Nature. Lond.* 167 (1951) 681
[30] THOMPSON, F. C. and MILLINGTON, W. E. W. *J. Iron and Steel Institute* 109 (1924) 67
[31] MATHEWSON, C. H. *Trans. Amer. Soc. Met.* 32 (1944) 38
[32] EDWARDS, O. S. and LIPSON, H. *Proc. Roy. Soc.* (A) 180 (1942) 268; *J. Inst. Met.* 69 (1943) 177
[33] BARRETT, C. S. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 123
[34] FULLMAN, R. L. *J. appl. Phys.* 22 (1951) 448
[35] MADDIN, R., MATHEWSON, C. H. and HIBBARD, W. R. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 86; 185 (1950) 527
[36] OGILVIE, G. J. and BOAS, W. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 102
[37] CAHN, R. W. *J. Inst. Metals* 79 (1951) 129
[38] NABARRO, F. R. N. *Rep. Conf. on Strength of Solids*, Physical Soc. London, (1948) p. 75
[39] SEITZ, F. *Phys. Rev.* 79 (1950) 723 890 1002 1003
[40] LOMER, W. M. *Phil. Mag.* 42 (1951) 1327
[41] COTTRELL, A. H. *Phil. Mag.* 43 (1952) 645
[42] KING, R. *Private communication*
[43] THOMPSON, N. and MILLARD, D. J. *Phil. Mag.* 43 (1952) 422
[44] COTTRELL, A. H. and BILBY, B. A. *Phil. Mag.* 42 (1951) 573
[45] MASON, W. P., MCSKIMIN, H. J. and SHOCKLEY, W. *Phys. Rev.* 73 (1948) 1213
[46] CHRISTIAN, J. W. *Proc. Roy. Soc.* (A) 206 (1951) 51
[47] COHEN, M., MACHLIN, E. S. and PARANJPE, V. G. *Thermodynamics in Metallurgy* Amer. Soc. Metals p. 242 1950
[48] FRANK, F. C. *Acta Metallurgica* (1953) To be published
[49] BOWLES, J. S. *Acta Cryst.* 4 (1951) 162
[50] LEIBFRIED, G. *Z. Phys.* 126 (1949) 781
[51] BILBY, B. A. *Proc. Phys. Soc.* (A) 63 (1950) 191
[52] PEACH, M. and KOEHLER, J. S. *Phys. Rev.* 80 (1950) 436

[53] NABARRO, F. R. N. *Phil. Mag.* 42 (1951) 213
[54] ESHELBY, J. D. *Phil. Trans. Roy. Soc.* (A) 244 (1951) 87
[55] SHOCKLEY, W. and READ, W. T. *Phys. Rev.* 75 (1949) 692; 78 (1950) 275
[56] LACOMBE, P. *Rep. Conf. on Strength of Solids* Physical Soc. p. 91 London (1948)
[57] JILLSON, D. C. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 1009
[58] PARKER, E. R. and WASHBURN, J. *University of California* Series No. 27, Issue No. 3 Oct. (1951)
[59] BURGERS, J. M. *Proc. Kon. Ned. Akad. Wet.* 42 (1939) 293
[60] VAN DER MERWE, J. H. *Proc. Phys. Soc.* (A) 63 (1950) 616
[61] SMITH, C. S. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 15
[62] SHUTTLEWORTH, R., KING, R. and CHALMERS, B. *Proc. Roy. Soc.* (A) 193 (1948) 465
[63] DUNN, C. G. and LIONETTI, F. *J. Metals* 1 (1949) 125
[64] AUST, K. T. and CHALMERS, B. *Proc. Roy. Soc.* (A) 204 (1950) 359
[65] LOMER, W. M. and NYE, J. F. *Proc. Roy. Soc.* (A) 212 (1952) 576
[66] ESHELBY, J. D. *Phil. Mag.* 40 (1949) 903; see also FOREMAN, A. J. E., JASWON, M. A. and WOOD, J. K. *Proc. Phys. Soc.* (A) 64 (1951) 156
[67] SOMIGLIANA, C. R.C. *Accad. Lincei* 24 (1915) 655
[68] NEMÉNYI, P. *Z. angew. Math. Mech.* 11 (1931) 59
[69] MANN, E. H. *Proc. Roy. Soc.* (A) 199 (1949) 376
[70] BOGDANOFF, J. L. *J. appl. Phys.* 21 (1950) 1258
[71] NYE, J. F. *Proc. Roy. Soc.* (A) 198 (1949) 190; (A) 200 (1949) 47
[72] ZENER, C. *Phys. Rev.* 69 (1946) 128
[73] ESHELBY, J. D., FRANK, F. C. and NABARRO, F. R. N. *Phil. Mag.* 42 (1951) 351
[74] STARR, A. T. *Proc. Camb. Phil. Soc.* 24 (1928) 489
[75] FRANK, F. C. *Pittsburgh Symposium* (reference 9) p. 89
[76] KOEHLER, J. S. *Phys. Rev.* 60 (1941) 397
[77] See, for example: ROSCOE, R. *Phil. Mag.* 21 (1936) 399; REHBINDER, P., LICHTMANN, V. I. and MASLENIKOV, V. M. *C. R. Acad. Sci. U.S.S.R.* 32 (1941) 125; KEMSLEY, D. S. *Nature, Lond.* 163 (1949) 404; HARPER, S. and COTTRELL, A. H. *Proc. Phys. Soc.* (B) 63 (1950) 331; MENTER, J. W. and HALL, E. O. *Nature. Lond.* 165 (1950) 611; ANDRADE, E. N. da C. and RANDALL, R. F. Y. *Nature, Lond.* 163 (1948) 890; 161 (1949) 1127; PHILLIPS, D. J. and THOMPSON, N. *Proc. Phys. Soc.* (B) 63 (1950) 839
[78] BARRETT, C. S. *Acta Metallurgica* (1953). To be published
[79] JASWON, M. A. and FOREMAN, A. J. E. *Phil. Mag.* 43 (1952) 201
[80] NABARRO, F. R. N. *Some Recent Developments in Rheology* British Rheologists Club, London 1950
[81] KUHLMANN, D. *Z. Phys.* 124 (1947) 168
[82] NABARRO, F. R. N. *Advances in Physics* 1 (1952) 271
[83] SEITZ, F. *Advances in Physics* 1 (1952) 43
[84] MOTT, N. F. *Proc. Phys. Soc.* (B) 64 (1951) 729
[85] GYULAI, Z. and HARTLY, D. *Z. Phys.* 51 (1928) 378
[86] MOLENAAR, J. and AARTS, W. H. *Nature, Lond.* 166 (1950) 690
[87] DUGDALE, R. A. To be published
[88] COTTRELL, A. H. and BILBY, B. A. *Proc. phys. Soc.* (A) 62 (1949) 49
[89] DIJKSTRA, L. J. *J. Metals* 1 (1949) 252
[90] SNOEK, J. L. *Physica* 8 (1941) 711
[91] MOTT, N. F. *Imperfections in Almost Perfect Crystals* John Wiley, New York 1952
[92] COTTRELL A. H. and JASWON, M. A. *Proc. Roy. Soc.* (A) 199 (1949) 104
[93] LOMER, W. M. *Private communication*
[94] ARDLEY, G. and COTTRELL, A. H. To be published

[95] SCHWARTZBART, H. and LOW, J. R. *J. Metals* 1 (1949) 637
[96] CHURCHMAN, A. T. and COTTRELL, A. H. *J. Metals* 1 (1949) 877; *Nature, Lond.* 167 (1950) 943
[97] HOLDEN, A. N. and HOLLOMON, J. H. *J. Metals* 1 (1949) 179. 879
[98] TÚRY, P. and KRAUSZ, S. *Nature, Lond.* 138 (1936) 331; 139 (1937) 30
[99] OROWAN, E. *Z. Phys.* 89 (1934) 634
[100] WAIN, H. L. and COTTRELL, A. H. *Proc. Phys. Soc.* (B) 63 (1950) 339
[101] SMITH, C. L. *Nature, Lond.* 160 (1947) 466
[102] COTTRELL, A. H. and GIBBONS, D. F. *Nature, Lond.* 162 (1948) 488
[103] GIBBONS, D. F. *Proc. Phys. Soc.* (B) 65 (1952) 193
[104] DUMBLETON, M. J. and HOWLETT, B. W. *Proc. Phys. Soc.* To be published
[105] SYLVESTROWICZ W. and HALL, E. O. *Proc. Phys. Soc.* (B) 64 (1951) 495; HALL, E. O. *Proc. Phys. Soc.* (B) 64 (1951) 742, 747
[106] BILBY, B. A. *Sheet Metal Industries* August (1950) 707
[107] COTTRELL, A. H. *Pittsburgh Symposium* (reference 9) p. 60; *Report of 9th Solvay Conference* (1952) 487
[108] HARPER, S. *Phys. Rev.* 83 (1951) 709
[109] COTTRELL, A. H. and CHURCHMAN, A. T. *J. Iron and Steel Inst.* 162 (1949) 271
[110] WERT, C. *J. appl. Phys.* 20 (1949) 943
[111] LEAK, G. M. and COTTRELL, A. H. *J. Iron and Steel Inst.* To be published
[112] HONEYCOMBE, R. W. K. *J. Inst. Metals* 80 (1951) 45; 49
[113] TAYLOR, G. I. *Proc. Roy. Soc.* (A) 116 (1927) 39
[114] KARNOP R. and SACHS, G. *Z. Phys.* 53 (1929) 605; GÖLER, F. VON and SACHS, G. *Z. Phys.* 55 (1929) 581
[115] ANDRADE, E. N. DA C. and HENDERSON, R. *Phil, Trans, Roy. Soc.* (A) 244 (1951) 177
[116] MASIMA, M. and SACHS, G. *Z. Phys.* 51 (1928) 321; 56 (1929) 394
[117] MASING, G. and RAFFELSIEPER, J. *Z. Metallk.* 41 (1950) 65
[118] LÜCKE, K. and LANGE, H. *Z. Metallk.* 43 (1952) 55
[119] RÖHM, F. and KOCHENDÖRFER, A. *Z. Metallk.* 41 (1950) 265
[120] PRATT, P. To be published
[121] CHEN, N. K. and MATHEWSON, C. H. *J. Metals* 3 (1951) 653
[122] LALOEUF, A. and CRUSSARD, C. *Rev. Met.* 48 (1951)
[123] FRANK, F. C. *Private communication*
[124] KUHLMANN, D. *Proc. Phys. Soc.* (A) 64 (1951) 140
[125] BROWN, A. F. *Advances in Physics* 1 (1952) 427
[126] See, for example. WILMS, G. R. and WOOD, W. A. *J. Inst. Metals* 75 (1948–49) 693; WOOD, W. A. and RACHINGER, W. A., *J. Inst. Metals* 76 (1949–50), 237; WOOD, W. A. and SCRUTTON, R. F. *J. Inst. Metals* 77 (1950) 423; RACHINGER, W. A. *J. Inst. Metals* 80 (1951–52) 415; GREENOUGH, G. B. and SMITH, E. M. *J. Inst. Metals* 77 (1950) 435; CALNAN, E. A. and BURNS, B. D. *J. Inst. Metals* 77 (1950) 435; RAMSAY, J. A. *J. Inst. Metals* 80 (1951) 167; WOOD, W. A. and SUITER, J. W. *J. Inst. Metals* 80 (1952) 501
[127] COTTRELL, A. H. and AYTEKIN, V. *J. Inst. Metals* 77 (1950) 389
[128] KUHLMANN, D., MASING, G. and RAFFELSIEPER, J. *Z. Metallk.* 40 (1949) 241
[129] BORELIUS, G. *Report of 9th Solvay Conference* (1952) 427
[130] MOTT, N. F. *Proc. Phys. Soc.* 60 (1948) 391
[131] TURNBULL, D. *Trans. Amer. Inst. Min. Met. Engrs.* 161 (1945) 140
[132] GUINIER, A. and TENNEVIN, J. *Progress in Metal Physics II* (Ed: B. Chalmers) Chapter VI 177 (1950)
[133] DUNN, C. G. and DANIELS, F. W. *Trans. Amer. Inst. Min. Met. Engrs.* 191 (1951) 147

[134] KELLAR, J. N., HIRSCH, P. B. and THORP, J. S. *Nature, Lond.* 165 (1950) 554
[135] HEIDENREICH, R. D. *J. appl. Phys.* 20 (1949) 943; *Bell System Tech. Journal*, Oct. (1951)
[136] TIEDEMA, T. J., MAY, W. and BURGERS, W. G. *Acta Cryst.* 2 (1949) 151
[137] BECK, P. A., SPERRY, P. R. and HU, HSUN *J. appl. Phys.* 21 (1950) 420
[138] ACHTER, M. R. and SMOLUCHOWSKI, R. *Phys. Rev.* 83 (1951) 163; *J. appl. Phys.* 22 (1951) 1260
[139] ANDERSON, W. A. and MEHL, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* 161 (1945) 140
[140] COOK. M. and RICHARDS, T. LL. *J. Inst. Metals* 73 (1946) 1
[141] BURGERS, W. G. and LOUWERSE, P. C. *Z. Phys.* 67, (1931) 605
[142] CAHN, R. W. *Proc. phys. Soc.* (A) 63 (1950) 323
[143] TIEDEMA, T. J. *Proc. Acad. Sci. Amsterdam* 53 (1950) 1422
[144] DAVIS, M. and THOMPSON, N. *Proc. Phys. Soc.* (B) 63 (1950) 347
[145] WYATT, O. H. *Nature, Lond.* 167 (1951) 866
[146] WYATT, O. H. *Proc. Phys. Soc.* To be published
[147] COTTRELL, A. H. *J. Mech. Phys. Solids* 1 (1952) 53
[148] BAILEY, R. W. *J. Inst. Metals* 35 (1926) 27
[149] OROWAN, E. *J. West Scotland Iron and Steel Inst.* 54 (1946–47) 45
[150] CARREKER, R. P., LESCHEN, J. G. and LUBAHN, J. D. *Trans. Amer. Min. Inst. Met. Engrs.* 180 (1949) 139
[151] SULLY, A. H. *Metallic Creep and Creep-resistant Alloys* Butterworths Scientific Publications, London 1949; STANFORD, E. G. *The Creep of Metals and Alloys* Temple Press, London 1949; ROTHERHAM, L. *Creep of Metals* Institute of Physics, London 1951
[152] MCLEAN, D. *J. Inst. Metals* 80 (1951–52) 507
[153] KING, R., CAHN, R. W. and CHALMERS, B. *Nature, Lond.* 161 (1948) 682
[154] KÊ, TING-SUI *Phys. Rev.* 71 (1947) 533; 72 (1947) 41; 73 (1948) 267
[155] ROTHERHAM, L., SMITH, A. D. N. and GREENOUGH, G. B. *J. Inst. Metals* 79 (1951) 439
[156] NABARRO, F. R. N. *Rep. Conf. on Strength of Solids* Physical Soc. London (1948) 75
[157] HERRING, C. *J. appl. Phys.* 21 (1950) 437
[158] MOTT, N. F. *Report of 9th Solvay Conference* (1952) 515
[159] HANSON, D. and WHEELER, M. A. *J. Inst. Metals* 45 (1931) 229
[160] BETTERIDGE, W. and FRANKLIN, A. W. *J. Inst. Metals* 80 (1951–52) 147

6

DIFFUSION IN METALS

A. D. Le Claire

THE last few years have been witness to a widening interest in the phenomenon of diffusion in metals and it is the purpose of the present chapter to summarize the more important advances that have been made, particularly since the time of writing of Volume I of this series,[1] which will be referred to hereafter as PMP.1.

Of greatest importance to our general understanding of chemical diffusion in polycrystalline materials, the case of most practical interest, has been the work on the Kirkendall Effect and on grain boundary diffusion processes and the first two sections are devoted to these topics. A knowledge of the fundamental atomic process by which diffusion occurs is essential for a complete understanding of any diffusion phenomenon and the discussion of this topic forms the subject of the next section. This includes an account of the first appreciable successes which have been achieved in relating experimentally determined diffusion coefficients to theoretical expressions for them. A few new and interesting methods for measuring diffusion coefficients are described in Section IV and in the last section an account is given of recent developments in the phenomenological and kinetic treatments of chemical diffusion.

Perhaps the most important result of the last few years work has been to provide a clearer understanding of the nature of the problems involved in studying diffusion in metals and so to point the way to the type of experiments which are most likely to further progress.

SECTION I

THE KIRKENDALL EFFECT

The original experiments of SMIGELSKAS and KIRKENDALL,[2] revealing what has now come to be known as the "Kirkendall Effect," were performed in 1946 and reviewed in PMP.1. Since then a considerable amount of experimental work has been devoted to confirming their results, to examining them in more detail and to demonstrating the occurrence of the Kirkendall effect in an extensive range of systems.

The importance of the Kirkendall effect lies in the assistance it gives to elucidating the precise atomic mechanism of diffusion, in its demonstrating the inadequacy of a single diffusion coefficient for fully

describing the diffusion behaviour of a binary alloy, and when studied in more detail, in revealing the complexity of the process which experiments on chemical diffusion seek to study.

The first of these points is dealt with in Section III. In this section we shall review the experimental evidence on which the second two points are made and include a reference to the more refined phenomenological description of binary diffusion which the occurrence of a Kirkendall effect demands.

As was described in more detail in PMP.1 SMIGELSKAS and KIRKENDALL laid a parallel set of fine molybdenum wires along each of a pair of opposite faces of a rectangular bar of 30 per cent α-brass and then plated it overall with a thick layer of copper. After annealing to produce diffusion the opposite sets of molybdenum wire markers were found to have moved in towards one another by an amount proportional to the square root of the annealing time and greater than could be accounted for by density changes due to changes in composition. This shift of markers placed at the original interface between two interdiffusing metals or alloys, relative to parts of the sample where no diffusion occurs, is what is known as the "Kirkendall Effect." It has now been confirmed for various compositions and temperatures in copper α-brass couples by CORREA DA SILVA and MEHL,[3] BÜCKLE and BLIN,[4] and BARNES[5] and shown by these investigators and by SEITH and KOTTMANN[6, 128] to exist also in many other single-phase interdiffusing systems such as Cu/Ni,[3, 5, 6, 128] Cu/Au,[3] Ag/Au,[3, 6, 128] Ag/Pd,[6, 128] Ni/Co,[6, 128] Ni/Au,[6, 128] and Fe/Ni[128] and in the α-ranges of Sn/Cu,[3] and Al/Cu.[3, 4]

A variation in technique employed in many of these experiments was the use of "sandwich samples" prepared by welding to the opposite faces of a plate or disc of one of the metals or alloys, on which the markers were placed, plates or discs of the other metal or alloy. In all cases where measurements were made to show it,[2–5] the extent of the marker shift was again found proportional to the square root of the annealing time, strongly suggesting it to be a true diffusion phenomenon; it also increases with temperature at fixed annealing time, and is larger the greater the initial difference in composition between the two halves of the diffusion sample. In all cases the movement is towards that side of the sample with the lower melting point[3]—towards copper in copper-nickel samples but towards gold in copper-gold samples. The observed shifts are of order 1/10 mm for the times and temperatures normally employed in diffusion experiments.

CORREA DA SILVA and MEHL[3] measured the Kirkendall effect in a large number of sandwich samples both by the direct SMIGELSKAS and KIRKENDALL method and by new methods. These included, for example, measuring relative to the centre of a foil marker, the shift

along the diffusion direction of fine wire markers inserted at the same interface ("foil reference method") or of the edge of the foil itself, which was found to keep pace with the wire markers ("foil bend method"). See Fig. 1. No diffusion occurs across the centre of the foil which, therefore, always defines the original position of the interface. Such methods avoid measuring small changes in distance between well separated sets of markers as in sandwich samples, which might be affected by secondary processes unconnected with diffusion, but at the

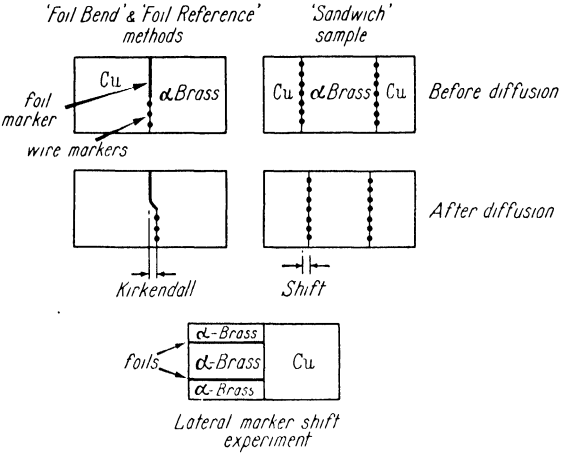


Fig. 1. Kirkendall Effect Experiments

same time might be expected to interfere much more than wire markers with the unidirectional diffusion flow. They demonstrated that whatever method of measurement was used and whatever type of marker was employed, be it alumina powder, carbonized silk thread or any one of a number of different metal wires or foils, the extent of the marker shift for any one system was always roughly the same for a given temperature and time of anneal; this was true, most strikingly, even for platinum wires which dissolved to produce as "markers" regions within the lattice itself of high platinum concentration. This confirms that the markers play only a passive role in the process, serving merely to reveal a movement relative to one another of different parts of the lattice: to be explicit we may assume they are fixed relative to the lattice. Using the copper α-brass case as an example, an immediate inference from the Kirkendall effect then is that relative to the lattice the number of zinc atoms flowing from the brass towards the copper must be greater than the number of copper atoms flowing in the opposite direction to take their place. The copper side of the interface must expand, new lattice planes being created, to accommodate the increased number of atoms and if a non-defective lattice in which all sites are occupied is to

be maintained the brass, which suffers a nett loss of atoms, must at the same time shrink. Thus is produced a motion towards regions richer in the more rapidly diffusing zinc of all lattice planes in the diffusion zone, and in particular of wire markers at the original interface, as observed.

Several processes have been suggested which may lead to the unequal opposite flows of the two species of atoms which the Kirkendall effect implies. It now seems unlikely that a preferential transport of one component through the vapour phase is in general important, as was proposed[7] for the zinc in α-brass-copper couples, if only because a Kirkendall effect is found in such systems as the copper-aluminium one[3, 4] where the vapour pressures are roughly equal and about 10^6 times less than that of zinc. In addition, successive investigations by different workers[2, 4, 3, 5] on 30 per cent α-brass-copper couples have, through changes in technique, effected a removal or variation of special facilities for possible vapour transport such as porosity in electroformed deposits and imperfections at the initial interface, without leading to any appreciable change in the magnitude of the Kirkendall effect. This insensitivity of results to a most probable variation also of grain structure provides evidence against a second possibility that the higher rate of flow of one component arises principally from its faster diffusion at grain boundaries. However, slight bulges in the moving interface in the direction of an increased Kirkendall effect have been observed in nickel-copper samples[5] near those nickel grain boundaries where suitable etching has revealed a preferential diffusion of copper, but these only represent a very small addition to the shift of interface apparent in regions well away from grain boundaries. Furthermore, Kirkendall effects have been observed[6, 128] in samples the grain size of which was purposely made large (3–15 mm). It appears then that we must assume the different rates of transfer to be associated principally with differences in the diffusion rates of the two atoms in the lattice itself.

It is only at this stage that we are able to reject the possibility of a ring diffusion mechanism for those alloy systems in which a Kirkendall effect has been found (see Section III).

The Kirkendall effect experiments, however, do not constitute the first or only evidence available to show that relative to the crystal lattice the two species of a binary alloy diffuse at unequal rates. As early as 1933 SEITH and KEIL[24] showed that the diffusion rates, at low concentration, of several elements in lead were substantially greater than the self diffusion rates of lead in the same alloys. JOHNSON,[13] also, in 1942 demonstrated by direct measurement the inequality of the self diffusion rates of silver and of gold in a 50 per cent alloy of silver and gold.

BÜCKLE and BLIN[4] have suggested that the two species diffuse by two different mechanisms, the zinc, say, by an interstitial mechanism

and the copper by a vacancy mechanism, so that unequal diffusion rates are to be expected. In the absence of any other evidence for interstitial solid solution in either α-brass or in any other system showing a Kirkendall effect, and in view further of HUNTINGTON and SEITZ's calculations (see Section III), this is not a very acceptable suggestion. SEITZ[59] first clearly pointed out that diffusion of both species by the vacancy mechanism is perfectly adequate to describe the greater transfer of, say, zinc than copper, for it is only necessary to assume that vacancies exchange with zinc atoms more readily than with copper.

The Phenomenological Description of Diffusion with a Kirkendall Effect and Further Phenomena Associated with Chemical Diffusion

Whatever the mechanism of diffusion the demonstration of unequal rates of flow for the two atomic species led DARKEN[8] to propose that two separate diffusion coefficients D_a and D_b should be defined, one for each species, specifying their respective rates of flow relative to the markers, or to planes of the lattice. With certain simplifying assumptions he then showed, as described fully in PMP.1, that on account of the difference in rates of flow, planes of the lattice, or markers, at points in the diffusion zone where the concentration gradient is $\partial N_a/\partial x$ will move relative to points where there is no concentration gradient and so well removed from the diffusion zone with a velocity v given by

$$v = (D_a - D_b)\partial N_a/\partial x \quad . \quad . \quad . \quad . \quad (1)$$

Also, the single chemical interdiffusion coefficient as normally measured experimentally, and referred strictly to the MATANO[9] interface, is given in terms of the separate or "partial chemical diffusion coefficients" D_a and D_b by

$$\tilde{D} = N_a D_b + N_b D_a \quad . \quad . \quad . \quad . \quad (2)$$

where N_a and N_b are the fractional concentrations of the two species at the composition to which $\tilde{D}$, D_a and D_b refer. Since $D_a \neq D_b$, $\tilde{D}$ necessarily varies with composition, although such variation is due also to D_a and D_b since $\tilde{D}$ does not appear to vary linearly with composition.

$\tilde{D}$ measures the rates of flow relative to a surface defined such that equal numbers of atoms of each species diffuse in opposite directions across it. This has the merit of simplicity in that such rates of flow are equal for the two species, and in many practical cases $\tilde{D}$ is an adequate measure of the diffusion behaviour of an alloy. But a fuller description is obtained in terms of D_a and D_b and a better understanding of diffusion in binary alloys would probably be obtained through study of these theoretically simpler and more fundamental coefficients than through study of $\tilde{D}$, as has so often been attempted in the past.

In sufficiently dilute solution of, say, A in B, $D_a \simeq \tilde{D}$. But in general D_a and D_b can only be calculated, using equations (1) and (2), from simultaneous measurement of $\tilde{D}$ and v. It is very often found that the marker shift x is proportional to $t^{1/2}$, t being the time of anneal. In such cases v can be obtained from a single measurement of x through the equation $v = x/2t$. However, LE CLAIRE[30] has pointed out that there is no obvious reason why in general x should be proportional to $t^{1/2}$, although such a relation holds when there is no concentration change at the markers, a condition approximately fulfilled for markers placed, as they usually have been, at the original interface. SEITH and KOTTMANN[128] have recently studied the movement of markers placed in positions in the diffusion zone other than at the original interface between the two dissimilar metals, and it can be seen from their results that, for some markers, approximately doubling the time of anneal more than doubles even the corresponding values of x. The relation $x \sim t^{1/2}$ is then not in fact generally true and v is best determined from a series of measurements of x and t.

But such determinations of D_a and D_b are based on the validity of the assumptions made in deriving these above equations. DARKEN assumed that changes in lattice parameter on diffusion could be ignored, that a non-defective lattice was fully maintained by complete shrinkage and that both expansion and shrinkage took place only along the direction of diffusion so maintaining constant the cross section through which diffusion occurs. The first assumption, never strictly valid of course except in self diffusion measurements, may be avoided simply by measuring distances in diffusion samples not in terms of centimetres but in numbers of lattice spacings. This practice has been adopted in recent years by many workers.[3, 10] The validity of the other assumptions is a matter for experiment.

A slight porosity in the brass side of their copper α-brass samples was reported by SMIGELSKAS and KIRKENDALL[2] and has been confirmed by later workers.[3–5] A similar porosity, consisting of large numbers of tiny holes on that side of the interface which the direction of the marker shift indicates is suffering a nett loss of atoms, has been found also in many other systems, such as Fe/Ni,[128] Ag/Au,[128] Cu/Ni,[5, 6, 128] Cu/Al,[3, 4] Ag/Pd,[6, 128] and Ni/Au.[6, 128] The amount of porosity varies considerably from system to system and sometimes may only be revealed by very careful mechanical polishing. It generally increases with increasing amount of diffusion and is greater the larger the initial difference in concentration. The density of holes is often not uniform throughout the side of the sample losing atoms but shows a well defined maximum in one region, the position of which has been observed in at least one case[128] to move as diffusion proceeds so as to occur always at a particular concentration in the diffusion zone, its distance from the interface varying

parabolically with the time of anneal. The pores may in some cases occupy, at the position of maximum density, a considerable proportion of the cross section of the diffusion zone in that region, and may join up in places to form larger cavities and fissures. SEITH and KOTTMANN[128] for example found this proportion to be as much as 30–40 per cent for a pure gold/pure silver sample annealed for 98 hours at 900°C. Occasionally the holes appear[4, 5] to be octahedral in shape with faces parallel to {111} planes. For this reason, for the fact that they are produced as readily in samples made from exclusively vacuum cast material[4] and because they are uninfluenced by variations in the atmosphere in which samples are annealed,[5] the porosity does not appear to be caused in any way by dissolved gases or by secondary gas reactions. From its location in the diffusion zone then the existence of porosity strongly suggests that shinkage is incomplete. A sensitive test for complete shrinkage is to compare the actual marker shift it produces with that to be expected from the total nett transfer of atoms across the marker interface, as determined, for example, by subsequent chemical analysis. The two are only equal when complete shrinkage occurs; otherwise the observed will be less than the calculated shift. SMIGELSKAS and KIRKENDALL[11] found that this latter was the case in their experiments and attributed the difference to their observed porosity. Now in a sandwich sample of brass between copper the inward marker shift is due to the shrinkage of brass, but in a sandwich of copper between brass the outward marker movement is due to the expansion of the copper. BÜCKLE and BLIN[4] compared the shifts in these two types of samples and showed that the expansion of the copper, which was quite regular from sample to sample, was greater than the shrinkage in the brass, which showed more variation between successive samples, by an amount which corresponded quantitatively with the measured total volume of the holes that they found in the brass. Similar results were obtained by SEITH and KOTTMANN who, by placing additional markers outside the diffusion zones at the ends of sandwich samples, were able to compare for each sample the two measurements of the Kirkendall shift made relative to points, outside the diffusion zone, on each side of an interface. The end markers in such samples are also generally found[5, 6, 128] to have moved apart after diffusion revealing an overall increase in length of the samples. This is of course to be expected if a porosity or incomplete shrinkage occurs, as also is a decrease in the overall density. Such changes of length and density are most apparent when magnified by using as diffusion samples a stack of strips or foils of the two metals packed alternately. In this way BARNES[5] showed the change in length to be proportional to the number of diffusion interfaces and to the square root of the annealing time and the change in density to correspond roughly to the total volume of holes.

The available evidence, both quantitative and qualitative, is then that this porosity is intimately connected with the diffusion process, and in view of its location in the diffusion zone, is indicative of an incomplete shrinkage of that part of the diffusion zone suffering a nett loss of atoms. As a typical example, BÜCKLE and BLIN[4] found the marker shift referred to points outside the diffusion zone on the brass side of copper/α-brass samples, to be between 1/2 and 2/3 of the more consistent values measured relative to points on the copper side. However, CORREA DA SILVA and MEHL[3] considered that porosity was negligible and shrinkage effectively complete, for in the many samples they examined they found the Matano interface, as computed graphically from plots of concentration as a function of distance expressed in lattice spacings, to coincide with the original position, relative to the ends of the sample, of the initial interface between the two halves of the sample. While this is a perfectly valid test for complete shrinkage, it is, as BALUFFI and ALEXANDER[12] point out, a very insensitive one, for the graphically computed Matano interface varies only slightly in position for considerable changes in porosity.

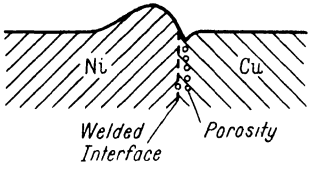


Fig. 2. Lateral Dimension Changes in Cu-Ni Diffusion Sample (Ref. 6)

CORREA DA SILVA and MEHL[3] measured the lateral shift of foil markers placed parallel to the direction of diffusion (Fig. 1) in a copper/α-brass specimen and concluded that since it was negligible all dimensional changes were in fact unidirectional as assumed by DARKEN.[8] But even the shift to be expected from lattice parameter changes at the interface, which would be in the same direction as that due to a lateral shrinkage, was not observed, and it may well be that an expansion produced by porosity compensated for this expected contraction. However, examination of the free surface of a specimen, where there is less constraint upon the tendency to expand or shrink, imposed by surrounding material, has provided evidence that such dimensional changes may take place normal to the direction of diffusion. Fig. 2 shows the situation observed in nickel/copper couples by SEITH and KOTTMANN,[6] and similar effects have been observed[128] by these authors in all the systems they studied. The Kirkendall shift is towards the copper in this system and an apparent expansion and shrinkage leading to an increase and decrease in the diameter of the sample on the sides of the interface suffering respectively a nett gain and a nett loss of atoms, is clearly evident. Similar effects have been noticed before by JOHNSON[13] and by FRAENKEL and HOUBEN[14] in gold/silver diffusion samples. JOHNSON attributed the whole of the effect he found to evaporation but SEITH and KOTTMANN prevented free evaporation by packing their samples in

alumina. In addition they observed in silver/palladium samples an overall decrease in length which they attribute to the nett lateral swelling of samples and which presumably is greater than the increase in length the porosity produced. This could not arise from vapour transport. That there is a tendency operating across the whole of the sample to shrink on one side of the interface and expand on the other is evident from the marked bending, observed by BARNES,[5] in nickel/copper samples that were very thin in the direction of diffusion. The curvature, uniform across the sample, was such that copper was on the concave side. Thermal expansion effects would have produced the opposite curvature in this experiment. SEITH and KOTTMANN[128] estimate that for one of their silver-gold samples (98 hr at 900°C) only two-thirds of the expansion on the gold side of the interface occurred along the direction of diffusion, the rest occurring laterally to produce an increase in diameter in that part of the diffusion zone.

These results show that the assumptions underlying equations (1) and (2) may not always be even approximately satisfied experimentally. The occurrence both of porosity and of lateral dimension changes means that the cross section through which diffusion is occurring is not constant in time or uniform throughout the diffusion zone and may be very appreciably reduced in the region of maximum porosity. This may seriously impair the accuracy of measurement of $\tilde{D}$. SEITH and KOTTMANN[128] determined from their Ag/Au and Fe/Ni samples, using the Matano method, the variation of $\tilde{D}$ with composition for these systems and in each case the results showed what is most probably a false minimum for $\tilde{D}$ near those compositions at which maximum porosity occurs.

The same effects will also in general influence the value of v. The best value of v for use in equation (1) is obtained by measuring the marker shift relative to points outside the diffusion zone, on that side of the interface richer in the slower diffusing component: i.e. by using samples in which the shift measured is produced by the expansion of that part of the diffusion zone receiving a nett addition of atoms. The direct effect of porosity in producing a lower value of the shift when this is measured relative to points on the other side of the diffusion zone is thereby avoided. But there remains the indirect effect of porosity, through its influence on the general course of diffusion, and the direct effect of the expansion being to some extent lateral to the direction of diffusion.

However, it has usually been experiments in which two dissimilar pure metals are diffusing into one another that have revealed marked porosity and lateral dimension changes. When the initial difference in concentration is smaller, as for diffusion between one metal and an

alloy of it with another, these phenomena are less marked and their effect on measurements of $\tilde{D}$ and v may be well within the normal experimental errors. But no attention has been given in the past to their possible influence and many experimental results may be suspect for this reason, especially those from samples covering a large concentration range.

The Vacancy Mechanism of Diffusion and the Kirkendall Effect

We now return to the vacancy mechanism description of diffusion in a binary system, using the copper/α-brass case for convenience in discussion. If the vacancies exchange with zinc atoms more readily than with copper atoms then the flow of zinc atoms towards the copper, arising from a concentration gradient of zinc atoms, will be greater than the flow of copper atoms towards the brass, as the Kirkendall effect indicates is actually the case. But since each atom which moves in one particular direction causes a vacancy to move in the opposite direction, there will be a nett flow of vacancies from the copper towards the brass equal to the nett flow of atoms, which is towards the copper. (A mathematical expression for the nett vacancy flow in terms of different probabilities of exchange of a vacancy with the two species of atoms has been given by SEITZ[15, 59] and is dealt with in Section V.) The interdiffusion of brass and copper is then accompanied by a pumping of vacancies from copper to brass, and we might expect the vacancy concentration to fall below that corresponding to equilibrium in the copper and to exceed it in the brass. But the observed nett transfer of atoms towards the copper is much greater than could be effected by that number of vacancies which are normally in thermodynamic equilibrium with the volume of copper usually employed in the experiments, so it is clear that some source of vacancies must be operating in the copper to produce the large number of extra vacancies required. Also, the nett number of atoms lost from the brass to the copper is replaced by an equal number of vacancies in the brass and since the volume of a vacancy is about the same as that of an atom the vacancies must disappear in some way if the brass is to shrink and the markers are to move towards the brass side of a diffusion sample.

Possible sources for vacancies are at the surface of the specimen, or at the grain boundaries or dislocations it contains. If the surface acted as the predominant source, as once suggested by SEITZ,[59] then we should expect the Kirkendall effect to depend on the volume of the diffusion sample and on whether the outer surface was of copper or of brass, but such dependence has never been found experimentally.[3–5] In particular the marker shift is always towards the brass and by about the same amount in samples with copper sandwiched between brass and with brass sandwiched between copper, any difference between the

two being satisfactorily accounted for as due to porosity. In the absence of experiments on the Kirkendall effect in single crystals which might decide whether or not grain boundaries alone were the predominant source, the most attractive and simple sources of vacancies are the dislocations.[16] Fig. 3 shows a simple edge type dislocation. If an atom at S moves down to T a vacancy is left at S and may then diffuse away by exchange with other atoms. If this is repeated by all S atoms in planes parallel to the figure, the whole dislocation will have moved along ZS by one lattice spacing without change of form and one vacancy will have been produced for each lattice plane. If a dislocation moves in this way right across from one side of a crystal to another, a new plane of atoms will have been inserted into the lattice and the two halves of the crystal will have moved away from one another by one lattice spacing to accommodate it. Accommodation is therefore provided for the nett number of atoms diffusing into the copper by the very process which produces the vacancies necessary for their transfer.

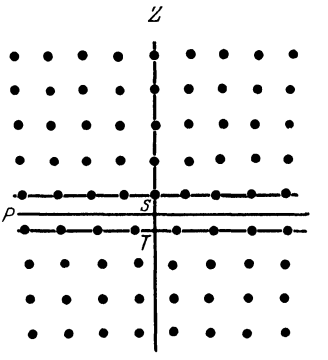


Fig. 3. Simple Edge Dislocation

There is clearly a limit to the extent to which any one simple edge dislocation may act as a source of vacancies, for the completion of the extra plane of atoms means the removal of that dislocation from the crystal. It is doubtful whether the density of dislocations, if all of the simple edge type, would be sufficient to sustain this loss without it influencing the course of the diffusion. But a dislocation line is not in general of the simple straight edge type. It may form a closed ring of one type or another or may meander across the crystal from one side to the other, assuming an edge character in some regions and a screw character elsewhere, much as in Fig. 4. Each section of edge character, such as mn, may act as a source of vacancies in much the same way as described above but extra atoms are not added to complete a single plane but so as to build up as a spiral surface continuous with the spiral structure of, and therefore extending, the screw dislocations joined at the ends of the edge. For each complete turn of the spiral the edge is merely moved along the slip direction without change of form. Atoms may then be added indefinitely, that half of the crystal to the left of mn moving relatively to the other half

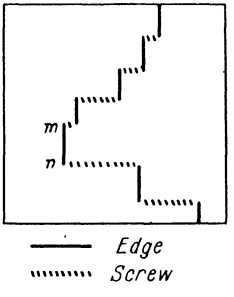


Fig. 4. General Dislocation line with Edge and Screw Components

in order to accommodate them, and any number of vacancies produced. The process is completely analogous to the mechanism of crystal growth proposed by FRANK[17] (see Chapter 5 of this volume), the edge dislocation to which atoms are added corresponding to the ledge to which they are added on the growing face of a crystal. Other types of dislocations which would act as extensive sources of vacancies have been described by SEITZ.[18]

When we come to consider mechanisms for removal of vacancies we notice at once that the above-mentioned processes have only to be reversed for dislocations to act as sinks for vacancies. Thus[16] a vacancy arriving on any of the sites surrounding the atom S may exchange with it and so raise the dislocation line one spacing along SZ. Vacancies condensing on the dislocation line in this way ultimately remove a whole plane of atoms and cause the two halves of the crystal to move towards one another by one atomic spacing. Similarly a general dislocation line with edge and screw components may act as an effectively infinite sink for vacancies merely by reversal of that process by which it acts as an infinite source.

An alternative proposal[3, 16] is that vacancies may aggregate on particular planes to produce platiform voids which, when they reach a critical size, collapse by inducing plastic flow in the surrounding material and with a resultant overall shrinkage. Such a collapsed platiform void has all the properties and characteristics of a general ring dislocation with a component of its Burgers Vector normal to its plane ("Prismatic Dislocation") which, as SEITZ points out,[60] are sufficient to explain all those properties of a crystal for which dislocations are usually invoked. In fact he suggests that the dislocation content of a crystal may arise solely in this way by condensation, during solidification and cooling, of the very large number of vacancies present in the molten metal. In particular such rings may act as further sinks (or sources) for vacancies in addition to those provided by other types and forms of dislocation.

The presence of a porosity in diffusion samples suggests that vacancies may also aggregate in a more isotropic manner to produce roughly spherical cavities. These we might expect to be stable when larger than some critical size and to grow by further accretion of vacancies into the visibly observable voids as first suggested by LE CLAIRE and BARNES.[19] Produced in this way, it is not surprising that they should have crystallographic shape, for vacancies will exchange preferentially with atoms whose removal from the surface of the cavity produces faces of minimum surface energy. The incomplete shrinkage of that side of a diffusion couple suffering a nett loss of atoms, may then be attributed to formation of relatively stable aggregates of some of those vacancies which have effected this nett loss. Complete shrinkage only occurs when

the vacancies in excess of those corresponding to thermal equilibrium are removed completely by condensation on dislocations.

The nett flux of vacancies at any point is given by

$$D_a \cdot \partial c_a/\partial x + D_b \cdot \partial c_b/\partial x = (D_a - D_b)\partial c_a/\partial x$$

and SEITH and KOTTMANN[128] suggest that the maximum porosity occurs in a region of the diffusion zone where the concentration is such that this expression is a maximum.

If the expansion in the copper and the contraction in the brass are to be only along the direction of diffusion, the planes of atoms removed or added during the operation of a dislocation as sink or source of vacancies must be normal to this direction. The dislocation must then have its Burgers Vector parallel to the diffusion direction. But it is dislocations of just this orientation which will be most employed as sources and sinks, for if of other orientation those parts of their dislocation edges which are advancing along the diffusion direction would soon run out of the diffusion zone, and the capacity of that dislocation as a source or sink for vacancies in consequence be reduced. BARDEEN and HERRING[20] propose this as an explanation of CORREA DA SILVA and MEHL'S result (see p. 272) that there is no appreciable lateral change of dimensions on diffusion. While we have seen that this result is not in general true, it is true that the expansion and shrinkage take place to a much greater extent along the direction of diffusion than would correspond to isotropic expansion and shrinkage. A further factor contributing to this may be that the surrounding undiffused material will impose constraints upon a lateral shrinkage or expansion of the diffusion zone i.e. upon the removal or growth of planes of atoms in a direction parallel to that of diffusion, but not in directions orthogonal to this. When the diffusion zone extends over the whole volume of a sample, so that there is no undiffused material, the dimensional changes accompanying diffusion are in fact more nearly isotropic, as recent experiments on diffusion into very thin foils and wires have shown.[129]

The growth or removal of planes of atoms in both single crystal and polycrystalline material will introduce stresses into the lattice, in particular between neighbouring grains in polycrystalline material. These may only be relieved by inducing recrystallization and grain growth, by introduction of new dislocations and, therefore, of further sinks and sources for vacancies, and by processes of plastic deformation which themselves contribute to the observed marker shift and lateral changes in dimensions. Additional stresses with similar consequence will arise from planes of atoms being forced to conform to changes in lattice parameter on diffusion. Recrystallization and grain growth in the diffusion zone between two metals has often been reported in the

past[21] and a more recent observation[5] of "polygonization" in the same region provides experimental evidence for the plastic flow.

Rough calculations[15, 20] have claimed to show that the normal dislocation content of a metal is adequate to maintain an equilibrium concentration of vacancies throughout the diffusion zone, despite the presence of a nett flux of vacancies. For such to be the case the average distance travelled by each vacancy during its lifetime must be very much less than that over which diffusion occurs, and it is considered that the normally assumed density of dislocations should be adequate to ensure this, even when by no means every encounter of a vacancy with a dislocation leads to its annihilation. However, the occurrence of a marked porosity in many diffusing systems would appear to invalidate these conclusions and there seems no reason to suppose that a vacancy concentration which is everywhere in thermodynamic equilibrium with the lattice is in fact maintained in all cases of diffusion.

The creation and annihilation of vacancies will take place more readily at irregularities along an edge dislocation, such as "jogs",[73] so that the effective number of points where such processes can occur may in fact be an even smaller fraction of the number of atoms along dislocation edges than these calculations have allowed for.

Conclusions, relating to the Theoretical Value of Chemical Interdiffusion Measurements

The demonstration of the various phenomena associated with the Kirkendall effect makes it difficult to know what theoretical significance may be attached to the results of chemical diffusion measurements. We have already discussed the difficulties associated with the measurement of D_a and D_b through use of equations (1) and (2), arising from porosity and changes in lateral dimensions, but even when these are sufficiently small not to affect the calculation of D_a and D_b from measurements of $\tilde{D}$ and v, there would still appear to be several factors controlling the values obtained which are not normally taken account of, and which may be very difficult to take account of, in any theoretical understanding of their magnitudes.

Theoretical treatments usually assume that the rates of chemical diffusion in an alloy depend only upon the vacancy concentration and upon the different probabilities of exchange of the two atoms with a vacancy in the presence of a concentration gradient, these probabilities being determined only by the local composition. It is assumed in theoretical treatments that the vacancies are in equilibrium with the lattice but in the case of chemical diffusion there exists the possibility of a vacancy concentration which is not only removed from its equilibrium value but also continuously changing during the diffusion. Also, it has been suggested[22] that vacancies may aggregate to form double

vacancies in regions of high vacancy concentration and that these would be more mobile than single vacancies and so lead to a high diffusion rate for the atoms. Abnormally large diffusion rates may then occur in that side of a diffusion couple which suffers a nett loss of atoms. We may mention that two independent diffusion coefficients D_a and D_b may only be legitimately defined by Fick's Law, and so equations (1) and (2) are only strictly valid, when the vacancies are everywhere in equilibrium with the lattice. Otherwise the rate of flow of, say, A atoms depends not only on the concentration gradient of A atoms, but also on that of the vacancies. This is discussed more fully in Section V. Furthermore, there is the possibility of stresses influencing the rate of diffusion and of inhomogeneous plastic flow concurrent with the diffusion, which later may generate additional vacancies and produce distorted lattice regions in which abnormal diffusion rates occur. The last two factors are associated with the shrinkage and expansion which accompany chemical diffusion and, with the addition of stresses arising from the interdiffusion of atoms of unequal size, will tend to produce recrystallization and so an enhancement of the grain boundary contribution to the overall flow.

The diffusion rates then appear dependent upon a number of quantities many of them changing during diffusion and varying with the nature of the sample and all difficult to determine separately. The relative importance of these several factors in determining the measured values of chemical diffusion coefficients and their variation with composition and the details of their operation in chemical diffusion, remain matters for further experimental study, but their presence may well be found to be not unconnected with the often very considerable disagreement between the diffusion coefficient measurements made on the same system by different observers. Since the Kirkendall effect, and therefore phenomena associated with it, is larger the greater the difference in concentration between the two halves of a diffusion sample, we might expect the most reliable and repeatable chemical diffusion coefficient measurements to be those conducted with the smallest practicable concentration gradients. A study of the dependence of $\tilde{D}$ upon concentration gradient at fixed concentration might reveal to what extent these complicating features, associated with a Kirkendall effect, do, in fact, influence the measured value of $\tilde{D}$. But the rates of self diffusion of the two species of atoms in a chemically homogeneous alloy depend *only* upon the vacancy concentration and upon the probabilities of exchange of the atoms with vacancies. Since chemical homogeneity is maintained throughout the diffusion process the vacancy concentration remains always at its equilibrium value and none of the other complicating factors which may influence chemical diffusion rates and their measurement arise at all. Self-diffusion coefficients

of the atoms of an alloy have then a much simpler and direct significance and furthermore give directly the most fundamental quantity underlying all diffusion phenomena, namely the average number of random jumps made per second by each atom in a homogeneous alloy in thermodynamic equilibrium. Measurements of self-diffusion coefficients would probably be far more profitable in furthering a detailed understanding of diffusion in alloys than the already very considerable number of chemical diffusion coefficient measurements have been.

SECTION II

Grain Boundary and Surface Diffusion

It has always been appreciated that in the neighbourhood of a grain boundary and on the surface of a metal the rate at which diffusion takes place may differ appreciably from that within a single crystal or grain because of the different structure which these regions present. We may define coefficients D_l, D_{gb}, and D_s specifying respectively the rates of flow by diffusion in the lattice, in grain boundaries, and on the surface of a metal and a full description of the diffusion behaviour of any system requires a knowledge of all these coefficients and their variation with concentration. Such a knowledge is clearly important for a proper understanding of the practically important diffusion rates in polycrystalline materials for in general these will be determined by D_{gb} and D_s as well as D_l. Many of the discrepancies between the values of diffusion coefficients obtained by different workers on the same system for both self and chemical diffusion may well be connected with the very common failure to take proper account of the possible contribution to the observed diffusion flow by grain boundary processes, and which may have been more significant in some cases than others. D_s may vary with the orientation of the surface and D_{gb} with both the mutual orientation of the grains adjoining the boundary and with the direction of the grain boundary for fixed orientation. Studies of grain boundary and surface diffusion may then also provide information on the nature of those regions in which they occur.

Whereas lattice diffusion can by suitable means be studied apart from grain boundary and surface diffusion, these processes themselves must necessarily be accompanied by diffusion in the lattice of the crystals or grains with which they are so intimately associated, and so cannot be studied apart from lattice diffusion. Experimental methods are, therefore, directed towards a comparison of the rates of flow by grain boundary or surface diffusion with that by lattice diffusion.

Grain Boundary Diffusion

Two principal methods of investigation have been employed for grain boundary diffusion studies.

I. The comparison of the overall rates of diffusion in materials of various grain sizes but similar in all other respects. It has generally been assumed in the past[21] but as we shall see, erroneously, that only when the measured bulk diffusion coefficient varies with grain size is D_{gb} appreciably different from D_l.

II. By examination of individual grain boundaries in sections cut from diffusion samples parallel to the direction of diffusion the relative amounts of diffusion at the grain boundaries and in the neighbouring grains may be directly observed. This method usually requires the use of a suitable etch which will reveal differences in composition. It is clearly an essential condition for full advantage to be taken of this more informative method that there should be no changes in the grain boundary pattern during the diffusion. If the grain boundaries do move as a result of recrystallization or grain growth any difference between grain boundary and lattice diffusion rates may be completely masked, and even if not completely so the results will anyway be of indefinite quantitative significance.

In the comparatively few studies of grain boundary diffusion that have been made, it has been reported that D_{gb} and D_l do not always differ but that when they do, they, and the corresponding activation energies Q_{gb} and Q_l, where these have been estimated, are such that D_l is less than D_{gb} and Q_l greater than Q_{gb}. This is not surprising in view of the more open structure presented for diffusion at grain boundaries. This early work and its results have been fully reviewed on other occasions.[21, 25, 26, 50]

More recently a much needed mathematical analysis of simultaneous grain boundary and lattice diffusion has been presented[27] and has provided a method much less objectionable than hitherto, for the quantitative measurement of the ratio D_{gb}/D_l. Quantitative estimates of D_{gb} have been made in the past but these have usually been based on a far less satisfactory analysis, often employing, for example, the Dushman-Langmuir equation.[28] Application of the analysis has been made to a few systems and it has also shown that method I will not necessarily always reveal a difference between D_l and D_{gb} even if a very appreciable difference does in fact exist. This method must then be applied with caution and many reports for various systems studied by it, that D_{gb} and D_l are appreciably the same, may, therefore, be false. A second important new advance has been a study for the first time of the dependence of grain boundary diffusion rates upon the orientations with respect to the boundary of the two adjoining grains.

Methods I and II have both been used to reveal a preferential grain boundary diffusion in a few more systems in which it had not previously been observed, although much of this has been in connexion with the work mentioned in the last paragraph.

Simultaneous Grain Boundary and Lattice Diffusion

FISHER[27] has considered diffusion into a semi-infinite bi-crystal with its grain boundary normal to the surface, as shown in section in Fig. 5. Several assumptions are made to simplify the analysis. (1) The concentration C_0 of diffusing material at the surface $y = 0$ is maintained constant throughout the anneal. (2) The grain boundary is

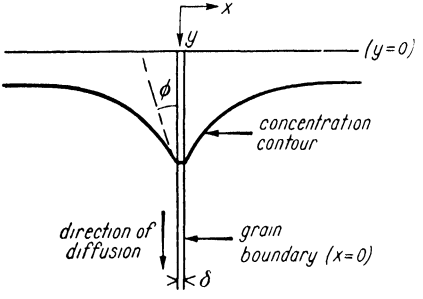


Fig. 5. Diffusion in the Neighbourhood of a Grain Boundary

regarded as a uniform slab of material of width δ in which the diffusion coefficient D_{gb} is much greater than that for diffusion in the adjoining crystals, D_l: the grain boundary region is not, of course, so well defined a region as this. Material diffusing more rapidly along the grain boundary than elsewhere will be lost from it laterally into the adjacent grains so producing an effective penetration of material which is greater in the neighbourhood of the grain boundary than the uniform penetration by lattice diffusion alone in regions well removed from the grain boundary. It is assumed (3) that this lateral loss occurs by diffusion only in a direction normal to the grain boundary, (4) that D_l and D_{gb} are independent of concentration, and (5) that the concentration in the grain boundary is uniform across its width δ; this is taken into account in drawing in Fig. 5, a line of constant concentration or a "concentration contour."

With these assumptions a numerical integration of the appropriate differential equations, for diffusion in the grain boundary and for lateral diffusion in the adjoining grains, gave a solution the behaviour of which had certain features which suggested a simple analytical expression for C_{gb}, the concentration of *laterally diffused material* in the vicinity of the grain boundary.

$$C_{gb} = C_0 \exp\left(- y_1\sqrt{2}/(\pi t_1)^{1/4}\right) Erfc(x_1/2t_1^{1/2}) \quad . \quad . \quad . \quad . \quad (3)$$

where x_1, y_1, and t_1 are dimensionless quantities given by

$$x_1 = x/\delta,\ y_1 = y/[\delta(D_{gb}/D_l)^{1/2}],\ t_1 = D_l t/\delta^2$$

It is to be emphasised that while this represents a useful and practicable relation for studying grain boundary diffusion, and so far as it has yet been tested[29] leads to results which are confirmed by experiment, it is obtained only by making very considerable simplifications in the mathematics and a more detailed and rigorous study of the problem would be very welcome. Failure to bear in mind the many assumptions upon which equation (3) is based may lead to serious errors in the interpretation of experimental results.

Let us now consider diffusion into a polycrystalline sample. For simplicity we assume this to have a columnar structure so that all the grain boundaries are parallel to the direction of diffusion and intersect the free surface normally as does the one in Fig. 5. In any thin section, of width Δy cut normal to the direction of diffusion at a distance y from the free surface $y = 0$, a certain amount A_l of the material which has diffused into the section will have been transferred only by lattice diffusion through the grains and the remainder A_{gb} by diffusion down the grain boundaries and by subsequent lateral diffusion from them. If there are n grains with average linear dimensions R normal to the columnar or diffusion direction, than A_l is given by

$$A_l = nR^2 \,.\, \Delta y \,.\, C_0 Erfc(y/2(D_l t)^{1/2}) \qquad . \ . \ . \ . \qquad (4)$$

where C_0 is the supposedly constant surface concentration of diffusing material and t is the annealing time. If we assume that the total amount of material actually in the grain boundaries is negligible compared with that which has been lost by lateral diffusion, we may write for A_{gb}

$$A_{gb} = 2nR \,.\, \Delta y \int_{-\infty}^{+\infty} C_{gb} \,.\, dx \qquad . \ . \ . \ . \qquad (5)$$

where C_{gb} is given by equation (3). TURNBULL[26] has used equations (4) and (5) to study the relative importance of transfer by lattice and by grain boundary diffusion in diffusion experiments on polycrystalline material, and has presented some results, applicable to a typical set of experimental conditions, in the form of Table 1. This shows for various values of D_{gb}/D_l the grain size for which A_l and A_{gb} are equal, when

TABLE 1

D_{gb}/D_l . . .	10^6	10^5	10^4	10^3
Grain size for which $A_{gb} = A_l$ (*cm*) .	0·26	0·0014	$2{\cdot}15 \,.\, 10^{-5}$	$< 10^{-6}$

$y = 0{\cdot}2$ cm, $t = 10^5$ sec, and $D_l = 10^{-9}$ cm^2/sec, conditions such that the concentration at y of material penetrating by lattice diffusion is ~ 15 per cent of c_0. If we may define lattice diffusion as predominant when $A_l \geq 10A_{gb}$, then since $A_l/A_{gb} \sim R$, experiments in which a measured diffusion coefficient showed no change when the grain size was varied from 2·6 to 0·014 cm would not be evidence for assuming that D_{gb} and D_l were appreciably the same, but only that $D_{gb}/D_l < 10^5$. Thus we may only expect method I to reveal a difference between D_{gb} and D_l when a very considerable range of grain sizes is employed. Alternatively, since Q_{gb} is usually less than Q_l, the ratio D_{gb}/D_l is increased with decreasing temperature, so that method I may produce positive results at sufficiently low temperatures where D_{gb} is so much greater than D_l that A_{gb} is comparable with A_l for the grain sizes used.

When D_{gb} is appreciably greater than D_l material will diffuse much further into the sample in the neighbourhood of a grain boundary than elsewhere and if sections are cut at a greater distance from the surface $y = 0$ than the maximum depth of penetration by lattice diffusion alone through the grains, their content of diffusing material will be given by equation (5). Performing the integration and taking logs, we find[27] for $A_{gb}/\Delta y$ the content per unit thickness,

$$\log (A_{gb}/\Delta y) = -[2^{1/2}(D_l/D_{gb}\delta)^{1/2}/(\pi D_l t)^{1/4}]y + \text{constant} \qquad (6)$$

where the constant varies with t but does not contain y.

This relation has been used by HOFFMAN and TURNBULL[29] to determine for the first time a grain boundary self-diffusion coefficient, that of silver. A thin layer of radioactive silver was electroplated on to one face of a number of polycrystalline samples, which were then annealed at temperatures in the comparatively low range 500–700°C. It was assumed that surface diffusion and diffusion in the highly disorganized thin electroplated layer would be sufficiently rapid to maintain a fairly constant surface concentration necessary for the validity of equation (3). The log of the specific activity of thin layers cut parallel to the surface after anneal was found to vary linearly with their distance y from the surface. Transfer by purely lattice diffusion would have required a proportionality to y^2; that this was not found may be regarded as evidence in favour of FISHER's equations (3) and (6). Knowing D_l (see equation (8)) and assuming a value of $5 \cdot 10^{-8}$ cm for δ, they found from equation (6) that for high purity silver

$$D_{gb} = 0{\cdot}03 \exp(-20{,}200/RT) \text{ cm}^2/\text{sec} \quad . \quad . \quad . \quad . \qquad (7)$$

At about 500°C, therefore, $D_{gb}/D_l \sim 10^5$.

To define D_{gb} properly requires the validity of assumption (2) above and to calculate it from equation (3) or (6) requires a knowledge of δ,

for the equations only give the product ($D_{gb}\delta$). But it may be more useful[30] to characterize grain boundary diffusion behaviour in terms of this measured quantity rather than D_{gb} alone for it is more significant in representing, without reference to δ, the amount of material diffusing in unit time along the grain boundary and across unit length of it normal to the diffusion direction, the concentration gradient also being unity.

HOFFMAN and TURNBULL[29] have also measured, by standard methods, the self-diffusion coefficients of silver in both single crystal and polycrystalline material, with the results shown in Fig. 6. At high tempera-

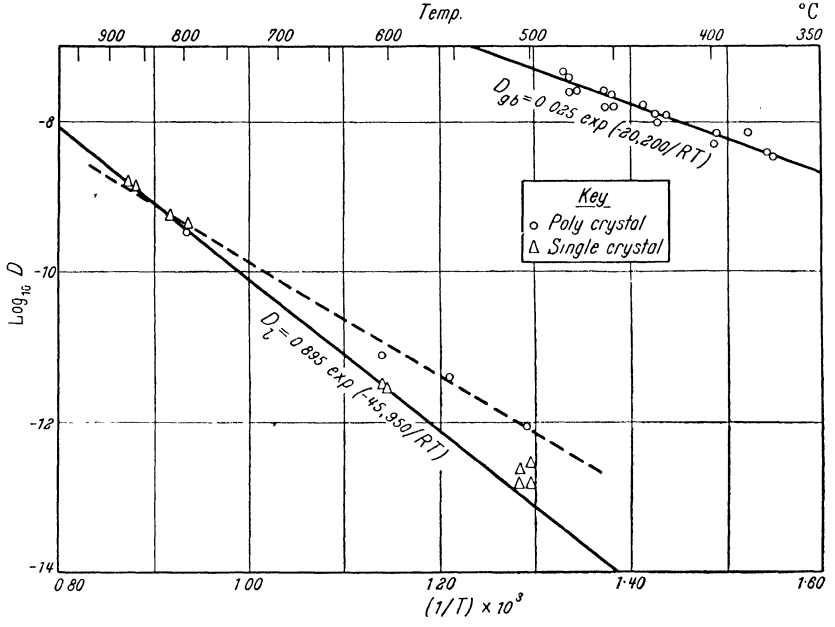


Fig. 6. Lattice and Grain Boundary Diffusion in Silver

tures single crystal and polycrystalline samples show the same self-diffusion coefficients and these results, which fit closely the earlier high temperature measurements of JOHNSON[31] on coarsely polycrystalline silver, may, therefore, be taken as representing the true lattice diffusion coefficient. The equation to the full line is

$$D_l = 0{\cdot}895 \exp(-45{,}950/RT) \text{ cm}^2/\text{sec} \quad . \quad . \quad . \quad . \quad (8)$$

so that $Q_{gb} \simeq \frac{1}{2}Q_l$. But at lower temperatures the apparent diffusion coefficient for polycrystalline material is appreciably greater than that for single crystals and for experiments between 500 and 650°C fit more closely the equation

$$D_{\text{app.}} = 2{\cdot}3 \,.\, 10^{-5} \exp(-26{,}400/RT) \text{ cm}^2/\text{sec} \quad . \quad . \quad . \quad . \quad (9)$$

It is clear that at high temperatures and with the grain sizes employed ($\sim 10^{-2}$ cm) D_{gb} is not sufficiently greater than D_l for the contribution to the overall diffusion flow by grain boundary diffusion to be significant, but that at lower temperatures where D_{gb}/D_l is greater, grain boundary diffusion is contributing appreciably.

These experiments are quoted in full because they illustrate a very important point in connexion with all experimental determinations of diffusion coefficients in polycrystalline materials which purport to give a measure of the lattice diffusion coefficient. It is usually considered that the more extensive the temperature range over which such measurements are made, the more accurate will be the resulting values of D_0 and Q (in the equation $D = D_0 \exp(-Q/RT)$) derived from them. We now see that they may represent lattice diffusion only at high temperatures but that abnormally high values of D may be obtained at the lower temperatures due to appreciable grain boundary diffusion. Experiments are not usually sufficiently accurate or numerous to reveal a departure at low temperatures from a linear relation between log D and $1/T$, so a "best straight line" drawn uncritically through both high and low temperature results, when plotted as log D against $1/T$, may then produce values of D_0 and Q which are of no particular theoretical significance and both less than the true values for lattice diffusion, which it is often assumed they represent. In fact the results for high temperatures in Fig. 6 with those at low temperatures for polycrystalline material, together lie quite closely to a straight line with equation

$$D = 8 \cdot 10^{-3} \exp(-38{,}000/RT)$$

D_0 is very appreciably less than, and Q lies between, the true values for lattice and for grain boundary diffusion.

HOFFMAN and TURNBULL took a contact autoradiograph of a taper section of a diffusion sample, cut in the diffusion zone and at a small angle to the plated surface. Blackening was strongest along a series of lines which corresponded exactly with the traces of the grain boundaries revealed by etching the same sample, and the lines thicker on that side of the section nearer the plated surface. This indicates, independently of FISHER'S analysis, that D_{gb} is, in fact, greater than D_l for self-diffusion in silver. The same experiment has been performed with tin by FENSHAM,[32] with similar results.

Method II (p. 281) has been used in the past only to reveal quantitatively any difference between D_{gb} and D_l. By suitable choice of etch, or by use of contact autoradiography in grain boundary self-diffusion studies, it may yield results from which concentration contours as in Fig. 5 could be obtained. Their shape at distances along the grain boundary greater than the maximum depth of penetration by lattice

diffusion alone is governed by equation (3) and can, therefore, be used to test this equation and to derive values of $(D_{gb}\delta)$.

LE CLAIRE[30] pointed out that simple differentiation of equation (3) shows the angle ϕ at which a concentration contour meets the grain boundary to be simply related to $D_{gb}\delta$, independently of concentration, by the equation

$$D_{gb}\delta/D_l = 2(\pi D_l t)^{1/2} \cot^2 \phi \quad . \quad . \quad . \quad . \quad (10)$$

This is derived directly from equation (3) without further assumptions, so that observations on this one simple geometrical feature of concentration contours can be used to check FISHER's equation. If valid then when D_{gb} and D_l are assumed independent of concentration ϕ is the same whatever the concentration at the contour, and for observations on the same grain boundary after successive anneals, $\tan^4 \phi$ should be proportional to t.

It also provides a quick and convenient method for determining D_{gb} and its variation with temperature and orientation. BARNES[33] has used method II to show that copper diffuses more rapidly in the grain boundaries than in the grains of nickel and the etchant used was such as to reveal several concentration contours.

Application of equation (10) showed that at 1000°C for this system

$$D_{gb}/D_l \sim 8 \cdot 10^5$$

The depth of penetration by grain boundary diffusion in this system is also greater than that by lattice diffusion by an amount which decreases markedly with increase in temperature, suggesting that Q_{gb} is again less than Q_l.

Orientation Dependence of Grain Boundary Diffusion

BECK, SPERRY and HU[34] have suggested that their observations, that grain boundary migration during recrystallization is more rapid the larger the disorientation between the grains flanking the boundary, may be explained by assuming that the rate of diffusion of atoms within the grain boundary varies with orientation. But the first direct study of the variation of grain boundary diffusion rates with relative grain orientation has been made recently by ACHTER and SMOLUCHOWSKI,[35] for silver diffusing in copper. BERMAN and HARRINGTON[36] had previously shown qualitatively for this system that silver diffused into copper more rapidly in the grain boundaries than in the adjoining grains.

For convenience in interpretation of the results, ACHTER and SMOLUCHOWSKI used specimens of "columnar" copper, in which all grains had approximately a common cubic direction and their grain boundaries parallel to this direction. Two co-ordinates completely specify the configuration of each grain boundary in such a sample,

rather than the five required in the more general case. The relative orientation of the two grains flanking the boundary is given by the angle θ between a pair of cube axes normal to the columnar direction, and the direction of the grain boundary by the angle β between one of the cube axes and the plane of the boundary. Diffusion couples of Cu

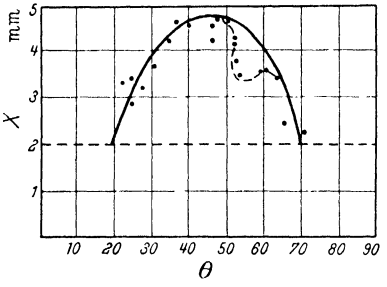


Fig. 7. Extent of Diffusion X vs. Orientation Difference
141 hr at 725°C
····· Extent of Volume Diffusion
——— Extent of Grain Boundary Diffusion

and Cu + 4 per cent Ag were welded together with the columnar direction in the copper normal to the interface, so that all grain boundaries were parallel to the direction of diffusion of the silver. After annealing, samples were quenched to retain the silver in solid solution. The maximum depth of penetration of silver down the grain boundaries was determined by etching consecutive sections lapped parallel to the interface and noting at what depth from it for each grain boundary silver, as revealed by the etchant, could no longer be detected. The amount of silver in each grain boundary decreased with increasing

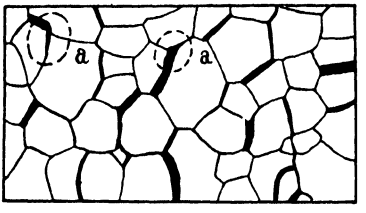


Fig. 8. Columnar Diffusion Sample cut parallel to Interface

distance from the interface and the maximum depth of penetration was found to vary with the relative orientation θ, as determined by etch pit methods, in the interesting manner shown in Fig. 7. The appearance of a typical etched section is shown in Fig. 8. At angles $\theta < 20°$ and $> 70°$ the silver does not penetrate along the grain boundaries any further than it does into the grains by lattice diffusion alone, but for $\theta >$ about 20° the grain boundary penetration depth exceeds the lattice

penetration by an amount which increases as θ increases from 20°, reaches a maximum at 45° and then falls to zero again at 70°. The results also suggested, but only vaguely, that the preferential grain boundary penetration may be least where the grain boundary is symmetrical with respect to the lattices of the adjoining grains, so that $\beta = \theta/2$.

It would appear then that grain boundary diffusion rates are not simply or directly related to the grain boundary energy, since READ and SHOCKLEY[37] have shown the energy to increase continuously as θ increases from zero up to about 20°.

The results were based upon observations made on pairs of grains which deviated from the perfect columnar structure by less than 7–8° but there were many pairs[41] with an angle γ between those $\langle 100 \rangle$ directions most nearly parallel to the general columnar or diffusion direction, which was appreciably greater than 8°. At a junction of three grains in a section perpendicular to the diffusion direction it often appeared that silver at a silver rich boundary had diffused perpendicularly to the columnar direction a little way into a silver free boundary as at (*a*) in Fig. 8. Such situations enabled a comparison to be made of diffusion rates in two mutually perpendicular directions in one grain boundary. It appeared[41] that even when θ was less than 20°, so that no preferential grain boundary diffusion occurred along the columnar direction, an appreciable diffusion perpendicular to the columnar direction in the same boundary could occur provided γ was greater than 20°.

These results reveal not only a dependence of D_{gb} upon the relative orientation of the grains but also, for fixed orientation, upon the direction within the grain boundary, i.e. an anisotropy of grain boundary diffusion.

In an earlier series of experiments using a slightly different method, ACHTER and SMOLUCHOWSKI[38] estimated the activation energy for grain boundary diffusion of Ag in Cu of about 23,800 cal/mole compared to the lattice diffusion activation energy estimated at the same time, of about 38,300.

Quite recently all this work has been criticized by DAWSON,[39] who suggests that it may only be a study of grain boundary precipitation as a function of solute supersaturation and orientation difference. He supposes that the silver diffuses into the copper everywhere by normal lattice diffusion and that on cooling the grain boundaries merely act as preferred sites for precipitate nucleation. However, ACHTER and SMOLUCHOWSKI[40] point out that the manner in which the silver is distributed in and around the grain boundaries is inconsistent with this viewpoint as also is the observed anisotropy and low activation energy. But even more recently FLANAGAN and SMOLUCHOWSKI[42] have extended

this work to the diffusion of zinc in copper, a system free from precipitation effects but in which a variation of grain boundary penetration with orientation was observed very similar in nature to that of the silver-copper system. DAWSON's criticisms seem then unfounded. A grain boundary activation energy of 24,500 cal/mole has been estimated[42] for Zn diffusing in Cu compared to a lattice diffusion activation energy of 34,000 cal/mole.

Grain Boundary Models and Diffusion

Many of the above results have been discussed by ACHTER, FLANAGAN and SMOLUCHOWSKI[35, 42, 130] in terms of some of the various models of a grain boundary that have been proposed by other workers (see PMP.1, Chapter 3). If we suppose the boundary to consist of a suitable array of edge dislocations then an inference from the anisotropy observations is that diffusion is most rapid in a grain boundary along the lines of the edge dislocations which compose it provided there is a difference in orientation between adjacent grains of more than 20°, but no faster than ordinary lattice diffusion in a direction perpendicular to the dislocation lines. For when γ is small and $\theta > 20°$ the dislocation lines run most nearly along the columnar direction, in which the grain boundary diffusion is rapid, but when $\gamma > 20°$ and θ small, as in some pairs of grains deviating considerably from the columnar structure, dislocation lines run most nearly perpendicular to the general columnar direction and it is only in this direction now that grain boundary diffusion is rapid.

That a disorientation of as large as 20° is necessary before any preferential grain boundary diffusion becomes apparent, strongly suggests that the mobility of atoms along individual edge dislocations is not appreciably different from their normal lattice mobility, for provided the angle is less than about 20° the dislocations composing the grain boundary are still individually distinguishable and separate from one another. But the angle of 20° at which the grain boundary rate begins to be appreciably greater than that of lattice diffusion, is just in the range where the distance between dislocations becomes so small that their individuality is lost; the grain boundary is now better represented by such a model as that of MOTT,[43] who pictures it as consisting of islands of good lattice fit surrounded by regions of misfit, which we may regard perhaps as formed by the grouping together of now individually indistinguishable dislocations. This change in structure of the grain boundary as θ, increasing, passes through a value of about 20° has been demonstrated by LOMER and NYE[44] with the BRAGG bubble model, and appears to be made manifest by the grain boundary diffusion behaviour. While the isolated regions of misfit described as single dislocations provide no special facilities for enhanced grain boundary diffusion, because perhaps the void space associated with them is too small to play the role

of vacancy and permit passage of a diffusing atom, (we should not then expect such sub-structural features of single crystals as mosaic boundaries to influence diffusion rates) the more extensive regions of misfit, which they combine to produce at larger angles, do allow more rapid diffusion because perhaps of their increased void space. Twin boundaries are regions of good lattice fit even if the relative orientation of the twin grains is large, and accordingly it is found that the diffusion rate in them is less than corresponds to the relative orientation, as shown by the cusp in Fig. 7 at angles corresponding to twinning on {210} planes. BARNES[33] found no evidence for any increased diffusion of copper at the presumably {111} twin boundaries of nickel. The islands of good fit will be elongated, and the arrears of misfit, therefore, less interrupted, along the same direction as that of the dislocation lines at lower angles of disorientation, so that grain boundary diffusion will be preferred along this direction and the anisotropy as observed at angles greater than 20° is still explicable.

It will be clear that FISHER's assumption (2), and therefore his equation (3), is most nearly valid for large angle boundaries where the regions of misfit will be extensive flat areas covering a large proportion, if not all, of the grain boundary. Such boundaries are believed[26] to have been predominant in the polycrystalline silver used by HOFFMAN and TURNBULL in the experiments described above. At smaller angles the regions of misfit will be less extensive, and for θ only a little greater than $\sim 20°$ may be regarded[132] as rods or channels of distorted structure, separated by relatively undistorted areas. D_{gb} and Q_{gb} are now to be taken as referring to diffusion in the distorted misfit regions only of a grain boundary. Now the maximum depth to which material penetrates by diffusion down a grain boundary is determined both by D_{gb} and by the rate, which depends on D_l, at which material is lost from the boundary into the adjoining grains. For low angle grain boundaries this loss of material, being in all radial directions from each of the rod shaped regions of misfit, will, relative to the flux of material down the misfit region themselves, be greater than for high angle grain boundaries where the loss is principally by planar diffusion from the extensive flat areas of misfit. The maximum depth of penetration, for this reason alone, will therefore increase with increasing values of θ, above a certain value, as shown in Fig. 7.

Since $Q_{gb} < Q_l$, the rate of loss of material from the boundary increases more rapidly with temperature than the rate of flow down the boundary itself. Thus the *apparent* activation energy, calculated from measurements of the variation with temperature of the maximum depth of penetration, will be less than the true activation energy Q_{gb} and will vary with the structure of the boundary, i.e. with the angle θ, being lowest and possibly even negative for the lower angle boundaries

($\sim 20°$), where the loss into the grains is relatively more important. For $\theta < \sim 20°$, of course, where there is no preferential grain boundary diffusion, the apparent activation energy equals Q_l.

FLANAGAN and SMOLUCHOWSKI[42] have recently studied the temperature dependence of the grain boundary diffusion of zinc in copper and have found an orientation dependence for the apparent activation energy of just the type expected from the above described model. Taking into account the changing structure of the grain boundary with θ, SMOLUCHOWSKI[130] has presented a theoretical discussion of grain boundary diffusion, with particular reference to the apparent activation energy, and obtained values for this which vary with θ in a manner closely agreeing with that observed in the zinc in copper experiments.

Little consideration has yet been given to the precise mechanism of diffusion in a grain boundary. Although the tendency is to discuss this in terms of the vacancy mechanism, the distinction between the various mechanisms which have been proposed for lattice diffusion is much less clear in the disorganized regions of a grain boundary; it may be quite impossible for example to distinguish in the boundary between an interstitial position and a vacant lattice site.[35] The observed anisotropy and the nature of the variation of D with orientation indicate that the increase in diffusion rates at grain boundaries arises mainly from the effect their changed structure has on the mobility of the atoms composing them rather than that it is principally due to the grain boundaries providing a higher concentration in their neighbourhood than elsewhere of, say, vacancies, and so increasing diffusion rates over a region on either side of the grain boundary more extensive than that over which the lattice is irregular. In the latter case δ would be appreciably greater than the rough value of $5 \cdot 10^{-8}$ cm usually assumed for it.

It was at one time suggested[45] that the same unit atomic process was responsible for grain boundary slip, for constant rate creep in single crystals and for volume diffusion because their activation energies appeared to be the same, at least for α-Fe, α-brass and Al. The unit process considered was an exchange or re-arrangement of atoms in what KÊ described[61] as "disordered groups" of atoms. These exist, he supposed, in certain concentration within the grains and in much higher concentration where they formed grain boundaries. Since they would also give rise to grain boundary diffusion, the inference is drawn that Q_{gb} should also be equal to the activation energy for the three above-mentioned processes. Of particular interest is the predicted equality of Q_{gb} and Q_l. But apart from the fact that experiment shows Q_{gb} to be always quite appreciably less than Q_l, a recent and more accurate redetermination of the activation energy for self-diffusion in iron[46] and a study of self-diffusion[47] and of grain boundary slip[48] in tin show that

the activation energies for Kê's three processes are certainly not equal for tin and iron, the activation energies for self-diffusion being appreciably less than those for grain boundary slip and for creep. These processes are not then all intimately related to one another through a common unit process, or to grain boundary diffusion and there is no reason for expecting Q_{gb} to equal Q_l. Should they be equal then differences between D_{gb} and D_l would reside wholly in the constant term in the Arrhenius equation; a grain boundary model which would then explain the observed differences has been discussed by Le Claire.[30]

It has often been suggested that impurities in metals may have a tendency to segregate in grain boundaries and the first convincing demonstration of a case where this is so has recently been given[125]—the case of Po impurities in PbBi alloys. The polonium impurity was produced by decay of radioactive Bi incorporated in the alloy and although present at as low a concentration as 1 part in 10^{10}, was shown by autoradiography to be present in the alloy after annealing in much higher concentration at the grain boundaries than elsewhere. Impurities which do segregate in this way will most probably have an effect on grain boundary processes out of all proportion to their total concentration. Thus Kê found[60] that very small amounts of impurity completely block grain boundary slip in aluminium, although Hoffman and Turnbull[29] found little change with purity in the observed grain boundary self-diffusion rates in silver. Birchenall and Mehl,[49] however, found that while at high temperatures the self-diffusion coefficients in polycrystalline samples of iron of various purities were substantially the same, at low temperatures it was greater in the iron of commercial purity than in more pure material. They suggest that perhaps in a pure metal D_{gb} and D_l are normally the same but that impurities accumulating in grain boundaries have the effect, at least in commercial iron, of increasing D_{gb} and so also the apparent low temperature self-diffusion coefficient. But these results are not conclusive because they had no quantitative evidence on the relative grain sizes of the various purity materials; the apparently smaller grain size in the less pure iron, if in fact sufficiently smaller than that of the pure iron would equally well explain the results.

It would be surprising if impurities alone were responsible for the increase in grain boundary diffusion over lattice diffusion rates but their importance in affecting values of D_{gb} is a matter which merits further and more extensive investigation.

Surface Diffusion

A number of interesting experiments have demonstrated the very rapid mobility of surface metal atoms and summarizing reviews[26, 92] of these are already available in the literature. In a very few cases quantitative

estimates of D_s and of Q_s have been made which show Q_s to be less than both Q_{gb} and Q_l, and D_s to be greater than D_{gb} and D_l which from the nature of the process is not unexpected. In the last few years some further contributions have been made in this field which generally confirm these results.

It will be clear that FISHER'S analysis[27] of grain boundary diffusion is, with similar assumptions, directly applicable to surface diffusion if in his equations D_{gb} is replaced by D_s and δ by $\frac{1}{2}\delta_s$, δ_s being the effective width of the diffusing surface layer. The factor $\frac{1}{2}$ arises because material is, of course, only lost by lateral diffusion into the grain from one side of the surface layer. But no use yet appears to have been made of this analysis in surface diffusion experiments.

NICKERSON and PARKER[51] have made measurements on the surface self diffusion of silver and obtained the result

$$D_s = 0{\cdot}16 \exp(-10{,}300/RT) \text{ cm}^2/\text{sec} \quad . \quad . \quad . \quad . \quad (11)$$

Radio active silver was deposited at one small region of a bundle of silver wires and the surface diffusion coefficient calculated from the distribution along the wire of the active silver after annealing at various temperatures between 225 and 350°C. To measure the distribution the wire was scanned with a counter tube fitted with a lead collimator and thin slit so that at each position the activity of only a small section of the wire was measured. However, the results were analysed without regard to loss of material from the surface into the grains: the wires used were, in the absence of any indication to the contrary, presumably polycrystalline so that surface atoms would also be lost into grain boundaries. For these reasons and in view of the considerable scatter among the results equation (11) represents only a rough estimate of D_s. However, taken with HOFFMAN and TURNBULL'S results we now have a more complete account of self-diffusion processes in Ag than in any other metal.

A similar method has been used by FRAUNFELDER[52] to study the surface diffusion of copper (Cu^{64}) on silver with the single result that at 750°C, $D_s = 7 \cdot 10^{-2}$ cm^2/sec.

MÜLLER[53] has used an ingenious method to study the surface self-diffusion of tungsten. The potential difference required to maintain a given electron emission from a fine tungsten point is a function of its radius of curvature. Changes in shape occur when the point is maintained at a high temperature (1600 to 2500°C) and assuming such changes are due to surface diffusion MÜLLER has calculated from the change in potential with time that the activation energy for surface self-diffusion of tungsten is 106,500 cal/mole. This is appreciably less than the value of 143 cal/mole for lattice self-diffusion estimated by VAN LIEMPT.[54] It is interesting to note that the ratio Q_s/Q_l is much

higher for tungsten than for silver but it should be remembered that VAN LIEMPT'S estimate is not from direct radio active measurements. He merely assumes that his values of the coefficient for diffusion of iron in tungsten, being measured at very low iron concentration, represent also the self-diffusion coefficient of tungsten. His value for Q_l may then be an underestimate of the true value.

SECTION III

THE MECHANISM OF DIFFUSION

1. *Diffusion in Interstitial Solid Solutions*

It has never been doubted that the diffusion of solute atoms in interstitial solid solutions occurs by successive jumps of the solute atoms from one interstitial position to another, for it is difficult to imagine any alternative mechanism in this case.

Recently there have become available new methods for measuring interstitial diffusion coefficients (see Section IV) which extend the temperature range over which such measurements may be made and so provide more accurate values than have before been available of the theoretically important quantities D_0 and Q in the usual equation employed to express the temperature dependence of D, viz—

$$D = D_0 \exp(-Q/RT) \quad \ldots . \quad (12)$$

Prompted by this important development, renewed attention has recently been given, notably by WERT and ZENER,[56, 57, 58] to a theoretical interpretation of D_0.

Treating the process of diffusion as a random walk problem, it may readily be shown that for a cubic lattice the diffusion coefficient D for any one species of atom is given in terms of τ, the average time between successive jumps of each atom of that species, by the expression

$$D = \alpha a^2/\tau \quad \ldots . \quad (13)$$

where a is the lattice constant and α is a geometric factor depending on the disposition round any one equilibrium site of neighbouring sites into which the atom may diffuse by jumping. α is easily computed and turns out, for interstitial diffusion, to be 1/24 for a b.c.c. and 1/12 for a f.c.c. lattice. The remaining problem is to determine $1/\tau$.

WERT[57] and WERT and ZENER[56] have applied for this purpose the essential assumptions and methods of EYRING'S[62] theory of reaction rates but have adapted it to this particular problem more carefully than previous workers. The most general expression for the rate of an activated process is, according to EYRING

$$1/\tau = p \,.\, (kT/h) \,.\, F^*/F$$

where k, T and h have their usual meaning and p is the number of equivalent paths for diffusion of an atom from any one position of equilibrium (for interstitial diffusion p is 4 in b.c.c. and 12 in f.c.c. lattices). F and F^* are the partition functions for the system in its initial and activated states respectively i.e. for an atom in its normal equilibrium position and at the top of the potential barrier separating it from the next position to which it is diffusing. Now in evaluating F^*, the degree of freedom at the saddle point corresponding to movement over the barrier is omitted. F^*/F is usually written as $\exp(-\Delta G'/RT)$, where $\Delta G'$ is referred to as the "free energy" change in passing from the equilibrium to the activated state. But this quantity is not strictly a free energy for it is the ratio of two partition functions with unequal numbers of degrees of freedom. In order to obtain a similar ratio of partition functions with equal numbers of degrees of freedom WERT and ZENER[56] write

$$F = F_\nu \,.\, F_1$$

where F_ν is the partition function for a single linear oscillator of frequency ν, this being considered to represent the vibrational motion of the atom about its equilibrium position and along the direction of impending diffusion. F_1 is then the partition function for the system with the atom constrained to move in a plane normal to the direction of diffusion and passing through the equilibrium position. For the temperatures at which diffusion measurements are made it is quite sufficient to approximate the usual expression for F_ν by $F_\nu = kT/h\nu$ so that we arrive at the result

$$1/\tau = p\nu \,.\, F^*/F_1$$

Since F^* and F_1 now refer to two states with equal numbers of degrees of freedom, we may, in writing

$$1/\tau = p\nu \exp(-\Delta G/RT) \qquad \ldots\ (14)$$

legitimately refer to ΔG as a free energy change and so equate it to the isothermal work done in moving an atom from an equilibrium position to the top of the potential barrier separating it from the next equilibrium position, the atom only being constrained during this motion to move in directions normal to the line joining the two equilibrium positions. ΔG in (14) has therefore a clear physical significance.

A derivation of equation (14) from first principles has been given by WERT.[57]

Assuming the isothermal work to be done at constant pressure we may make use of the thermodynamic relation

$$\Delta G = \Delta H - T \,.\, \Delta S$$

in writing, from equations (13) and (14), the expression for D.

$$D = p\alpha a^2 \nu \exp(\Delta S/R) \exp(-\Delta H/RT) \quad . \quad . \quad . \quad . \quad (15)$$

ΔS and ΔH are referred to respectively as the "entropy" and "enthalpy" (or heat) of activation.

We also have the additional relation

$$\left(\frac{\partial \Delta H}{\partial T}\right)_P = T\left(\frac{\partial \Delta S}{\partial T}\right)_P$$

so it follows that

$$\frac{\partial \log D}{\partial(1/T)} = -\frac{\Delta H}{R}$$

and hence that ΔH in equation (15) corresponds with Q in equation (12). D_0 is, therefore, given by the relation

$$D_0 = p\alpha a^2 \nu \exp(\Delta S/R) \quad . \quad . \quad . \quad . \quad (16)$$

WERT and ZENER estimate ν by assuming the potential energy of a solute atom to vary sinusoidally as it moves from one interstitial lattice position to another with an amplitude taken as equal to the heat of activation ΔH. With such an assumption they obtain

$$\nu = (\Delta H/2m\lambda^2)^{1/2}$$

where λ is the distance between interstitial positions and m is the mass of the atom.

The only unknown quantity now appearing in the expression for D_0 is ΔS. This may be evaluated by noting the further thermodynamic relation.

$$\Delta S = -\frac{\partial \Delta G}{\partial T}$$

If we assume that all the work ΔG appears as strain energy in the lattice surrounding the diffusing atom when at the saddle point, then we may put

$$\Delta G = (1/2)\mu e^2 v \quad . \quad . \quad . \quad . \quad (17)$$

where μ is some shear modulus of the lattice, e a measure of the shear strain, and v the volume over which the strain is apparent. The variation with temperature of that part of the experimentally observed modulus which is associated only with lattice elasticity, and which does not contain a contribution from grain boundary relaxation processes, is found to be quite closely linear.[63] Therefore, we may write

$$\mu = \mu_0 + T\partial\mu/\partial T \quad . \quad . \quad . \quad . \quad (18)$$

and if we ignore any possibly small temperature variations of v and e, we find

$$\Delta H = (1/2)\mu_0 e^2 v$$

and

$$\Delta S = -\Delta H \,.\, \partial(\mu/\mu_0)/\partial T. \qquad \ldots \quad (19)$$

$\partial\mu/\partial T$ is always found to be negative so that ΔS is essentially positive and exp $(\Delta S/R)$ always greater than one.

Inserting the known or now estimated values of p, α, ν and a^2 into equation (16) values of $\Delta S_{\text{exp.}}$ can be calculated from experimental

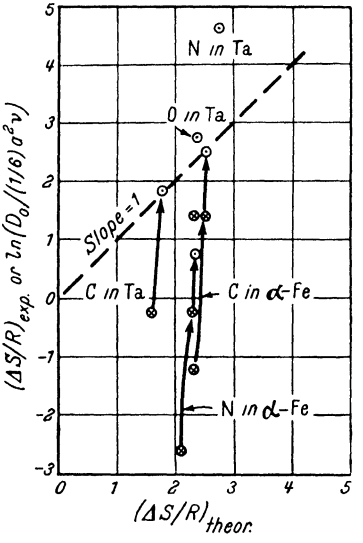


Fig. 9. Theoretical and Experimental Values of $\Delta S/R$ for Interstitial Diffusion[58]

Oxygen in Ta—KÊ[110] (1948)
Nitrogen in Ta—KÊ[111] (1948)
Carbon in Ta—KÊ[111] (1948) WERT[58] (1950)
Carbon in αFe—SNOEK[102] (1941) STANLEY[109] (1949) WERT and ZENER[56] (1949) WERT[57] (1950)
Nitrogen in αFe—SNOEK[102] (1941) WERT and ZENER[56] (1949) WERT[58] (1950)

values of D_0 and compared with the values of $\Delta S_{\text{Theo.}}$ given by equation (19). For this latter calculation WERT has used the temperature coefficients of the polycrystalline bulk shear moduli as deduced from KÖSTER'S[64] measurements for a large number of materials of the temperature dependence of Young's moduli. A comparison of theory with experiment has been made by WERT[58] by plotting $\Delta S_{\text{exp.}}$ against $\Delta S_{\text{Theo.}}$ as in Fig. 9. Full agreement with the foregoing theory is indicated for those points which lie on the line of slope one. The most

accurate experiments to date are those for C in α-Fe[57] in view of the very extensive temperature range (— 40 to 800°C) over which they have been made, and in this case the agreement is excellent. WERT points out that both for this and other cases, repetition of measurements by different observers appears to lead to continuously improving agreement between experiment and theory and so by implication attributes remaining discrepancies to there still being room for improvements in the accuracy of D_0. The arrows in Fig. 9 run from earlier to later experiments.

In an earlier paper, WERT and ZENER[56] consider the possibility that not all the energy ΔG may go into straining the lattice in which case ΔS may be allowed to take any value between zero and that given by equation (19). The experimental values suggest, however, that ΔS always takes its full value.

This work on interstitial solutions represents a considerable advance in the subject of diffusion, being the first successful theoretical interpretation of observed values of D_0. Its success is due in no small measure to the recent availability of more accurate values of D_0 for previous attempts[1] in this direction have been made by use of an equation very similar in form to number (15). The equation for $1/\tau$ was often derived simply by multiplying the probability that an atom had the activation energy necessary to carry it over the potential barrier, as given by the Boltzmann factor, by the number of times per second the atom was moving in the right direction to diffuse, i.e. ν. Since the total energy of activation rather than the free energy appeared in these expressions, the equations resulting did not contain the term $\exp(\Delta S/R)$. But attempts were sometimes made to remove the discrepancy between experiment and theory which thereby resulted by taking explicitly into account the temperature dependence of the total energy of activation. When the energies involved vary linearly with temperature, such a procedure leads to much the same expressions for D_0 and Q as the above and so with reliable data to work on, and perhaps in some cases a little more careful attention to geometric factors, could have succeeded in relating theory with experiment. However, it may be shown[65] to be thermodynamically quite unsound to employ Boltzmann factors containing total energies (in the sense of the first law of thermodynamics) which are functions of temperature; the correct procedure in such cases is to replace the total energies normally appearing in the exponential by the corresponding free energies of the states to which they refer. Although the derivation of $1/\tau$ which we have discussed is the more rigorous and much to be preferred, the earlier simple kinetic derivations would have produced the same results, without recourse to thermodynamically unsound practices, had the correct form of the Boltzmann factor been employed.

2. *Diffusion in Substitutional Solid Solutions and Pure Metals*

It is by no means so obvious in this case by what atomic process diffusion occurs for even on casual reflection several possibilities occur. It becomes necessary then to consider these possibilities in turn and to choose from them, as a model upon which to base discussion, the one, the consequences of which can be related the more closely to experimental measurements and observations on diffusion phenomena.

(*a*) *Comparison of Observed and Calculated Activation Energies*—The three types of process which have been the subject of most discussion in the past are the direct exchange mechanism, the vacancy mechanism, and mechanisms involving atoms in interstitial positions. These were described and discussed in PMP.1, where it was pointed out that while the vacancy mechanism had often been preferred in the past because it was felt that it would require a much lower energy of activation than the other mechanisms, this feeling was only confirmed, at least for the close packed metal copper, by HUNTINGTON and SEITZ's[66, 67] direct quantum mechanical calculation of the activation energy Q for all three processes. (In the particular type of interstitial mechanism chosen for consideration by these authors, the interstitial atoms, assumed present in very small concentration in thermodynamic equilibrium with the lattice, diffuse by moving on to an adjacent substitutional site, pushing the atom occupying it on to the next interstitial site. This has been termed "Interstitialcy Diffusion" by SEITZ[15] to distinguish it from the more usual concept of interstitial diffusion, where the atoms move directly from one interstitial position to the next, and which would require a larger activation energy.)

Their estimated values of Q were,[66] in kcal/mole, 400 for the exchange mechanism, 230 for the interstitialcy mechanism, and, with a more refined calculation,[67] 64 for the vacancy mechanism. When these are compared with the experimentally observed value of ~ 50 for self-diffusion in copper, it becomes clear that the vacancy mechanism is by far the most likely of the three. A recent re-estimation[68] of the value for the exchange mechanism reduces it to about 240 and so does not vitiate this conclusion. The inference is usually drawn from these results that for all close packed lattices the vacancy mechanism represents most closely the actual process given rise to diffusion in them. But the calculations allow of no similar conclusion regarding the mechanism in non-close packed lattices. In fact, since it is the ion core repulsions which play the dominant part in rejecting the possibility of the exchange and interstitialcy mechanisms in such a metal as copper, these mechanisms would be in much closer competition with the vacancy mechanism in calculations on less close packed metals, where the ion core repulsion is much reduced.

For a long time serious consideration was given only to the three above-mentioned mechanisms, but recently other mechanisms have been suggested representing more generalized versions of these. Thus, for example, ZENER[69] has pointed out that the simple exchange mechanism is only a special case of what he terms "ring diffusion" in which a number n of atoms lying on a ring, so that each is a nearest neighbour of the one before and after it, diffuse simultaneously by synchronous movement round the ring. Fig. 10 illustrates this for $n = 4$ in a f.c.c. and b.c.c. lattice. The exchange mechanism is clearly the simplest case of $n = 2$. ZENER has carried out detailed calculations, similar to those of HUNTINGTON and SEITZ, of the activation energy required for ring-diffusion processes in copper. He shows that contrary

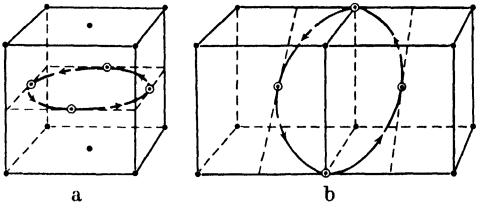


Fig. 10. 4-Ring Diffusion[69]
(*a*) f.c.c. lattice
(*b*) b.c.c. lattice

to what might at first sight be expected the activation energy does not increase with n, but is appreciably less for $n = 4$ than for other values. Although more atoms are involved in a 4-ring than in a simple exchange or 2-ring process, the disturbance attending it in the surrounding lattice and hence the energy required for its performance, is less. In fact, the coulomb interaction of the closed shells at the saddle point is almost negligible and it is only the exchange interaction which contributes significantly to the activation energy. ZENER finds this to be $\sim$ 90 kcal, an estimate which, although very considerably nearer than those for the simple exchange and interstitialcy mechanisms to the observed value of $\sim$ 50 kcal/mole, is still by no means in as close agreement with it as the estimate of 64 for the vacancy mechanism: and such refinements to the calculations which might be expected to bring it closer to the experimental value would, at the same time tend further to improve the agreement with experiment of the estimate for vacancy diffusion.[15] There would then appear to be no very strong grounds for replacing the vacancy by the ring mechanism as the most likely process of diffusion in close packed metals. As we shall see shortly there is in fact stronger evidence against the ring mechanism: but in the absence both of such evidence and similar calculations for less close packed metals ring diffusion must remain in their case a possible process to be considered.

PANETH[70] has calculated the activation energies for diffusion by the above-mentioned processes in the alkali metals and in this way has been led to consider yet another process of diffusion representing a generalization of the interstitialcy process. He found that while the energy to form an interstitial atom in sodium is very much less than for copper, the energetically most favourable position for the interstitial atom is actually midway between a pair of nearest neighbours, for an atom in such a position produces a very considerable amount of relaxation by the neighbours and by atoms lying along the line of the neighbours. A defect is thereby produced, termed a "crowdion" which may be described as a row of n nearest neighbour atoms compressed so as to occupy a distance normally occupied by $n - 1$ atoms, or as PANETH puts it, as an interstitial atom diluted over several atomic spacings. Diffusion then occurs by propagation of the region of compression along the direction of the crowdion and PANETH shows that a crowdion of about 8 atoms has a much lower energy of activation for diffusion than either the vacancy or ring mechanism. PANETH'S estimate of c . 7 kcal/mole for sodium awaits experimental confirmation.

By way of a generalization of the simple vacancy model we may mention that SEITZ[22] has discussed the possibility that diffusion may occur by the migration of coupled pairs of vacancies and suggests that this may be faster than diffusion by single vacancies because of the lower closed shell interaction during the exchange of an atom with either of the vacancy pair. Such double vacancies are thought[71] to play a role in the changes in resistance of metals on annealing after plastic deformation, and may be of importance in chemical diffusion on account of the Kirkendall effect, as mentioned in Section I.

(*b*) *Evidence from the Kirkendall Effect*—A necessary and important property of the exchange or ring diffusion mechanisms is that they can never lead to any nett displacement of atoms relative to the crystal lattice or to any changes in volume or shape as a result of diffusion, other than those arising from differences in the atomic sizes of the interdiffusing species. The work of SMIGELSKAS and KIRKENDALL, reviewed in PMP.1 and many more recent experiments designed mostly to confirm and extend it and reviewed in Section I, have together demonstrated that all these forbidden phenomena are, in fact, associated with diffusion, at least in f.c.c. lattices, and so provide the most direct evidence to date against an exchange or ring mechanism in such structures. A Kirkendall effect has now been observed in about nine different systems and as these are all of the face centred cubic structure we may probably assume that a ring mechanism is most improbable in all alloys of this structure. No Kirkendall effect experiments have yet been performed which would assist in elucidating the mechanism of diffusion in alloys of other structures. As shown in Section I the

vacancy mechanism does provide a satisfactory picture of all the phenomena associated with the Kirkendall effect, and although an interstitial mechanism might also be invoked to do the same, the calculations referred to in the last section, with the evidence mentioned here, provide very strong evidence for a vacancy mechanism.

(c) *Miscellaneous Evidence Relating to a Vacancy Mechanism*—It has often been proposed that if diffusion occurs by vacancies then since the equilibrium concentration of vacancies increases rapidly with temperature it ought to be possible to quench a metal so as to produce a concentration of vacancies in excess of that corresponding to equilibrium at the temperature to which the material is quenched. But it appears from studies[15] of the problem that the normal density of dislocations in a metal, which can act as sources and sinks of vacancies, is, at the temperatures at which conventional diffusion experiments are made, sufficient to restore equilibrium in a time very short compared with that required for the measurements. The excess vacancy concentration in such quenched materials will not then lead to an abnormally high diffusion coefficient if one attempts to measure this by conventional means. Thus quite recently HOFFMAN and TURNBULL[29] found no difference between the self-diffusion coefficients of silver samples whether these were air quenched or slowly cooled before the experiments.

However, the recently developed anelastic methods (see Section IV) which allow measurements of diffusion coefficients at much lower temperatures and in shorter times than by the conventional methods, have been used by NOWICK[72] successfully to demonstrate that D is, in fact, larger in a quenched alloy (of 29 atoms per cent Zn in Ag) than in the same alloy more slowly cooled. The method is such, of course, that any number of diffusion experiments may be made on the same sample and so one may follow the rate at which the increased diffusion coefficient falls to the normal value appropriate to slowly cooled or annealed alloys, as the vacancy concentration returns to its equilibrium value. In this way NOWICK estimates that ΔH_1, the activation energy for the formation of vacancies, is $\sim$ 12·5 kcal/mole and ΔH_2, that for their exchange with atoms, $\sim$ 20·4 kcal/mole, corresponding to the total activation energy Q for diffusion of 32·9 kcal/mole. It is interesting in connexion with remarks in the next part of this section that $\Delta H_1/Q \simeq 0{\cdot}4$.

Thus is confirmed a necessary consequence of the supposition that diffusion occurs by vacancies.

There are many properties of metals which are intimately related in some way to their diffusion behaviour so that a study of such properties may be of use in furthering a knowledge of the diffusion processes themselves. By assuming that vacancies exist in the lattice and that diffusion occurs by the migration of vacancies, many of these properties have been successfully interpreted using the known values of diffusion

coefficients, or conversely have provided diffusion data agreeing with that obtained by direct measurement. Such work then strengthens the evidence in favour of the vacancy mechanism in the f.c.c. structures for which these other properties have been successfully so interpreted. The work on sintering and on the creep of wires will be discussed in Section IV. These involve changes of shape at rates determined by diffusion processes and so like the Kirkendall effect rule out the possibility of an exchange or ring mechanism. It has been suggested[73, 74] that the increase in electrical resistance of metals after cold work is due to the generation of an increased vacancy concentration by the movement of the dislocations. MANINTVELD[71] has recently studied the rate of return of the resistance to its normal value on annealing, a rate determined by the migration rate of the vacancies, and has calculated the activation energies for the movement of single vacancies in Au, Ag and Cu to be 16, 15 and 20 kcal/mole, respectively. Although appreciably less than half the corresponding total activation energies for diffusion, these are not unlikely values. The same experiments provide possible evidence for migration of pairs of vacancies, with activation energies rather less than those for single vacancies.

There is also an increase in resistance of pure metals after neutron irradiation. Since such treatment can only produce vacancies or interstitials, it is most interesting that the activation energy for the recovery of the resistance on annealing turns out, at least for platinum,[131] to be equal to that for recovery of the resistance produced by cold work in the same metal.

It should be mentioned that such experiments as these constitute "evidence" for a vacancy mechanism only because the calculations of HUNTINGTON and SEITZ show an interstitial mechanism to be most unlikely on energetic grounds. In the absence of such calculations, NOWICK's experiments, for example, could as readily be explained on the supposition that diffusion occurs by any mechanism involving lattice defects whose concentration increases with temperature.

Attempts have also been made[75, 76] to obtain evidence for a vacancy mechanism by studying diffusion in alloys such as FeAl, NiAl and CoAl which are known to possess a defect structure in that in certain ranges of composition a large number of lattice sites are vacant. Thus, for example,[77] on the cobalt rich side of approximately 50–50 atomic per cent b.c.c. alloys of CoAl, the cobalt atoms replace aluminium, atom for atom, but on the aluminium rich side, although the b.c.c. structure is maintained, the number of atoms per unit cell is less than two and many sites are in consequence vacant. If such vacancies facilitate diffusion then a marked increase in the diffusion coefficient should occur as the concentration passes through 50 per cent towards higher aluminium content. Such an increase has in fact been observed by NIX and

JAUMOT[76] for the diffusion of cobalt in CoAl alloys. BURGESS and SMOLUCHOWSKI,[75] however, found no similar increase in Ni Al alloys and attributed this to the inability of the vacant sites in such alloys to play the role required of them in diffusion on account of their small size compared to that of the aluminium atom.

One would expect a similar effect with no increased diffusion in the Co/Al alloys and so the evidence from these experiments is not very helpful, but it is in any case difficult to see why such experiments could be taken as providing evidence for a vacancy mechanism except in these rather special defect alloys. Any increase in diffusion in the concentration range where the alloys are defective may certainly be due to a vacancy mechanism, but this throws no light on the mechanism in the non-defective structures, which, being body centred cubic, is much needed.

(*d*) *Theory of* D_0—Diffusion coefficients for self-diffusion and for chemical diffusion in substitutional solid solution are well known to vary with temperature through the Arrhenius type equation (12). Just as in Section 2 (*a*) a comparison with experiment of the values of Q to be expected from the various proposed models for diffusion has enabled a choice to be made between these models, so in general may we expect it to be possible to make or confirm the choice by comparison with experiment of observed values of D_0 for the various models.

The most forbidding feature of the reported values[78] of D_0 for chemical diffusion is that they cover the alarming range of 10^{-9} to 10 cm^2/sec. For self-diffusion D_0 lies, however, in the far more moderate range of 10^{-1} to 10 cm^2/sec and so for this reason and because it is the simplest possible type of diffusion in metals, it is natural that attempts should first be made to understand the values of D_0 associated with it.

A careful reassessment[79, 80] of the reported work on self-diffusion measurements has provided what are on the evidence available, the most likely values of D_0 and Q for the f.c.c. metals Au, Ag, Pb and Cu. An interpretation[79, 81] of this data by a method which is an extension of that used so successfully by WERT and ZENER[56–58] for interstitial diffusion, has shown that the values of D_0 for self-diffusion in these metals are consistent only with, and, therefore, confirm the operation of, the vacancy model of diffusion. The same methods applied to the b.c.c. metals suggest that self-diffusion occurs in them by a 4-ring diffusion process.

(i) *Self Diffusion*—Equation (13) for D holds equally well for vacancy diffusion provided it can be assumed that successive jumps of any individual atom are random in direction and independent of one another. This assumption, valid certainly for dilute interstitial solid solution, is not precisely valid for vacancy diffusion for an atom which has just exchanged with a vacancy will be more likely to make its next

move back to its original position rather than in any other direction. The error involved in making it ($\sim$ 10 per cent)[20] is well within the error of most experiments and so at present unimportant. α in equation (13) is readily shown to be 1/12 for an f.c.c.[105] and 1/8 for a b.c.c. lattice.

If β is the fractional concentration of vacancies, then in an f.c.c. lattice 12β is the probability that an atom has a vacant site as neighbour. Only then, of course, may the atom move.

If n' is the rate of jumping of an atom into an adjacent vacant site, then clearly

$$1/\tau = 12\beta n' \quad . \quad . \quad . \quad . \quad (20)$$

where τ is as previously defined (p. 295). Similarly for a b.c.c. lattice

$$1/\tau = 8\beta n' \quad . \quad . \quad . \quad . \quad (21)$$

If ΔG_1 is the change in free energy on forming a vacancy at constant temperature and pressure, apart from that contributed by changes in mixing entropy, then the equilibrium concentration of vacancies is readily shown to be

$$\beta = \exp(-\Delta G_1/RT) \quad . \quad . \quad . \quad . \quad (22)$$

n' may be evaluated by methods identical with those used by WERT[57] for interstitial diffusion, when we find

$$n' = \nu \exp(-\Delta G_2/RT) \quad . \quad . \quad . \quad . \quad (23)$$

where ν is the frequency of vibration of an atom on a lattice site and ΔG_2 is an amount of isothermal work defined exactly as on p. 296. Using the same thermodynamic methods and relations as for the interstitial case we find from equations (13) and (20) to (23) that for self diffusion in both b.c.c. and f.c.c. lattices,

$$D = a^2\nu \exp[(\Delta S_1 + \Delta S_2)/R] \exp[-(\Delta H_1 + \Delta H_2)/RT] \quad . \quad . \quad (24)$$

where ΔS_1 and ΔH_1 are the entropy and enthalpy associated with the free energy ΔG_1, and similarly for ΔS_2 and ΔH_2 . $(\Delta H_1 + \Delta H_2)$ corresponds to Q, the observed activation energy, and the constant term D_0 is given by

$$D_0 = a^2\nu \exp[(\Delta S_1 + \Delta S_2)/R] \quad . \quad . \quad . \quad . \quad (25)$$

As before, if the work ΔG_2 appears as strain energy in the lattice ΔS_2 will be positive. And also[81] ΔS_1 will be positive for when a vacancy is formed in the lattice the neighbouring atoms acquire a greater freedom of movement so that their entropy must increase. Thus the total entropy of activation $(\Delta S_1 + \Delta S_2) = \Delta S$ must be positive. ZENER[81] points out that this is, in fact, true for self-diffusion in all cubic metals.

As we have seen, the calculations of HUNTINGTON and SEITZ provide values of ΔH_1 and ΔH_2 agreeing quite closely with experiment, but no

such corresponding calculations, necessary to complete the theoretical interpretation of D, have yet been made for $\Delta S_1 = -\partial \Delta G_1/\partial T$ and $\Delta S_2 = -\partial \Delta G_2/\partial T$. Very rough estimates have from time to time been made (see PMP.1) and do produce values of D_0 in order of magnitude agreement with experiment. An exact calculation of ΔS_1 and ΔS_2 would be very difficult, but we may, as in the interstitial case, at least obtain an expression for ΔS in terms of ΔH, although the analysis is more complicated here by there being two terms in the activation energy. However, we proceed in a similar way and propose a correlation between both ΔG_1 and ΔG_2 and other physical properties of the lattice to which we might expect these quantities to be most closely related. Thus we write again equation (17) for ΔG_2. For geometrically similar lattices e will be the same and we may suppose v to be proportional to the atomic volume M/ρ, where M is the atomic weight and ρ the density, so that we may write

$$\Delta G_2 = k_2 M \mu/\rho \quad . \quad . \quad . \quad . \quad (26)$$

where k_2 is a constant for metals of one structure. ΔH_1 is the change in internal energy on forming a vacancy and it has often been proposed that this will be closely related to the latent heat of sublimation L_s, so we put[82]

$$\Delta H_1 = k_1 L_s \quad . \quad . \quad . \quad . \quad (27)$$

Assuming μ and ρ to vary linearly with temperature we then find, using the equation $\Delta S = -\partial \Delta G/\partial T$, that

$$Q = k_1 L_s + k_2 M \mu_0/\rho_0 \quad . \quad . \quad . \quad . \quad (28)$$

$$\Delta S = \left[\Delta S_1 - k_2 \frac{M \mu_0}{\rho_0} \{(\mu/\mu_0)' - (\rho/\rho_0)'\}\right] \quad . \quad . \quad . \quad . \quad (29)$$

where the subscript $_0$ refers to the constant term in the linear expressions for μ and ρ as in equation (18) for μ, and the dash denotes differentiation with respect to temperature. The second term in (29) is negative, so that ΔS remains positive.

We may eliminate k_2 from (29) using (28) and so find for ΔS

$$\Delta S = [\Delta S_1 - (Q - k_1 L_s)\{(\mu/\mu_0)' - (\rho/\rho_0)'\} \quad . \quad . \quad . \quad . \quad (30)$$

The experimental value of ΔS is given by $R \log (D_0/a^2 \nu)$ so k_1 may be calculated for each of the cubic metals for which D_0 and Q are known by inserting in equation (30) appropriate values for the other quantities. ΔS_1, the change in entropy on forming a vacancy, arises principally from the lowering of the vibrational frequency of the atoms surrounding the vacant site, but a proper calculation of this frequency change does not appear to have been made. It is to be noted, however, that the local values near a vacant site, certainly of μ and most probably of $(\mu/\mu_0)'$

will be less than the bulk values, and these changes in elastic properties and in ν will be closely related. It is the local values of $(\mu/\mu_0)'$ which should of course be used in (30). But since k_1 derived from equation (30) decreases with decreasing ΔS_1 and increases with increasing $(\mu/\mu_0)'$ we shall assume for simplicity that k_1 is unaffected if we put $\Delta S_1 = 0$ in equation (30) and employ in the calculations the bulk value of $(\mu/\mu_0)'$. For ν we may take the Debye frequency. Using for $(\mu/\mu_0)'$ the same source of data as employed in Part 1 of this section, it is found[79] that k_1 is very appreciably the same for the four f.c.c. metals for which all the relevant data exists, as shown in Table 2.

TABLE 2

	Pb[84]	Ag[29, 31]	Au[83]	Cu[80]		α-Fe[46]	W[54]
D_0 cm^2/sec .	4·02	0·895	0·157	0·178		5·8	11·5
Q $kcal/mole$.	27·5	45·9	53·0	46·5		59·7	142
k_1 . .	0·21	0·21	0·20	0·23	Mean $k_1 = 0·21$	− 0·075	− 0·147
k_2 . .	0·23	0·22	0·25	0·18	Mean $k_2 = 0·22$	—	—

The constancy of k_1 for the f.c.c. metals justifies the correlations (26) and (27) and the assumption of a vacancy mechanism on which they were based. The mean value of k_1 gives for Cu, for example, 17·4 kcal/mole as the energy required to form a vacancy. This, being about 0·4 of the total activation energy is a very reasonable value, although a little lower than the value of $0·64Q$ estimated by HUNTINGTON and SEITZ.[66, 67] But since both this and the present estimate are only approximate, the discrepancy does not invalidate the present analysis.

However, we see from Table 2 that the values of k_1 for the b.c.c. metals are not only well removed from those for f.c.c. but are in addition negative; a negative k_1 in equation (27) is meaningless. This may indicate inaccuracies in the measurements for W (see also p. 295) and α-Fe or that in their case a vacancy mechanism of diffusion does not operate. It is very interesting that the observed values of D_0 and Q for these two metals can be satisfactorily related only on the assumption that diffusion occurs in them by a 4-ring exchange mechanism.

The diffusion coefficient for self-diffusion by a ring mechanism may be shown[81] to be given by

$$D = \gamma \nu a^2 \exp(-\Delta F/RT)$$

where γ is an integer always greater than 1 and depending on the number of the ring and upon the type of lattice. ΔF is the free energy of activation for the ring diffusion process and, as before, we may put

$$\Delta F = kM\mu/\rho$$

and so find

$$\Delta S_{\text{ring}} = Q\{(\mu/\mu_0)' - (\rho/\rho_0)'\} = R \log (D_0/\gamma a^2 \nu)$$

Thus γ may be calculated from the equation

$$R \log \gamma = R \log (D_0/a^2\nu) - Q\{(\mu/\mu_0)' - (\rho/\rho_0)'\} \quad . \quad . \quad . \quad . \quad (31)$$

and the values turn out, for f.c.c. lattices, to be not only far from constant but all small fractions inconsistent with the supposition of a ring mechanism. But for W and α-Fe they are greater than 1 and both closer to the value of γ (= 6) for a 4-ring diffusion process than for any other numbered ring.

An analysis similar to the above, with the same type of assumptions, has been carried out[79] for interstitial diffusion but fails to provide values of ΔS and, therefore, D_0, which are consistent with experiment either for b.c.c. or f.c.c. metals.

The methods outlined here, which avoid any direct calculation of D_0 show that observed values of it may be satisfactorily interpreted on the basis only of a vacancy mechanism for f.c.c. metals and of a 4-ring mechanism for b.c.c. metals. Both results are quite consistent with previous evidence on the mechanism of diffusion in these structures.

Equation (30) may be written in a simpler form by noting, as was pointed out by STEIGMAN, SHOCKLEY and NIX[119] sometime ago (see also ref. 31), that the ratio Q/L_s is fairly constant from metal to metal and of mean value about 0·6. Putting $L_s = Q/0{\cdot}6$ in equation (30) we find

$$\Delta S = 0{\cdot}65Q\{(\mu/\mu_0)' - (\rho/\rho_0)'\} \quad . \quad . \quad . \quad . \quad (32)$$

If further we ignore $(\rho/\rho_0)'$ in the above analysis, as being smaller than $(\mu/\mu_0)'$ and recalculate a mean value of k_1 (the separate values of k_1 turn out to be not quite so constant as in Table 2), we find that

$$\Delta S = 0{\cdot}55Q(\mu/\mu_0)' \quad . \quad . \quad . \quad . \quad (33)$$

which is exactly the result obtained by ZENER[81] who plotted log $(D_0/a^2\nu)$ against $Q(\mu/\mu_0)'/R$ and found the points for f.c.c. metals to be fairly close to a line of slope 0·55. ZENER, however, provides no clear understanding of the significance of the numerical factor. As would be expected from the foregoing, the points for b.c.c. metals lie well above the line of slope 0·55 and ZENER interpreted this as probable evidence for a ring diffusion process.

With the now known value of k_1, k_2 can be calculated from equation (29) or preferably equation (28). An examination of the displacements of its neighbours during the diffusion of an atom shows that these correspond crudely, when resolved in two directions at right angles, to atomic movements associated with shears for which C_{44} and $(C_{11} - C_{12})/2$

are the appropriate moduli. Using the sum of these as a rough estimate of the appropriate modulus in equation (28), k_2 is found[79] to be fairly constant, as shown in Table 2. Ideally, of course, the temperature coefficient of $(C_{44} + (C_{11} - C_{12})/2)$ should be used in equation (30) in determining k_1, but the necessary measurements do not appear to have been made. We might expect the value of $(\mu/\mu_0)'$ to be less sensitive to the type of modulus chosen than is μ itself, so that the calculation of k_1 may not be seriously affected by the modulus we employed not being the most appropriate. k_2, which requires a knowledge of μ and $(\mu/\mu_0)'$, may be more sensitive to choice of correct modulus and a more careful study of the problem may reveal a modulus which gives more constant values of k_2 than does the one we have used.

Equations (28) and (29) provide a useful working rule for estimating those self diffusion coefficients of metals which have not yet been directly measured.

(ii) *Chemical Diffusion*—It might be assumed that equation (24) applies equally well, just as it stands, to diffusion of the solute in a dilute solid solution, where the mutual interaction of solute atoms may be neglected. However, equation (24) is based on the assumption that successive moves of a diffusing atom are independent of one another and for chemical diffusion such independence does not necessarily hold. It does not hold for example when diffusion occurs by migration of solute atom—vacancy molecules, as JOHNSON[86] suggested would take place if vacancies were more readily formed near to solute atoms than elsewhere. However, ZENER has analysed this case and concludes that equation (24), unchanged, is still applicable. But a more extensive investigation of other possible situations would be very welcome. For example, if, when vacancies are formed equally readily on all sites so as to be distributed at random in the lattice a vacancy exchanges much more readily with a solute atom than with a solvent atom, then when a vacancy meets a solute atom it will exchange it with it a large number of times before moving away again: successive moves of a solute atom are not then independent of one another and an analysis of this situation[86] suggests that the equation should be multiplied by some factor of the order of magnitude of the co-ordination number.

However, if we calculate values of ΔS, using the equation $\Delta S = R \log (D_0/a^2\nu)$ from the extensive available data on diffusion of solute atoms in dilute solid solution, we find[81] the great majority of them are negative and often numerically very large. It seems unlikely that refinements to equation (24) would appreciably change this finding and, in fact, the one we have mentioned would produce even more negative values of ΔS. Now we have seen that negative values of ΔS are inadmissible so it appears that very many reported values of D_0 for chemical diffusion in dilute solid solutions are smaller than can be

accounted for by an extension of the theory which so successfully accounts for values of D_0 in self diffusion.

It is important to note that, very often, the lower the experimental values of D_0 for the diffusion of a series of solutes in the same solvent the lower also is the value of Q, relative to the self-diffusion activation energy of the pure solvent. If the unaccountably low values of D_0 are low because of experimental errors in the measurements, this correlation with Q values suggests that such errors *may* arise from the use of polycrystalline material in the measurements. Grain boundary diffusion processes, by contributing appreciably to and so increasing above the true lattice diffusion values the low temperature measurements of D will lead to abnormally low values of Q and at the same time of D_0 as discussed more fully on page 286, and these errors will be related to one another because both arise from the same cause. We may obtain a rough estimate of this relation as follows. Values of D are most likely to represent true lattice diffusion at high temperatures and certainly near the melting point T_M. We may, therefore, assume that the straight line representing log D for true lattice diffusion plotted against $1/T$, may be obtained from the "best straight line" drawn through experimentally obtained high and low temperature values of D by a clockwise rotation about the point where it passes through T_M. It is then readily shown that log D_0 and Q derived from the experimental straight line will differ from the true lattice diffusion values by amounts Δ log D_0 and ΔQ, which are related by

$$\Delta \log (D_0/a^2\nu) = \Delta Q/RT_M \qquad \ldots . \quad (34)$$

Relation (34) is most interesting in connexion with a correlation observed by DIENES[87, 127] between D_0 and Q. DIENES showed that if experimental values of log $(D_0/a^2\nu)$ for both self diffusion and chemical diffusion in dilute solutions are plotted against Q/RT_M the points all lie, although roughly, along a line of slope one and so of equation

$$\log (D_0/a^2\nu) = Q/RT_M - \text{constant} \qquad \ldots . \quad (35)$$

It can be shown[79] that because of a loose correlation between $(\mu/\mu_0)'$ and $1/T_M$ this equation may be derived with the same value of the constant as a very rough approximation to our previous expression for log $(D_0/a^2\nu)$: but our theory does not allow the negative values of R log $(D_0/a^2\nu) = \Delta S$, which this equation claims to cover. It is clear then why the experimental values of D_0 and Q for self-diffusion lie near the line of equation (35). The very extensive range of values of $D_0/a^2\nu$, covering about nine cycles of ten, which are included in the correlation, tends to mask the fact that many points lie away from the line by amounts which correspond to errors of a factor as great as 10 in

the values of $D_0/a^2\nu$. If then we take equation (24) as valid for chemical diffusion in dilute solid solution and assume that, say, equation (33), even approximately represents the value of ΔS for chemical diffusion, then since $\Delta S = R \log D_0/a^2\nu$ is proportional to Q and since further $(\mu/\mu_0)'$ for the alloy may not differ much from that for the pure solvent, we might expect the true values of D_0 to give points fairly close to DIENES' line of slope 1. But even if the experimental values of D_0 and Q are in error, for reasons such as mentioned above, so that these errors are related by equation (34), then by virtue of this equation they will continue to lie near the line of slope 1 when plotted as log $(D_0/a^2\nu)$ against Q/RT_M, but in a region removed from that corresponding to the true values of D_0 and Q by an amount which is larger the larger the errors in D_0 and Q.

The theory of chemical diffusion even in dilute solid solution is by no means complete and the above discussion of it in consequence very approximate, but it does appear from this work that many of the reported values of D_0 and Q are not truly representative of lattice diffusion. That D_0 and Q are roughly correlated in the way shown by DIENES does not necessarily mean that all values of D_0 and Q are at least approximately correct and that the correlation is a true consequence of lattice diffusion, for it is as readily explained if the departures of D_0 and Q from the values for true lattice diffusion are related in a way approximately represented by equation (34). One such way as we have seen is through the operation of grain boundary diffusion processes, but any effects which produce enhanced overall diffusion at low temperatures, by providing what ZENER[81] calls "short circuiting paths," with a lesser tendency to do so at high temperatures, will have a similar influence on reported values of D_0 and Q. Many of the complicating features associated with the Kirkendall effect are just such as might be more important at low temperatures. Thus for example the lattice distortion accompanying the shrinkage and expansion which occurs in chemical diffusion and which will favour increased diffusion rates, will persist more readily at low than at high temperatures. Both Kirkendall effects, through the influence of these features, and grain boundary effects in polycrystalline material occurring even in the absence of a Kirkendall effect, may lead to values of Q and D_0 which do not represent true lattice diffusion. The latter may be avoided by using single crystal material or, if polycrystalline, by working over a temperature range such that for the grain sizes used, grain boundary diffusion is negligible, but the Kirkendall effects are to be avoided only by working with very low concentration gradients, or with zero gradients as in chemical self-diffusion studies.

Evidence that these considerations may be quite important is provided by recent measurements at extremely low concentrations of

Au (0·001 per cent) in Cu[88] and of Co in Ni,[89] for these both yield values of D_0 which correspond to positive values of ΔS, even agreeing roughly with the theoretical values given by equation (33). This is very far from true for the many dilute solution values of D_0 for other elements diffusing in Cu and which have been measured with larger concentration gradients. The few results of chemical self-diffusion coefficient measurements and of measurements of diffusion coefficients by the anelastic methods discussed in Section IV, both of which employ chemically homogeneous materials, also lead to positive values of ΔS. It is also of great interest that Q for Au in Cu is quite close to that for self-diffusion in Cu in contradistinction to a general result that Q for low concentration solute diffusion in a solvent is always appreciably less than that for self-diffusion in the pure solvent (see PMP.1). This one result throws doubt on this commonly accepted generalization which was the basis of JOHNSON'S[86] solute atom-vacancy "molecule" diffusion mechanism. Further investigation may possibly show it to have been a consequence of Kirkendall or concentration gradient effects, absent in self-diffusion, introducing errors into chemical diffusion measurements even at quite small concentrations of a few per cent.

Many of these matters have been discussed at length by NOWICK,[90] who also makes use of the foregoing theory of D_0 to calculate lattice values for D_0 and Q from those reported measurements of D in various low concentration binary systems which were made at the highest temperature used in the investigations, and so are most likely to represent true lattice diffusion behaviour. Using ZENER'S approximate equation (33) we have for D

$$D = \nu a^2 \exp\left(0{\cdot}55Q\{\mu/\mu_0\}'/R\right) \exp\left(-Q/RT\right) \quad . \quad . \quad . \quad . \quad (36)$$

Hence a single high temperature value of D serves to provide a value for Q which, as NOWICK points out, even if equation (36) is correct only to an order of magnitude, will be more accurate than obtained from a conventional plot of $\log D \sim 1/T$, for such plots often yield values of D_0 which appear to be in error by factors as great as 10^8. In this way, using values of $(\mu/\mu_0)'$ for the pure solvent, he finds that the recalculated activation energies for the diffusion of a series of solutes in the same solvent are, in fact, all much closer (within $\sim$ 15 per cent) to that for self-diffusion in the pure solvent than the reported values. Also the solute activation energies appear often to be greater than the solvent self-diffusion activation energies for those solutes which according to KÖSTER and RAUSCHER[91] increase the elastic modulus of the pure solvent and to be less for those solutes which decrease the modulus, as might be expected from equation (28). From these values of Q NOWICK also calculated the corresponding values of D_0 from equation (33) but these are much more sensitive to any inaccuracies in equation (33) than

are the Q to inaccuracies in equation (36), as the putting of appropriate quantities into this equation when solved for Q will show.

Little has been said about the applicability to chemical diffusion of the theoretical expressions for ΔS we have discussed for self diffusion. While equation (30) may be formally applicable it is, particularly in the absence of a clearer knowledge of the mechanism of chemical diffusion, more difficult to know what values to insert for the various quantities it contains. Such matters, and to what extent equation (33) is still valid as an approximation to (30) remain for further study, but it is to be noted that when experimental measurements lead to positive values of ΔS, these agree, at least within an order of magnitude, with the values given by equation (33).

SECTION IV

Experimental Methods

The various conventional experimental methods by which chemical and self-diffusion coefficients in the various metals and alloys have been measured have been adequately reviewed elsewhere,[1, 50, 55, 92] together with the mathematical equipment required for their interpretation. These generally require the measurement of the rate at which an inhomogeneity in chemical composition, or in content of radioactive atoms, is changed by the process of diffusion and a compilation of the data accumulated by these methods has recently been published in Smithells' *Metals Reference Book*.[78] It is to be emphasised, however, that much of this is probably of uncertain theoretical significance for many reasons, discussed either in PMP.1 or in the present chapter.

That diffusion does take place can only mean that single atoms can migrate from one site in the lattice to another and in theoretical interpretations of diffusion coefficients the rate of this migration is regarded as the fundamental quantity. But any property of a metal which can be considered to depend upon the migration of atoms may in principle be employed to measure the rate of migration and so to provide information on diffusion.

Thus, for example, Alexander, Dawson and Kling[93] have studied the variation with temperature of the creep rate of gold wires under load and assuming this to occur by the vacancy mechanism of creep suggested by Nabarro[16] have calculated from the results a value of 48 kcal/mole for the activation energy for gold self-diffusion. Using a more refined theory due to Herring[94] they obtain a value of 50 kcal/mole, agreeing better with the directly measured value of 53.[83]

But two properties in particular have been made use of recently to determine diffusion coefficients and these provide several advantages

over the conventional methods of measurement. These are, certain types of internal friction in metals and their ability to sinter.

1. *Sintering Method for Measuring Self-diffusion Coefficients*

KUCZYNSKI[95–98] and others have studied the rate of sintering of spherical particles to a flat plate of the same material or of a fine wire to the surface of a cylinder on which it is wound, when the two become bonded together in a way shown schematically in section in Fig. 11. Various processes may be considered as responsible for the obvious transfer of material that takes place. If this is to be by volume diffusion then KUCZYNSKI[95] supposes that, just as the equilibrium vapour pressure over a convex surface of a liquid is greater than that over a plane surface so, assuming a vacancy mechanism of diffusion, is the equilibrium vacancy concentration over the highly convex surface just inside the neck or bond in Fig. 11 greater than that prevailing elsewhere. This difference in concentration is such as to cause vacancies to flow away from the neck and, therefore, atoms into it, enabling the bond to grow in width. With such a model for the growth of the bond between particle or wire and plate KUCZYNSKI obtains expressions from which the self-diffusion coefficient can be calculated. For a particle bonding to a plate

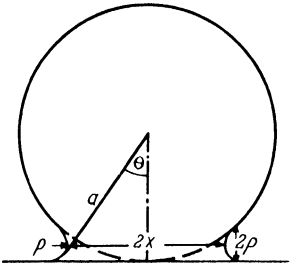


Fig. 11. Section of Wire or Spherical Particle sintered to flat surface (Schematic)[95]

$$D = kTx^5/40\gamma\delta^3ta^2 \quad . \quad . \quad . \quad . \quad (37)$$

and for a wire bonding to a plate or cylinder of large radius

$$D = 3kTx^5/80\pi\gamma\delta^3ta^2 \quad . \quad . \quad . \quad . \quad (38)$$

where k is Boltzmann's constant, T the absolute temperature of anneal, γ the surface energy of the metal, δ the interatomic distance, and a the radius of the particle or wire. The experiments require a measurement of the width $2x$ of the bond formed after a time of anneal t. This is done by microscopic measurements on samples suitably mounted and sectioned. The necessary dependence of x^5 on t demanded by equations (37) and (38) has been confirmed, at least for not too small sized particles ($> 15\mu$), by experiments on silver[95, 97] and on copper[95, 96, 98] and the measurements used to calculate the lattice self-diffusion coefficients for these two metals.

Since x^5 is proportional to t it is important to obtain an accurate value for $2x$. But the present methods of experiment seem to give a considerable spread of values for $2x$ from particle to particle or in different parts of the wire when this technique is used, and a mean value

must be taken for a large number of observations. This must detract from the accuracy of the measurements, and for this reason, because of the approximate nature of the derivation of equations (37) and (38) and because of uncertainties in the value of γ, it is surprising that the results for silver should agree so closely with those measured by conventional radioactive methods. Thus KUCZYNSKI finds for silver,[97] using the wire and cylinder method, which is reported as being the most convenient and accurate, $D = 0{\cdot}9 \exp(-45{,}700/RT)$ compared with JOHNSON'S results,[31] recently well confirmed by HOFFMAN and TURNBULL,[29] of $D = 0{\cdot}89 \exp(45{,}900/RT)$.

The results for copper are in rather less good agreement but still fairly close to those obtained by conventional radioactive methods.

KUCZYNSKI[95] considered that the transfer of matter was definitely by volume diffusion for his examination of the other possible processes of viscous or plastic flow, of evaporation and condensation and of surface diffusion suggested that these require the relation between x and t to be, respectively, $x^2 \sim t$, $x^3 \sim t$, and $x^7 \sim t$, whereas the experiments clearly produced a fifth power law consistent only with the supposition of volume diffusion. However, both CABRERA[99] and SCHWED[100] have criticized KUCZYNSKI'S treatment of surface diffusion and pointed out than an $x^5 \sim t$ law should hold for both surface and volume diffusion: SCHWED further suggests that for very small particles the law should become $x^3 \sim t$. The experiments on silver and copper seem though to support the view that sintering is in fact by volume diffusion and not surface diffusion, at least for the particle and wire sizes used, particularly because the activation energies are very close for silver and for copper to those found for volume diffusion by more direct methods.

The method is one of promise and seems capable of providing quite accurate values for D, although it has not yet been fully developed experimentally: another way of measuring the width of the bond which has been suggested[101] is to follow the change in contact resistance between two metallic spheres or hemispheres as these sinter together. It would benefit from a more thorough theoretical treatment of the sintering process. It has the great advantage over conventional methods in that it could be used to measure the self-diffusion coefficients of metals for which no suitable radioactive tracers exist, such as aluminium and magnesium.

2. *Anelastic Methods*

In both interstitial solid solutions in b.c.c. metals[102] and substitutional solid solutions in f.c.c. metals[103] there is a type of anelasticity which has been interpreted as arising from a redistribution of solute atoms in the presence of an applied stress, and which is, therefore, intimately related to the diffusion properties of the solute atoms. Measurements on

this type of anelasticity have recently been employed to measure the diffusion coefficients of solute atoms in both these types of solid solution.

A solute atom in an octahedral interstitial position will produce a distortion in the surrounding lattice with tetragonal symmetry, the tetragonal axis being along one of the three ⟨100⟩ directions.[102] In the absence of any applied stress the tetragonal axes of distortion will be distributed equally among these three ⟨100⟩ directions, but if a tensile stress is applied along one of these directions atoms will tend to move into those positions for which the tetragonal axis is along the same direction, for the stress is then to some extent relieved. A certain time will be required for the establishment of the new stress induced preferential distribution of atoms so that the strain will lag behind the stress. This hysteresis will give rise to relaxation and other anelastic effects such as internal friction.

In substitutional solid solution isolated solute atoms produce a local distortion with cubic symmetry so that no redistribution occurs on applying a stress. But in a random distribution of solute atoms there will occur pairs of solute atoms which are nearest neighbours to one another and the local distortion produced around such pairs will be of tetragonal symmetry. On applying a stress the pair axes will tend to orient themselves along those crystallographically permissible directions which make the least angle with the direction of stress, for by so doing the stress is to some extent relieved and the strain energy reduced. A pair reorientation is achieved by the migration of either member of the pair to a mutual nearest neighbour position, and it is to such a process that these particular internal friction effects, observed only in alloys, have been attributed.

Clearly, one would expect the relaxation strength, or the magnitude of the internal friction, to be greater, for equal solute concentration, for interstitial solid solution in b.c.c. lattices than in f.c.c. substitutional solutions, for in the latter case it is only the randomly occurring pairs of atoms which contribute to the anelastic effects. Thus, these may be observed with very low concentration of for example C in α-Fe (·001 per cent) whereas a concentration greater than ~ 10 per cent is necessary with substitutional alloys. The larger the difference in size between solute and solvent atoms the greater the relaxation strength and so the lower the concentration at which observable effects are obtained. But, at these high concentrations the concept of the reorientation of pairs requires considerable modification, for much larger groupings of solute atoms will occur and isolated pairs will be relatively uncommon. However, calculations based on the simple pair concept do seem to provide from experimental measurements, values of D which, when extrapolated to the higher temperatures at which conventional diffusion coefficient measurements are made, agree fairly well with the

values the latter give. This, and other evidence, confirms that the internal friction effects are intimately associated with the diffusion properties of the alloy, but a more careful theoretical treatment would be welcome. The anelastic effects in the b.c.c. interstitial solutions are more clear in origin and free from such difficulties of theoretical interpretation.

These anelastic effects[104] are characterized by the "relaxation time" τ_R required for the establishment of the new stress induced equilibrium redistribution of atoms or pairs of atoms and τ_R is clearly closely related to τ, the mean time of stay of any one atom in one lattice position. The diffusion coefficient D for the solute atoms is given by equation (13), so that a determination of τ_R serves to determine D.

For interstitial solid solution[126, 57] in b.c.c. lattices $\tau = 3\tau_R/2$ and for substitutional solid solution in f.c.c. lattices it is sufficiently accurate to take $\tau = \tau_R$ for polycrystalline material, although for single crystals the relation between τ and τ_R varies with the direction of the stress.[105]

τ_R may be determined by measurements either of the internal friction or of the elastic after-effect.[104]

The first method consists in measuring the temperature at which the maximum occurs in a plot of internal friction as a function of temperature. If $\bar{\omega}$ is the vibrational frequency at which the experiments are being conducted then the τ_R at this temperature $= 1/\bar{\omega}$. Measurements of τ_R at other temperatures are made by varying the frequency $\bar{\omega}$. With the specimen as a wire forming the suspension of a torsional pendulum and with photographic recording of the vibrations, measurements may be made in the range $\sim$ 1 to 100 c.p.s. For higher frequencies and, therefore, also temperatures, it is more convenient to study the damping of a bar of material set into flexural oscillations, using electronic methods of timing.

In the second method, applicable at lower temperatures and for values of τ_R greater than a minute or so, the specimen in wire form is twisted and kept so for a time of the order of τ_R. On releasing the stress the elastic strain is recovered immediately, but the anelastic strain only gradually and a measure of this strain as a function of time determines τ_R.

Using both methods values of τ_R from about 1 hr to, say, 10^{-4} sec (for $\bar{\omega} \sim 10K$ c.p.s.) may be measured covering 8 cycles of 10 in the value of D. The corresponding temperatures are such as to extend the temperature range over which diffusion coefficients may be measured from the comparatively restricted range covered by the conventional methods right down to below room temperature for C in α-Fe, for example,[57] and to within about 200°C of room temperature for alloys such as CuZn[106] and AgZn.[107]

This very considerable advantage of the method means that very

accurate values of D_0 and Q may be obtained, which, as mentioned on p. 313 give positive values for the entropy of activation. Kirkendall or concentration gradient effects are without influence on the results for measurements are made, of course, on chemically homogeneous material, and, therefore, also provide values of D at well-defined compositions. Grain boundary diffusion plays no part in determining τ_R but, at least for the substitutional alloys, single crystal or quite coarsely polycrystalline material must be used to avoid or reduce anelastic effects

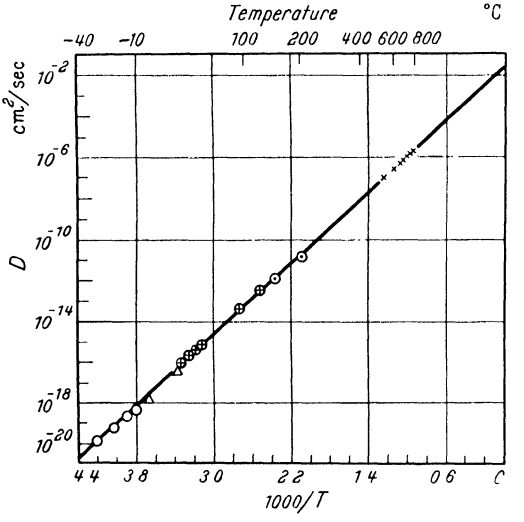


Fig. 12. Diffusion Coefficient of —

× = Conventional Diffusion Measurements—STANLEY[109]

○ = Elastic After-effect Measurements—WERT[57]

△ = Elastic After-effect Measurements—RICHTER[124]

⊕ = Internal Friction Measurements—WERT[57]

○ = "Precipitation Measurements"—WERT[57] (Not described in text—See ref.)

associated with grain boundaries such as grain boundary slip,[108] which might otherwise contribute more appreciably to the observed anelastic effects than the process being studied.

For substitutional solid solutions the method has certain limitations. The solute and solvent atoms must differ appreciably in size and the concentration must not be so low that the number of solute atom pairs is inappreciable.

These methods have been employed most extensively to interstitial solid solution, in particular of C[126, 102, 56, 57] and N[102, 56, 58] in α-Fe and of O_2,[110] N_2,[111] and C[110] in Ta and have provided[56–58] the accurate values of D_0 and Q made use of in the work discussed in Section III. Fig. 12

shows the results for C in α-Fe largely by WERT,[57] together with those of STANLEY[109] measured by the conventional methods. The extension of the measurements to lower temperatures clearly provides for an accuracy in D_0 and Q quite unattainable by the conventional methods alone.

Comparatively little use has yet been made of the method with substitutional solid solutions. NOWICK[107] has studied the diffusion of Zn (15 per cent) in Ag and LE CLAIRE and CHILDS[106] of Zn (10–30 per cent) in Cu and in each case the results give a very accurately linear relation between log D and $1/T$, which when extrapolated to high temperatures give values of D agreeing fairly well with those obtained by conventional methods. Earlier measurements[108, 122] of the internal friction in 30 per cent Zn α-brass have since been made use[105] of to give values for D. An interesting result of the CuZn experiments is that between 15 and 30 atomic per cent Zn, the activation energy is fairly constant at 38,000 kcal/mole, contrary to previous measurements which show a marked variation. D increases with composition but not so rapidly as indicated by RHINES and MEHL.[123]

An interesting use of this method stemming from the advantages of the short times in which and the low temperatures at which diffusion coefficients may be measured by it, has already been described (p. 303).

These advantages arise, of course, because the internal friction methods of measurement only require the diffusion of atoms over distances of the order of the atomic spacing, whereas the conventional methods require diffusion distances of ~ 1/mm or so. In order to obtain the latter in a reasonable time, experiments have to be conducted at high temperatures where the jump frequency is large.

The internal friction methods are therefore of importance in any study of the effect on diffusion rates of any condition of the material which is appreciably influenced by the temperatures at which conventional measurements are made, in a time short compared with that required for their execution. Such conditions are the presence of cold work and of radiation damage as well as of a non-equilibrium vacancy concentration.

SECTION V

THERMODYNAMIC AND KINETIC THEORY OF CHEMICAL DIFFUSION

Darken's Equations

DARKEN,[8] it will be remembered, in order to provide a phenomenological description of binary diffusion which would embrace the Kirkendall effect, proposed that a separate and independent diffusion coefficient should be ascribed to each of the two diffusing species defining their

respective rates of flow relative to a system of markers fixed in the system. It is usual to assume that the markers define lattice planes so if we generalize DARKEN'S proposal, the rates of flow J_i relative to the lattice of the several species of an n component system are supposed given by

$$J_i = -D_i \partial c_i / \partial x \quad . \quad . \quad . \quad . \quad (39)$$

In the case of a binary alloy these equations lead, subject to the previously discussed assumptions, to equations (1) and (2) of Section I.

Now equation (39) implies that the necessary condition for diffusion to occur between two points is that there exists a concentration gradient between them and assumes the simplest mathematical expression of this fact by putting the rate of flow proportional to the concentration gradient. The constants of proportionality D_i then characterize the diffusion behaviour of the system. Although this is a reasonable deduction from experience, it is thermodynamically more sound to expect that the essential condition for diffusion of any one species to occur between two points is that there should exist between them a difference in the chemical potential of that species, for a system is only in equilibrium when the chemical potential of each species present is the same throughout the system. The simplest mathematical expression of this fact is to assume the rate of flow of any one species to be proportional to the gradient of its chemical potential μ_i. The rates of flow of the several species are then to be given by the equation

$$J_i = -L_i \partial \mu_i / \partial x \quad . \quad . \quad . \quad . \quad (40)$$

The diffusion behaviour of each species may now be specified by the corresponding coefficient L_i but it is more common practice to retain equation (39) as a definition of the diffusion coefficients and to express each D_i in terms of the quantities appearing in equation (40). In this way we find that (see PMP.1)

$$D_i = \frac{L_i}{c_i} \frac{\partial \mu_i}{\partial \log c_i} = RT \frac{L_i}{c_i} \left(1 + \frac{\partial \log \gamma_i}{\partial \log c_i}\right) \quad . \quad . \quad . \quad . \quad (41)$$

where use has been made of the usual expression, $\mu_i = \text{const} + RT \log c_i \gamma_i$ for μ_i in terms of the activity coefficient γ_i. The term L_i/c_i corresponds to the "mobility" G_i in terms of which the phenomenological theory of diffusion was discussed in PMP.1.

These phenomenological equations (40) and (41) are frequently used in preference to equation (39) in dealing with diffusion, particularly in liquids, and although their use in describing metallic diffusion processes has been suggested from time to time in the literature, little application appears to have been made of them until the work of BIRCHENALL and MEHL[112] and of DARKEN.[8]

The first important application of these equations, by DARKEN, was to provide a relation between the individual chemical diffusion coefficients D_i and the corresponding self-diffusion coefficients D_i^*. For this purpose DARKEN assumes that the mobility L_i/c_i, defining the average velocity of flow by diffusion under unit free energy gradient (see PMP.1), is the same, for any one chemical species, under condition of both self-diffusion and chemical diffusion, so that differences between D_i and D_i^* arise entirely from differences in the chemical potential gradient in the two cases. In self diffusion the chemical potential gradient involves only the entropy of mixing so that

$$D_i^* = RTL_i/c_i \quad \ldots . \quad (42)$$

With the above expression for D_i, we then find

$$D_i = D_i^*(1 + \partial \log \gamma_i/\partial \log c_i) \quad \ldots . \quad (43)$$

relating D_i and D_i^*.

For the special case of a binary diffusing system we may now relate D_a^* and D_b^* with the chemical interdiffusion coefficient $\tilde{D}$ and the rate of marker movement v (equations (1) and (2)) by making use of the Gibbs-Duhem relation for a binary system. $\partial \log \gamma_a/\partial \log c_a = \partial \log \gamma_b/\partial \log c_b$.

Thus

$$\tilde{D} = (N_a D_b^* + N_b D_a^*)(1 + \partial \log \gamma_a/\partial \log c_a) \quad \ldots . \quad (44)$$

and

$$v = (D_a^* - D_b^*)(1 + \partial \log \gamma_a/\partial \log c_a)\partial c_a/\partial x \quad \ldots . \quad (45)$$

In support of equation (44) DARKEN quoted the earlier results of JOHNSON,[13] who measured the self-diffusion coefficients D_{Au}^* and D_{Ag}^* of gold atoms and silver atoms respectively in a chemically homogeneous alloy of 50 per cent composition, and also the chemical interdiffusion coefficient $\tilde{D}$ at the same composition. For this latter purpose a diffusion couple was used the compositions of the two halves of which differed only slightly from their mean of about 50 per cent. DARKEN showed that JOHNSON'S values of D_a^*, D_b^* and $\tilde{D}$ were satisfactorily related by equation (44) with $N_a = N_b = \frac{1}{2}$ and with the already known thermodynamic data on Au/Ag alloys being used to calculate the last factor. Equations (39), (40), (41) and (43) as well as (1) and (2) are generally referred to in the metallurgical literature as Darken's equations, although most of them appear to have been proposed independently by workers in other fields.[114]

More recently LE CLAIRE and BARNES[19] have pointed out that the marker shift actually observed in Au/Ag alloys by CORREA DA SILVA and MEHL[3] agrees quite well with that calculated from equation (45), using JOHNSON'S measurements of D_a^* and D_b^*. This calculation required a rough estimate of the value of $\partial c_a/\partial x$ in CORREA DA SILVA

and MEHL'S experiments which was not reported by them. SEITH *et al*[115, 128] actually measured the marker shift in a Au/Ag couple identical with the one used by JOHNSON and used his reported data to give the exact value of $\partial c_a/\partial x$. Their observed shift of $2{\cdot}90 \cdot 10^{-3}$ cm agrees fairly well with the value of $3{\cdot}10 \cdot 10^{-3}$ cm calculated from equation (45).

These measurements on Au/Ag alloys do then appear to provide a fairly satisfactory check of DARKEN'S equations but more extensive experimental work of a similar nature would be very welcome and, indeed, necessary to establish the equations firmly. A case might be made that any discrepancy between experiment and calculation based on Darken's equations was to be attributed to uncertainties in the measured value of the chemical diffusion coefficients arising from porosity and lateral changes in the dimensions of diffusion samples as discussed in Section I. Thus SEITH and KOTTMANN[128] point out that the observed value of v in their Au/Ag samples is *less* than the calculated value, as is to be expected if the discrepancy arises from these causes.

The relation between D_i^* and D_i is important for many reasons. It provides the essential link between the theoretically most simple and fundamental quantity underlying all diffusion phenomena, namely the average number of random jumps made per second by an atom in a homogeneous alloy, derived from D_i^*, and the practically significant rates of chemical diffusion given by D_i. The form of the relations will, in general, vary with the mechanism of diffusion and so a choice in principle may be made between proposed mechanisms by experimental study of the corresponding relations (see PMP.1).

A number of recent papers[15, 59, 20, 120, 121] have been concerned with the validity of DARKEN'S equation and of the assumptions upon which they are based. One important feature omitted from DARKEN'S treatment of diffusion is the part played by the vacancies, when diffusion occurs by this process. There is a definite nett flux of vacancies as the Kirkendall experiments show and it is natural to inquire to what extent this vacancy flow affects the validity of equations (39) and (40). It is by no means apparent why, as equation (40) states, the rate of flow of atoms of one species should be proportional only to the chemical potential gradient of atoms of that species when there is a nett flow of the vacancies whose movement alone can lead to a flow of the atoms, or to put it another way, when there is such an intimate coupling between the flow of atoms and of vacancies.

A study of the problem is conveniently made in terms of ONSAGER'S general theory of irreversible phenomena.[116–18]

General Thermodynamic Theory of Diffusion

The essential assumption is that the rate of flow J_i of each of the several n species present in a multicomponent system is linearly

dependent upon the chemical potential gradient of *each* of the species present, *including the vacancies*, so that we have as the "thermodynamic equations of motion"

$$J_i = \sum_{j=1}^{n} L_{ij} X_j \qquad \ldots . \quad (46)$$

where for short we write $-X_j$ as the chemical potential gradient of the jth species, $X_j = -\partial\mu_i/\partial x$.

The terms L_{ii} correspond closely to the terms L_i in equation (40) and, therefore, to the diffusion coefficients but the "cross-terms" L_{ij} $(i \neq j)$ take account of possible couplings between the several diffusion flows. It may be shown[116] that as a consequence of the principle of microscopic reversibility, a symmetry exists between these cross terms in equation (46) expressed by the "Onsager Relations."

$$L_{ij} = L_{ji} \ (i \neq j) \qquad \ldots . \quad (47)$$

(the same is not in general true when equation (46) is rewritten with $\partial c_i/\partial x$ in place of $\partial\mu_i/\partial x$).

The L_{ij} are functions of composition and will in general vary according to the choice of plane with reference to which the J_i are measured. Let this plane be such that the nett flow of all species across it is zero. Then

$$\sum J_i = 0 \qquad \ldots . \quad (48)$$

Since for each atom moving from one lattice plane to the next there is a vacancy moving in the other direction, equation (48) is satisfied for a plane fixed relative to the lattice, as required for comparison of the J_i given by equation (48) with those given by equation (39) or (40).

In addition to the "Onsager relations," equation (48) leads to a further reduction in the number of independent coefficients L_{ij} for it may very simply be shown[117] to lead to the result

$$\sum_{j=1}^{n} L_{ij} = 0 \qquad \ldots . \quad (49)$$

$$\sum_{i=1}^{n} L_{ij} = 0 \qquad \ldots . \quad (50)$$

Let us now regard the vacancies as the nth component and put $n = v$. With the aid of equation (49) we may now rewrite the equation (46) for J_i as

$$J_i = \sum_{j=1}^{n-1} L_{ij} X_j + L_{iv} X_v$$

$$= \sum_{j=1}^{n-1} L_{ij}(X_j - X_v) \qquad \ldots . \quad (51)$$

where $-X_j$ and $-X_v$ are the chemical potential gradients of the jth component and of the vacancies respectively.

The equations (51) give in the most convenient general form the rates of flow by diffusion of the vacancies and of the $n-1$ atomic species.

If the density of vacancies at each point in the system is such as corresponds to them being in equilibrium with the lattice, then the free energy of the vacancies is a minimum for small changes in their number, so that μ_v and also $X_v = 0$. Equation (51) now becomes

$$J_i = \sum_{j=1}^{n-1} L_{ij} X_j \quad \ldots . \quad (52)$$

If further we neglect all cross terms L_{ij} we arrive at

$$J_i = L_{ii} X_i \quad \ldots . \quad (53)$$

Neglecting cross terms but not assuming the vacancies are in equilibrium we find

$$J_i = L_{ii}(X_i - X_v) \quad \ldots . \quad (54)$$

Equation (53) proposes that each J_i is determined solely through the corresponding X_i. Such phenomenological equations are identical with equation (40) on page 321. We now see that their validity rests on the assumption that the vacancies are everywhere maintained in thermodynamic equilibrium with the lattice and that the cross terms, which give a contribution to the flow of each component proportional to the chemical potential gradient of each of the other components, are negligible.

It may well be in many cases that the vacancies are everywhere in thermal equilibrium so that at least equation (52) is valid. This will certainly be true in self-diffusion experiments, where chemical homogeneity is maintained and also in chemical diffusion experiments if there are sources or sinks for vacancies, such as dislocations, adequate to maintain equilibrium.

As regards the cross terms, at least those representing coupling between the flow of vacancies and the flows of the several atomic species are not to be neglected, as is to be expected from the very nature of vacancy diffusion. The nett flow of vacancies is certainly not zero as the Kirkendall effect shows and being given by

$$J_v = \sum_{j=1}^{n-1} L_{vj} X_j$$

may be regarded as determined entirely by the cross terms L_{vj} ($j \leq n-1$). The fuller and more general phenomenological theory expressed by equation (46) is then necessary because of the observed nett vacancy flow.

But no convincing demonstration has yet been given of the part played in vacancy diffusion by the cross terms between different

atomic species, and in particular of whether they are zero or sufficiently small to be neglected. Obviously, for an exchange or ring diffusion mechanism, the cross terms would certainly be significant.

It is to be noted that a simple diffusion coefficient D_i as given by equation (39) can only in general be defined for each species when cross terms are neglected and equation (53) is valid. With non-zero cross terms then it is not possible in an n component system to obtain any equation corresponding to equation (43).

However, in the simple but important special case of binary diffusion the rate of flow of each species may, when the vacancies are in equilibrium, be written as proportional only to its own concentration gradient even when cross terms are not ignored. We may, therefore, obtain diffusion coefficients for the two species, so that the use of DARKEN'S equations (1) and (2) is consistent with this more general theory. But DARKEN'S relation (43) and, therefore, (44) and (45) only follows when cross terms are zero. These results are simply shown as follows.

In a measurement of the self-diffusion coefficient of one component of a binary alloy we are dealing with a system of four identifiable species, inactive A atoms, active A^* atoms, inactive B atoms, and vacancies. Denoting these four species respectively by suffixes 1, 2, 3, 4 the "thermodynamic equations of motion" are[121]

$$\left.\begin{aligned} J_1 &= L_{11}X_1 + L_{12}X_2 + L_{13}X_3 + L_{14}X_4 \\ J_2 &= L_{21}X_1 + L_{22}X_2 + L_{23}X_3 + L_{24}X_4 \\ J_3 &= L_{31}X_1 + L_{32}X_2 + L_{33}X_3 + L_{34}X_4 \\ J_4 &= L_{41}X_1 + L_{42}X_2 + L_{43}X_3 + L_{44}X_4 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (55)$$

A self-diffusion coefficient measurement requires a determination of J_2 in a chemically homogeneous alloy of A and B, i.e. in which $X_3 = 0$ and $c_1 + c_2 = c_a$, the uniform and constant total concentration of A atoms. If we assume the vacancies are equilibrium $X_4 = 0$ and

$$J_2 = L_{21}X_1 + L_{22}X_2 \quad . \quad . \quad . \quad . \quad (56)$$

The Gibbs-Duhem relation in the form $\sum c_i X_i = 0$ enables us to eliminate X_1 from (56) and so obtain J_2 in the required form, from which we find for the self-diffusion coefficient of A atoms in the alloy D_a^*,

$$D_a^* = -RT\left(L_{22} - \frac{c_2}{c_1}L_{12}\right)\Big/ c_2 \quad . \quad . \quad . \quad . \quad (57)$$

The further condition that $J_1 + J_2 = 0$ and $J_4 = 0$ lead to the relations

$$c_2 L_{31} - c_1 L_{32} = 0 \quad . \quad . \quad . \quad . \quad (58)$$

and

$$c_2(L_{11} + L_{21}) - c_1(L_{22} + L_{12}) = 0 \quad . \quad . \quad . \quad . \quad (59)$$

Use of (59) enables us to write D_a^* in the more symmetric form[121]

$$D_a^* = -RT\left(L_{11} + L_{22} - \left\{\frac{c_1}{c_2} + \frac{c_2}{c_1}\right\} L_{12}\right) \Big/ c_a \quad . \quad . \quad . \quad . \quad (60)$$

A measurement of the chemical diffusion coefficient of A requires a determination of the total flow $J_1 + J_2$ of A in an alloy in which a chemical concentration gradient of A exists. In this case, not the sum of c_1 and c_2 but the ratio c_1/c_2 remains constant at every point. Thus, as c_1 and c_2 are concentrations of chemically identical species

$$c_1X_1 + c_2X_2 = c_aX_a \quad . \quad . \quad . \quad . \quad (61)$$

where $-X_a$ is the total chemical potential gradient of A atoms, and

$$X_1 = X_2 = X_a$$

With these and the Gibbs-Duhem relation we find for D_a

$$D_a = -RT\{(L_{11} + L_{22} + 2L_{12})/c_a - (L_{13} + L_{23})/c_3\}\left\{1 + \frac{\partial \log \gamma_a}{\partial \log c_a}\right\} \quad . \quad . \quad . \quad . \quad (62)$$

From (60) and (62) the general relation between D_a and D_a^* becomes,[90] using (58),

$$D_a = \left\{D_a^* + \frac{RTc_a}{c_1c_2c_3}(c_1L_{23} - c_3L_{12})\right\}\left\{1 + \frac{\partial \log \gamma_a}{\partial \log c_a}\right\} \quad . \quad . \quad . \quad . \quad (63)$$

It is clear that this is equivalent to DARKEN'S relation between D_a and D_a^* only when $L_{23}/c_3 = L_{12}/c_1$ or when the cross terms are both zero. The first condition would mean that the diffusion behaviour of the two species of atoms are identical whereas the Kirkendall effect clearly shows that such is not the case. (The analogous equation (58) is a direct consequence of the identity of A and A^* atoms.) This condition would, of course, hold for a mixture of three identical isotopes; since the thermodynamic factor is unity in this case we arrive at the trivial result $D_a = D_a^*$.

We have seen that there is a little experimental evidence in support of equation (43), which therefore suggests that the cross terms may be unimportant. But a theoretical justification of this conclusion, indicating that cross terms are, in fact, zero or at least sufficiently small to be neglected, can only come through a detailed study by kinetic methods of the process of vacancy diffusion in binary alloys. Although several attempts have been made to construct such a theory, the problem is not yet sufficiently clearly or convincingly resolved to warrant a full review of them.

It is interesting to note in passing that the cross terms do play an important part in the phenomena of diffusion in strong electrolytes.[113]

Kinetic Theories of Diffusion

BARDEEN[120] has shown how equation (54) may be derived by simple kinetic considerations, so that DARKEN'S equation (43) again follows if the vacancies are in equilibrium, but the assumptions introduced into the analysis have been criticized by SEITZ,[15] who regards them as "exceedingly restrictive." BARDEEN and HERRING[20] in a later paper claim to have avoided the more important of these assumptions, without influence on the results, but their treatment is not without objections. SHUTTLEWORTH[121] has considered a model in which one species of atom exchanges very much more readily with the vacancies than does the other, and finds a relation between D_a and D_a* which differs from equation (43) in that the factor $(1 + \partial \log \gamma_a/\partial \log c_a)$ is replaced by $(2 + \partial \log \gamma_a/\partial \log c_a)$ so that cross terms are clearly of importance in such a case.

It would hardly be profitable in the present incomplete state of the subject, to describe all these various attempts at the problem in detail, but SEITZ'S[15, 59] more general approach may be mentioned for it reveals in a fairly simple way and with the fewest assumptions, the requirement that the vacancies must be in equilibrium for the rates of flow of the two species to be expressible by diffusion coefficients.

SEITZ was the first clearly to point out that the Kirkendall effect was consistent with a vacancy mechanism of diffusion if it was assumed that vacancies exchanged more readily with one type of atom than with the other. He introduces the quantity $\pi_a(\xi\eta)$ defining the probability that a vacancy on a plane where the composition is ξ will exchange in unit time with an A atom on a neighbouring plane where the composition is η. A similar quantity $\pi_b(\xi\eta)$ defines the corresponding probability for a vacancy exchange with a B atom. The total probability that in unit time a vacancy will jump from plane 1 to plane 2 is clearly $\pi_a + \pi_b$, which sum we denote by π.

We suppose diffusion to occur between two neighbouring planes 1 and 2 separated by a distance λ and whereon respectively the planar densities of A atoms, B atoms, and vacancies are $n_{a1}n_{a2}$, $n_{b1}n_{b2}$, $n_{v1}n_{v2}$. If the total number of sites n remains constant during the diffusion and since n_v is very small the composition on each plane is completely determined when only n_{a1} and n_{a2} are given.

The nett flow of vacancies from plane 1 to plane 2 is

$$J_v = n_{v1}\pi(n_{a1}n_{a2}) - n_{v2}\pi(n_{a2}n_{a1})$$

Since n_{a1} and n_{a2} differ only slightly from their mean value of n_a, π may be expanded to the first order in λ to give

$$J_v = -\pi(n_a n_a)\frac{\mathrm{d}n_v}{\mathrm{d}x}\lambda + n_v\left(-\frac{\partial\pi}{\partial\xi} + \frac{\partial\pi}{\partial\eta}\right)\lambda\frac{\mathrm{d}n_a}{\mathrm{d}x}$$

We thus see that even when there is no concentration gradient of vacancies a flux of vacancies exists solely on account of the chemical concentration gradient.

The flow of A atoms is clearly

$$J_a = n_{v2}\pi_a(n_{a2}n_{a1}) - n_{v1}\pi_a(n_{a1}n_{a2})$$

which we may write as

$$J_a = \pi_a(n_a n_a)\frac{\mathrm{d}n_v}{\mathrm{d}x}\lambda + n_v\left(\frac{\partial\pi_a}{\partial\xi} - \frac{\partial\pi_a}{\partial\eta}\right)\lambda\frac{\mathrm{d}n_a}{\mathrm{d}x} \quad . \quad . \quad . \quad . \quad (64)$$

If the vacancies are everywhere in thermodynamic equilibrium with the lattice, then the vacancy concentration is a function only of chemical composition and we may write

$$\frac{\mathrm{d}n_v}{\mathrm{d}x} = \frac{\mathrm{d}n_v}{\mathrm{d}n_a}\cdot\frac{\mathrm{d}n_a}{\mathrm{d}x}$$

J_a in equation (64) may now be written proportional to the concentration gradient of A atoms, so we may obtain the chemical diffusion coefficient D_a as

$$D_a = -\pi_a(n_a n_a)\frac{\mathrm{d}n_v}{\mathrm{d}n_a}\lambda^2 - n_v\left(\frac{\partial\pi_a}{\partial\xi} - \frac{\partial\pi_a}{\partial\eta}\right)\lambda^2 \quad . \quad . \quad . \quad . \quad (65)$$

where n_a/λ has been replaced by c_a, the volume concentration of A atoms.

We see then again that equation (39) and, therefore, also (1) and (2) are valid for binary diffusion only when the vacancies are maintained in equilibrium.

It is readily shown that $D_a{}^*$ is given by

$$D_a{}^* = n_v\pi_a(n_a n_a)\,.\,\lambda^2/n_a \quad . \quad . \quad . \quad . \quad (66)$$

so that the expression for D_a and $D_a{}^*$ are intrinsically different in form for $D_a{}^*$ involves only π_a whereas D_a involves also its first derivatives.

This represents the least objectionable justification by kinetic methods available so far for the use of equations (1) and (2) in describing binary diffusion; it is only when simplifying assumptions are made as to the form of π that a relation between π and its derivatives and, therefore, between D_a and $D_a{}^*$ may be obtained. It is, in effect, at this stage, in the choice of simplifying assumptions, that the various attempts at constructing a kinetic theory of diffusion, diverge. And none of these assumptions is sufficiently unrestricted to lead to a convincingly general result.

References

[1] Le Claire, A. D., *Progress in Metal Physics Vol.* 1 (Ed., B. Chalmers) Chap. 7. Pergamon Press London 1949

[2] Smigelskas, A. D. and Kirkendall, E. O. *Trans. Amer. Inst. Min. Met. Engrs.* 171 (1947) 130–135

[3] Correa da Silva, L. C. and Mehl, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* 191 (1951) 155–73

[4] Bückle, H. and Blin, *J. Inst. Metals.* 80 (1951–2) 385–9

[5] Barnes, R. S. *Proc. Phys. Soc.* B 65 (1952) 512–25

[6] Seith, W. and Kottmann, A. *Naturwiss.* 39 (2) (1952) 40–41

[7] Smith, C. S. *Trans. Amer. Inst. Min. Met. Engrs.* (Discussion) 171 (1947) 135–42

[8] Darken, L. S. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 184–94

[9] Matano, C. *Japan J. Phys.* 8 (1933) 109–13

[10] Birchenall, C. E., Correa da Silva, L. C. and Mehl, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* (Discussion) 175 (1948) 197–200

[11] Smigelskas, A. D. and Kirkendall, E. O. *Trans. Amer. Inst. Min. Met. Engrs.* (Discussion) 171 (1947) 135–142

[12] Baluffi, R. W. and Alexander, B. H. *U.S. Atom. Energy Comm. Rept.* SEP–74

[13] Johnson, W. A. *Trans. Amer. Inst. Min. Met. Engrs.* 147 (1942) 331–46

[14] Fraenkel, W. and Houben, H. *Z. anorg. allg. Chem.* 116 (1921) 1

[15] Seitz, F. *Acta Cryst.* 3 (5) (1950) 355–63

[16] Nabarro, F. R. N. *Conference on the Strength of Metals* (Bristol) p. 75–90. The Physical Society, London 1948

[17] Frank, F. C. *Crystal Growth: Discuss. Faraday Soc.* No. 5, 67 (1949)

[18] Seitz, F. *Phys. Rev.* 79 (6) (1950) 1002

[19] Le Claire, A. D. and Barnes, R. S. *J. Metals* 3 (1951) 1060–2

[20] Bardeen, J. and Herring, C. *Atom Movements* American Society for Metals, 1951

[21] Rhines, F. N. and Wells, C. *Trans. Amer. Soc. Met.* 27 (1939) 625–57

[22] Seitz, F. *Advances in Physics* Quart. Suppl. to *Phil. Mag.* 1 (1952) 1

[23] Le Claire, A. D. *Trans. Amer. Inst. Min. Met. Engrs.* 175 (1948) 194–7

[24] Seith, W. and Keil, A. *Z. Physik Chemie* 22 (1933) 350–8

[25] Barnes, R. S. *Metal Treatment and Drop Forging.* 18 (1951) 531–8

[26] Turnbull, D. *Atom Movements* American Society for Metals (1951)

[27] Fisher, J. C. *J. appl. Phys.* 22 (1) (1951) 74–77

[28] Langmuir, I. *J. Franklin Inst.* 217 (1934) 543–69

[29] Hoffman, R. E. and Turnbull, D. *J. Appl. Phys.* 22 (5) (1951) 634–9

[30] Le Claire, A. D. *Phil, Mag.* 42 (1951) 468–74

[31] Johnson, W. A. *Trans. Amer. Inst. Min. Met. Engrs.* 143 (1941) 107–11

[32] Fensham, P. J. *Aust. J. Sci. Res.* 3 (1) (1950) 105–8

[33] Barnes, R. S. *Nature, Lond.* 166 (1950) 1032

[34] Beck, P. A., Sperry, P. R., and Hu, H. *J. appl. Phys.* 21 (1950) 420–25

[35] Achter, M. R. and Smoluchowski, R. *J. appl. Phys.* 22 (10) (1951) 1260–1264

[36] Berman, F. and Harrington, R. H. *Trans. Amer. Soc. Metals* 36 (1945) 362

[37] Read, W. T. and Shockley, W. *Phys. Rev.* 78 (1950) 275–289

[38] Achter, M. R. and Smoluchowski, R. *Physics of Powder Metallurgy* McGraw Hill 1951; and *Phys. Rev.* 76 (1949) 470

[39] Dawson, M. H. *J. appl. Phys.* 23 (1952) 373

[40] Achter, M. R. and Smoluchowski, R. *J. appl. Phys.* 23 (1952) 373–4

[41] —— —— *Phys. Rev.* 83 (1951) 163–4

[42] FLANAGAN, R., and SMOLUCHOWSKI, R. *J. App. Phys.* 23 (7) (1952) 785–7
[43] MOTT, N. F. *Proc. Phys. Soc.* 60 (1948) 391–4
[44] LOMER, W. M. and NYE, J. F. *Proc. Roy. Soc.* 212 (1952) 576–84
[45] KÊ, T'ING SUI *Phys. Rev.* 73 (1948) 267–8
[46] BUFFINGTON, F. S., BAKELAR, I. D. and COHEN, M. *Physics of Powder Metallurgy.* Chapt. VI, McGraw Hill 1951
[47] FENSHAM, P. J. *Aust. J. Sci. Res.* 3 (1) (1950) 91–104
[48] ROTHERHAM, L., SMITH, A. D. N. and GREENOUGH, G. B. *J. Inst. Met.* 79 (6) (1951) 439–54
[49] BIRCHENALL, C. E. and MEHL, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 144–9
[50] BARRER, R. M. *Diffusion in and through Solids* C.U.P. 1941
[51] NICKERSON, R. A. and PARKER, E. R. *Trans. Amer. Soc. Metals.* 42 (1950) 376–86
[52] FRAUNFELDER, H. *Helv. phys. Acta* 23 (4) (1950) 347–79
[53] MÜLLER, E. W. *Z. Physik.* 126 (1949) 642–65
[54] VAN LIEMPT, J. A. M. *Rec. trav. Chim.* 64 (1945) 239–49
[55] MEHL, R. F. *Trans Amer. Inst. Min. Met. Engrs.* 122 (1936) 1
[56] WERT, C. A. and ZENER C. *Phys. Rev.* 76 (1949) 1169–75
[57] —— *Phys. Rev.* 79 (4) (1950) 601–5
[58] —— *J. Appl. Phys.* 21 (11) (1950) 1196–7
[59] SEITZ, F. *Phys. Rev.* 74 (10) (1948) 1513–23
[60] —— *Phys. Rev.* 79 (5) (1950) 890–1
[61] KÊ, T'ING SUI *J. appl. Phys.* 20 (3) (1949) 274–80
[62] GLASSTONE, S., LAIDLER, K. J. and EYRING, H. *Theory of Rate Processes* McGraw Hill 1941
[63] KÊ, T'ING SUI *Phys. Rev.* 71 (8) (1947) 533–46
[64] KÖSTER, W. *Z. Metallk.* 39 (1948) 1–9
[65] RUSHBROOKE, G. S. *Trans. Faraday Soc.* 36 (10) (1940) 1055–62
[66] HUNTINGTON, H. B. and SEITZ, F. *Phys. Rev.* 61 (1942) 315–25
[67] —— *Phys. Rev.* (1942) 325–8
[68] —— *Phys. Rev.* 76 (1949) 1728
[69] ZENER, C. *Acta Cryst.* 3 (5) (1950) 346–54
[70] PANETH, H. R. *Phys. Rev.* 80 (4) (1950) 708–11
[71] MANINTVELD, I. *Nature, Lond.* 169 (1952) 623
[72] NOWICK, A. S. *Phys. Rev.* 82 (1951) 551
[73] MOTT, N. F. *Proc. Phys. Soc.* B. 64 (1951) 729–41
[74] SEITZ, F. *Phys. Rev.* 80 (1950) 239
[75] SMOLUCHOWSKI, R. and BURGESS, H. *Phys. Rev.* 76 (1949) 309–10
[76] NIX, F. C. and JAUMOT, F. E. *Phys. Rev.* 83 (1951) 1275
[77] BRADLEY, A. J. and SEAGER, G. C. *J. Inst. Met.* 64 (1939) 81
[78] *Metals Reference Book* (Ed. C. J. SMITHELLS) London 1949
[79] LE CLAIRE, A. D. To be published
[80] KUBASCHEWSKI, O. *Trans, Faraday Soc.* 46 (9) (1950) 713–722
[81] ZENER, C. *J. appl. Phys.* 22 (4) (1951) 372–5
[82] FISHER, J. C. and HOFFMANN, R. E. *Gen. Elec. Res. Lab. Rept.* No. RL. 330
[83] ZAGRUBSKIJ, A. M. *Bull. Acad. Sci. U.S.S.R. (Phys)* 6 (1937) 903–13
[84] SEITH, W. and KIEL, A. *Z. Metallk.* 25 (1933) 104–6. The values of Q and D_0 given in Table 2 were obtained by a recalculation from the data in this paper
[85] NOWICK, A. S. *J. appl. Phys.* 22 (9) (1951) 1182–6
[86] JOHNSON, R. P. *Phys. Rev.* 56 (1939) 814–18
[87] DIENES, G. J. *J. appl. Phys.* 21 (1950) 1189–92

[88] MARTIN, A. B. and ASARO, F. *North American Aviation Inc. Rept.* NAA–S.R.–83 (1950) and *Phys. Rev.* 80 (1950) 123
[89] RUDER, R. C. and BIRCHENALL, C. E. *J. of Metals* 191 (1951) 142–146
[90] LE CLAIRE, A. D. Unpublished work
[91] KÖSTER, W. and RAUSCHER, W. *Z. Metallk.* 39 (1948) 111
[92] SEITH, W. *Diffusion in Metallen* J. Springer, Berlin 1939
[93] ALEXANDER, B. H., DAWSON, M. H. and KLING, H. P. *J. appl. Phys.* 22 (4) (1951) 439–43
[94] HERRING, C. *J. appl. Phys.* 21 (1950) 437–45
[95] KUCZYNSKI, G. C. *J. of Metals* 1 (2) (1949) 169–178
[96] DEDRICK, J. H. and GERDS, A. *J. appl. Phys.* 20 (11) (1949) 1042–44
[97] KUCZYNSKI, G. C. *J. appl. Phys.* 21 (7) (1950) 632–5
[98] COHEN, G. and KUCZYNSKI, G. C. *J. appl. Phys.* 21 (12) (1950) 1339–40
[99] CABRERA, N. *J. of Metals* Trans. Amer. Inst. Min. Met. Engrs. 188 (1950) 667
[100] SCHWED, P. *J. of Metals* 3 (3) (1951) 245–6
[101] DEDRICK, J. H. and KUCZYNSKI, G. C. *J. appl. Phys.* 21 (12) (1950) 1224–5
[102] SNOEK, J. *Physica* 8 (1941) 711–733
[103] ZENER, C. *Phys. Rev.* 71 (1947) 34–8
[104] —— *Elasticity and Anelasticity of Metals* Chicago University Press 1947
[105] LE CLAIRE, A. D. *Phil. Mag.* 42 (1951) 673–88
[106] —— and CHILDS, B. G. To be published
[107] NOWICK, A. S. Quoted from *Atom Movements* Amer. Soc. for Metals 1951 & Phys. Rev. 82 (1951) 340
[108] KÊ, T'ING SUI *J. appl. Phys.* 19 (1948) 285–90
[109] STANLEY, J. K. *J. of Metals* Trans. Amer. Inst. Min. Met. Engrs. 185 (1949) 752–61
[110] KÊ, T'ING SUI *Phys. Rev.* 74 (1948) 9–15
[111] —— *Phys. Rev.* 74 (1948) 914–16
[112] BIRCHENALL, C. E. and MEHL, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* 171 (1947) 143–58
[113] ONSAGER, L. and FUOSS, R. M. *J. phys. Chem.* 36 (1932) 2689–2778
[114] HARTLEY, G. S. and CRANK, J. *Trans. Farad. Soc.* 45 (1949) 801–18 and PMP.1
[115] SEITH, W., HERMANN, T. and KOTTMANN, A. *Naturwiss.* 39 (2) (1952) 41
[116] ONSAGER. L. *Phys. Rev.* 37 (1931) 405–26, 38 (1931) 2265–79
[117] GROOT, S. R. DE *Thermodynamics of Irreversible Processes* North Holland Publishing Co. 1951
[118] DENBIGH, K. G. *Thermodynamics of the Steady State* Methuen, London 1951
[119] STEIGMAN, J., SHOCKLEY, W. and NIX, F. C. *Phys. Rev.* 56 (1939) 13–21
[120] BARDEEN, J. *Phys. Rev.* 76 (9) (1949) 1403–5
[121] SHUTTLEWORTH, R. Unpublished work
[122] ZENER, C. *Trans. Amer. Inst. Min. Met. Engrs* 152 (1943) 122–26
[123] RHINES, F. N. and MEHL, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* 128 (1938) 185
[124] RICHTER, G. *Ann. Physik* 32 (1938) 683
[125] STEWART, M. T., THOMAS, R., WAUCHOPE, K., WINEGARD, W. C. and CHALMERS, B. *Phys. Rev.* 83 (3) (1951) 657
[126] POLDER, D. *Philips Res. Repts.* 1 (1) (1945) 6–12
[127] DIENES, G. J. *J. appl. Phys.* 22 (1951) 848–9
[128] SEITH, W. and KOTTMANN, A. *Angewandte Chemie* 64 (4) (1952) 376–406
[129] BALUFFI, R. W. and ALEXANDER, B. H. *U.S. Atom. Energy Comm. Rept.* SEP–86 & 31 App. Phys. 23 (9) (1952) 953–6
[130] SMOLUCHOWSKI, R. *Phys. Rev.* 87 (3) (1952) 482–87
[131] DUGDALE, R. A. *Phil. Mag.* 43 (1952) 912–14

7

NUCLEATION

J. H. Hollomon and D. Turnbull

Introduction

In a large number of chemical reactions there are no recognizable centres at which the reaction begins and from which it propagates. In contrast to this mode of reaction most phase transformations begin at certain sites and propagate from these places. An element of the system undergoes no recognizable change until it is swept over by a boundary between the transformed and untransformed phases propagating from the centre at which the transformation began.

The initiation of transformation at certain recognizable centres is known as nucleation, and the propagation of the transformation from the centres is called growth. Even though the nucleation of transformations has been recognized and described for at least two centuries, until recent years it was difficult to make reproducible and quantitative measurements of the rate of nucleation. During the past two decades semi-quantitative measurements have been made on the rate of nucleation of liquid droplets in supersaturated vapours. These measurements largely confirmed the predictions of the theory of homogeneous nucleation developed by Volmer, Becker and co-workers. More recently the rate of nucleation in supercooled liquids has been measured quantitatively. Also our insight into the nature and mechanism of nucleation catalyzed by extraneous surfaces has been deepened by recent experimental and theoretical developments. Even though the understanding of nucleation in crystalline media is much less advanced than for gaseous and liquid media, an important theoretical groundwork for this understanding has been developing.

These advances in the understanding of nucleation have spurred great experimental and theoretical activity in the past few years. Descriptions of some of this activity are given in recent reviews by Smoluchowski,[1] Turnbull and Hollomon,[1a] Hollomon,[2] Bradley,[3] and La Mer, Pound and Reiss.[4]

In this paper we shall critically review the subject of nucleation in general but particularly emphasize the rate of nucleation in reactions involving crystalline phases.

Consider the formation of a phase β within a parent phase α by the process of nucleation and growth. If the transformation occurs isothermally the size of each β region, D, is some function of time $D = f(t)$. If no change in composition accompanies the transformation

β propagates at a constant rate, G, after nucleation and $D = G(t - \tau)$, where τ, usually a positive intercept on the time axis, is the nucleation period. Of course, all regions of β do not nucleate at the same time, so to describe the course of the transformation we must know more than one value of τ. It is most convenient to define the nucleation frequency, I, which is the number of β regions formed per unit of time per unit of untransformed volume. I is reciprocally related to a mean value of $\tau = \bar{\tau}$. To measure I experimentally we effectively must measure $D = f(t)$ for a large number of β regions at a given stage in the transformation, extrapolate all these to zero, and determine the number, N, of intercepts per unit time. Then $I = N/V$, where V is the untransformed volume corresponding to the particular time. In order for the I value to be of significance N must be very large. Where a composition change accompanies transformation, as in precipitation, often it is found that the growth of β is diffusion limited and $D = G^{1/2}(t - \tau)^{1/2}$. For such transformations I is determined similarly.

The kinetics of a large number of isothermal transformations, probably the majority, can be satisfactorily described by the two parameters I and G. Knowing I and G and taking account of the geometrical problem of the impingement of growing β regions, the dependence upon time of the fraction of the α sample transforming to β may be derived. Derivations of this dependence have been given by JOHNSON and MEHL[5] and AVRAMI[6] (now MELVIN). These derivations recently have been reviewed critically by BURKE and TURNBULL.[7]

Examples of phase transformations satisfactorily described as nucleation and growth processes are vapour $\leftrightarrows$ liquid, vapour $\leftrightarrows$ crystal, liquid $\leftrightarrows$ crystal, precipitation, allotropic transformation, and in a few instances disorder $\rightarrow$ order. A variety of processes that are not phase transformations in the ordinary sense can also be described by nucleation and growth kinetics. An example is the recrystallization of cold worked metals treated in detail in an earlier review.[7]

Nucleation that takes place at random within homogeneous elements of volume of a phase is designated *homogeneous nucleation.* Experience indicates that transformation nuclei often form preferentially at various interfaces in the system, such as free surfaces, internal surfaces, container walls, suspended crystallites, etc. Such interfaces are said to catalyze nucleation and this process of nucleation is called *heterogeneous nucleation.*

We will describe and explain nucleation in phase transformations categorized according to its complexity. In many phase changes, such as condensation of a pure vapour or crystallization of a supercooled liquid, only a change in *structure* is involved. The crystallization of a phase from a solution involves a change in *structure* and *composition.* Volume changes accompany most phase transformations; therefore,

when a phase forms in a crystalline medium strain energy is developed due to the transformation. If no composition change is involved, as in allotropic transformations in pure crystals, there is *structure* and *strain* to be considered. The precipitation of a phase from a solid solution is the most complex type of transformation in that it is accompanied by changes in *structure*, *composition* and *strain*.

Structure

Nucleation of Liquid from Vapour

General Theory—Even though our primary purpose in writing this review is to describe and explain nucleation in condensed metal systems, it is desirable to re-examine the theory for nucleation of liquids from supersaturated vapours and its experimental verification. The theories for the nucleation of a liquid from a supersaturated vapour have influenced the development of theories for nucleation in condensed systems.

The theory for homogeneous nucleation of liquids in vapours as now generally accepted is largely due to Volmer and Weber[8] and Becker and Döring.[9]† Volmer and Weber supposed that dense clusters of molecules exist in equilibrium in the vapour and that some of these clusters (called embryos) may become nuclei for the growth of liquid droplets. The essential assumption of the theory is that the standard free energy, ΔG, of these embryos can be calculated as though they are liquid droplets having the same liquid-vapour energy per area, σ, and the same free energy per volume, ΔG_v, with reference to the vapour phase, as the macroscopic liquid phase. Thus, for spherical embryos of radius, r, at a given temperature and external pressure,

$$\Delta G = 4\pi r^2\sigma + (\tfrac{4}{3}\pi r^3)(\Delta G_v), \qquad \ldots \quad (1)$$

where σ and ΔG_v are independent of r.

In general,

$$\Delta G = a\sigma i^{2/3} + b\Delta G_v i, \qquad \ldots \quad (2)$$

where i is the number of molecules per cluster and a and b are factors determined by the shape and density of the cluster.

A second assumption of the theory[8] is that embryos are born, grow to nuclei, or disappear by a sequence of bimolecular reactions. Designating α as a molecule of vapour and β_i an embryo of liquid containing i molecules, these sequences of bimolecular processes can be represented in the following way:

$$\begin{aligned} \alpha + \alpha &\leftrightarrows \beta_2 \\ \beta_2 + \alpha &\rightleftarrows \beta_3 \\ \beta_{i-1} + \alpha &\rightleftarrows \beta_i \end{aligned} \qquad \ldots \quad (3)$$

† Called hereafter the V.W.B.D. theory.

With these assumptions the absolute frequency of nucleation per volume, I, of liquid nuclei in supersaturated vapours can be derived.

Let the number of embryos per volume containing i molecules in equilibrium with vapour at a given pressure P and absolute temperature T be n_i. Then assuming ideal mixing of embryos of all sizes and that $n_1 \gg \sum_{i=2}^{i=\infty} n_i$, where n_1 is the number of single molecules per volume, it follows that:

$$n_i = n_1 \exp [- \Delta G/kT] \quad . \quad . \quad . \quad . \quad (4)$$

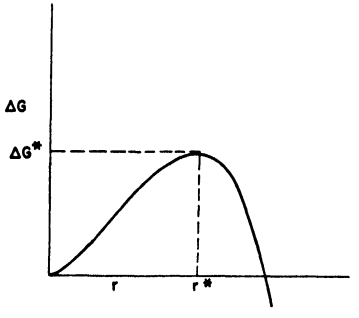


Fig. 1. Free energy of an embryo of a new phase as a function of its radius

When $\Delta G_v < 0$, $\Delta G = f(r)$ (equation (1)) has a maximum (see Fig. 1):

$$\Delta G^* = 16\pi\sigma^3/3(\Delta G_v)^2, \quad . \quad . \quad . \quad . \quad (5)$$

at

$$r = r^* = - 2\sigma/\Delta G_v. \quad . \quad . \quad . \quad . \quad (6)$$

Embryos having a radius smaller than this critical value will on the average disappear, but those larger will grow into liquid droplets. An embryo having the critical radius is by definition a nucleus. The number of embryos per volume-time becoming nuclei is the nucleation frequency,† I.

I is the net frequency per volume of the bimolecular process:

$$\beta_{i^*} + \alpha \rightleftarrows \beta_{i^*+1}, \quad . \quad . \quad . \quad . \quad (7)$$

where i^* is the number of molecules in the nucleus.

Assuming with VOLMER and WEBER an equilibrium distribution of embryos of all sizes, the number of nuclei per volume, n_{i^*}', is given by:

$$n_{i^*}' = n_{i^*} = n_1 \exp [- \Delta G^*/kT]. \quad . \quad . \quad . \quad (8)$$

Then neglecting the reverse reaction indicated in (7) we obtain the expression:

$$I = zs^*n_{i^*}, \quad . \quad . \quad . \quad . \quad (9)$$

where s^* is the surface area of the nucleus and z is the collision frequency of vapour molecules with the surface of the nucleus per area. From kinetic theory $z = P/\sqrt{2\pi mkT}$ and substituting (8) into (9) we have:

$$I = [P/(2\pi mkT)^{1/2}]s^*n_1 \exp [- \Delta G^*/kT]. \quad . \quad . \quad . \quad (10)$$

BECKER and DÖRING improved this treatment by taking into account the reverse process indicated in (7), and the decrease in concentration of

† This definition does not correspond exactly to the operational definition of I given in the introduction. However, on the basis of the V.W.B.D. theory the difference in the two I values is negligible in most instances.

embryos during a transition from the equilibrium concentration because of the rapid growth of nuclei to droplets.

Schematic embryo size distribution curves calculated from these two points of view are compared in Fig. 2. The expression derived by BECKER and DÖRING is:

$$I = zn_1 s^*(-b\Delta G_v/2\pi i^* kT)^{1/2} \exp(-\Delta G^*/kT) \quad . \quad . \quad . \quad . \quad (11)$$

Excepting for the factor $(-b\Delta G_v/2\pi i^* kT)^{1/2}$, this equation for I is identical with equation (10).

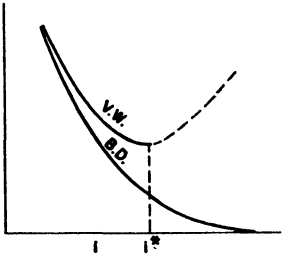


Fig. 2. Comparison between VOLMER–WEBER and BECKER–DÖRING distributions of embryos with respect to size

The basic assumption that the free energy of embryos can be calculated by assuming them to be small droplets of the macroscopic liquid phase, recently has been examined critically by KIRKWOOD and BUFF[10] and REISS.[11] The primary criticism derives from the fact that the nucleation rate can be perceptible only when the nucleus diameter is of the order of 10 molecular diameters or less (equation (6)). When the size is this small there is some ambiguity in the definition of surface tension. TOLMAN[12] and

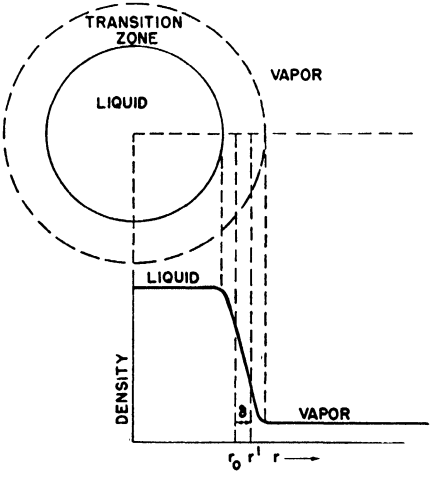


Fig. 3. Variation of density of matter across transition zone between liquid droplet and vapour

KIRKWOOD[13] have developed theories for the variation in the density of matter across the liquid-vapour interface which predict that the density changes from the liquid value to the vapour value in a transition zone (see Fig. 3), of the order of 2–5 molecular diameters in thickness.

As a consequence of this transition zone, these authors show that the

surface energy σ_r of a small droplet of radius r should not be identical with $\sigma = \sigma_0$ for plane surfaces but should be related to σ_0 by the equation:

$$\sigma_r = \sigma_0/(1 + 2\delta/r_0), \quad \ldots \quad (12)$$

where δ is the difference between the value of r at the Gibbs surface of tension and its value ($r = r_0$) where the superficial density of matter vanishes. According to TOLMAN and KIRKWOOD, δ is of the order of one to two molecular diameters—hence, when δ/r_0 is 0·10 to 0·20, σ_r is 0·8 to 0·7 σ_0.

KIRKWOOD and BUFF[10] and BUFF[14] though expressing reservations about the validity of using macroscopic thermodynamic parameters to describe the properties of nuclei, have modified equation (5) by taking into account the variation of σ with curvature. BUFF'S development gives, instead of equation (5), the following expression for ΔG^*:

$$\Delta G^* = \frac{4\pi y}{(\Delta G_v)^2}\{\sigma_0[y + 2\delta(\Delta G_v)] - y^{2/3}\} \quad \ldots \quad (13)$$

where $y = \sigma_0 + (\sigma_0^2 + 2\sigma_0\delta\Delta G_v)^{1/2}$.

The ambiguity in the definition of the energy of interfaces having radii of curvature of only a few molecular diameters is not confined to liquid vapour interfaces but is inherent in the corresponding definitions for all types of interfaces. This problem must be reckoned with in the treatments of nucleation of crystalline phases in condensed systems.

The V.W.B.D. theory and its modification tentatively proposed by KIRKWOOD and BUFF make substantially different predictions of the value of ΔG_v at which the rate of nucleation becomes perceptible. Alternatively, the V.W.B.D. theory predicts that the values of the interfacial energy, σ, calculated from rates of nucleation should be in close agreement with that of σ_0, measured for flat bulk liquid surfaces while the KIRKWOOD–BUFF modification predicts that σ should be substantially less than σ_0.

Transient Nucleation—BECKER and DÖRING'S final equation for I was derived on the assumption that I is time independent. However, when the driving free energy of a phase transition is changed very rapidly from one level, $(\Delta G_v)_1$, to another, $(\Delta G_v)_2$, there is a transient during which the embryo distribution characteristic of $(\Delta G_v)_1$ changes to the distribution characteristic of $(\Delta G_v)_2$. During this time the nucleation frequency I_t is time dependent.

The earliest formal treatment of transient nucleation is due to ZELDOVICH,[15] who derived an approximate relation $I_t = f(t)$. TURNBULL[16] pointed out the importance of transients in solid-solid transformations and evaluated $I_t = f(t)$ precisely by a numerical method for

certain simple cases. More recently KANTROWITZ[17] and PROBSTEIN[18] have developed the theory for transient nucleation of liquid in supersaturated vapours, which is of importance in interpreting the results of experiments in which the supersaturation is increased with great rapidity. With some simplifying assumptions ZELDOVICH derived from the BECKER–DÖRING theory the following equation for I_t:

$$I_t \approx I \exp [- i^{*2}/4zs^*t] \approx I \exp [- \tau/t], \quad . \quad . \quad . \quad . \quad (14)$$

where I is the steady nucleation frequency. KANTROWITZ takes into account the possible departure of the accommodation coefficient from unity and the difference in temperature between the embryos and their surroundings due to the heat liberated by the condensation. With these modifications, zs^* in equation (14) is replaced by a function of the appropriate accommodation coefficients and heat content quantities. Without making the corrections introduced by KANTROWITZ the order of time τ required for the transient in the nucleation of liquid droplets to expire is a microsecond.

Formal Theory of Heterogeneous Nucleation—It is probable that in most phase transformations the formation of nuclei is catalyzed by various foreign surfaces present in the system. In experiments designed to determine the rate of nucleation, the most difficult problem is the elimination of these surfaces or the specification of the nature of the substances effecting nucleation catalysis.

VOLMER[19] started with the basic assumptions of the V.W.B.D. theory and developed a formal theory for heterogeneous nucleation of liquid from supersaturated vapour. Suppose that a liquid embryo comes to equilibrium on a plane foreign surface with a contact angle θ, it follows that:

$$\sigma_{SV} = \sigma_{LS} + \sigma_{LV} \cos \theta, \quad . \quad . \quad . \quad . \quad (15)$$

where σ_{SV} is the interfacial energy per area between vapour and catalyst, σ_{LS} is the interfacial energy per area between liquid and catalyst, and σ_{LV} is the interfacial energy per area between liquid and vapour and is identical with σ in the theory of homogeneous nucleation. The equilibrium form of the embryo is a spherical cap and the free energy of formation is:

$$\Delta G_s = [\pi r^2 \sigma_{LV} + (\pi r^3/3)\Delta G_v][2 - 3 \cos \theta + \cos^3 \theta], \quad . \quad . \quad . \quad . \quad (16)$$

where r is the radius of the sphere. For all $\theta > 0$, ΔG_s goes through a maximum having the ΔG_s co-ordinate.

$$\Delta G_s^* = \Delta G^* f(\theta), \quad . \quad . \quad . \quad . \quad (17)$$

where $f(\theta) = (2 + \cos \theta)(1 - \cos \theta)^2/4$, and ΔG^* is given by equation (5). We may suppose that the embryos are in equilibrium with atoms

adsorbed on the catalyst surface from the condensing vapour. If the number of adsorbed vapour atoms is n_s/area, the equilibrium number of embryos per area of a given size on the catalyst surface, n_{s_i}, is

$$n_{s_i} = n_s \exp [- \Delta Gf(\theta)/kT]. \quad . \quad . \quad . \quad . \quad (18)$$

Following a development analogous to that in deriving equation (11) we find approximately for the nucleation frequency per area of catalyst surface, I_s, the following[20]:

$$I_s = zn_s s^*(- b\Delta G_v/2\pi i^* kT)^{1/2} \exp [- \Delta G^* f(\theta)/kT]. \quad . \quad . \quad . \quad . \quad (19)$$

Consider a unit volume of vapour containing a unit surface of a catalyst characterized by a single contact angle θ. Since $n_s \sim 10^{-8} n_1$, it follows from equations (11) and (19) that for this system,

$$\log_{10} (I_s/I) = - 8 + (\Delta G^*/2 \cdot 3kT)[1 - f(\theta)]. \quad . \quad . \quad . \quad (20)$$

I is perceptible when $(\Delta G^*/2 \cdot 3kT) \sim 25$; therefore, I_s is of the same magnitude or greater than I for all $f(\theta) \gtrsim 2/3$ corresponding to $\theta \gtrsim 102°$. When $\theta = 0$, I_s can be perceptible with an infinitesimal degree of supersaturation.

Comparison of Theory and Experiment—Nucleation of liquid phases is catalyzed by a large variety of surfaces, and in most vapour-liquid transitions the supersaturation is relieved by diffusion of vapour to liquid heterogeneously nucleated on these surfaces. In order to obtain sufficient supersaturation to cause a perceptible rate of homogeneous nucleation the vapour is freed, as far as possible, from suspended particles and it is rapidly chilled by adiabatic expansion so that the supersaturation ratio P/P_0,† builds up rapidly. With these conditions the value of P/P_0 required for perceptible homogeneous nucleation can be reached before an appreciable amount of vapour has diffused to surfaces on which liquid has been heterogeneously nucleated.

In experiments on nucleation of liquid droplets[21] in supersaturated vapours P/P_0 is progressively increased in a series of different adiabatic expansions until a fog of liquid droplets is seen to form in the body of the vapour. The significant datum obtained from the experiment is the minimum value of $(P/P_0)_m$ necessary to cause this fog formation. Experience indicates that the fog appears in these experiments at a fairly sharply defined and reproducible supersaturation ratio. This result is in agreement with the prediction of the V.W.B.D. theory that I increases sharply with increasing P/P_0, e.g. an increase of 2 per cent in P/P_0 may cause a tenfold increase in I. Because of this sharp dependence of the rate of nucleation on the supersaturation ratio, a

† P is the vapour pressure and P_0 is the vapour pressure in equilibrium with a flat surface of the corresponding liquid.

large error in assigning a value to I corresponding to $(P/P_0)_m$ has little effect on the value of σ calculated from the nucleation data, i.e. the exponential term dominates equation (11). Therefore, in the VOLMER–FLOOD type of experiment, $(P/P_0)_m$ is measured accurately and I is estimated within wide limits of uncertainty, within a factor of 100.

VOLMER and FLOOD[21] have measured $(P/P_0)_m$ for various temperatures for water and some organic compounds. For an ideal vapour,

$$\Delta G_v = (RT/V)\ln(P/P_0), \quad \ldots \quad (21)$$

where V is the volume per mole of the liquid phase. VOLMER and FLOOD[21] compared their measured values of $\ln(P/P_0)_m$ with values calculated from equation (11), where ΔG^* is given by (5), ΔG_v by (21), and assuming that σ is equal to σ_0.

Alternatively, σ may be calculated from (11), by substitution of the measured value of $(P/P_0)_m$, and compared with σ_0. The comparison of σ with σ_0, the measured liquid-vapour interfacial energy for infinite radius of curvature, is shown in Table 1.

TABLE 1

Comparison of measured liquid-vapour interfacial energies with values calculated from VOLMER–FLOOD experiments.
After TURNBULL and HOLLOMON[1a]

Substance	*Temperature* °K	σ	
		Calculated from VOLMER–FLOOD *data*	*Measured*
Water . . .	275·2	75·5	75·2
Methyl Alcohol .	270·0	37·5	24·8
Ethyl Alcohol .	273·0	24·4	24·0
n-Propyl Alcohol .	270·0	26·2	25·4
iso-Propyl Alcohol .	265·0	23·6	23·1
n-Butyl Alcohol .	270·0	26·2	26·1
Nitromethane .	252·0	41·1	40·6

With the single exception of the result for methyl alcohol, where the deviation is 50 per cent, σ and σ_0 agree within 2–3 per cent and well within the experimental uncertainty. From these results it is concluded that the V.W.B.D. theory describes the frequency of homogeneous nucleation of liquid droplets in pure vapours to within the limits of accuracy attained in these experiments. The diameter of the nucleus

calculated from the VOLMER–FLOOD data is of the order of 5 molecular diameters and certainly of the same order of magnitude as the δ of the TOLMAN and KIRKWOOD theories. The data are in disagreement with the prediction of BUFF'S modified theory that $\sigma \sim 0{\cdot}8\sigma_0$ for the experimental conditions of VOLMER and FLOOD.[21]

SANDER and DAMKÖHLER[22] have measured the rate of nucleation of water droplets in supersaturated water vapour at lower temperatures than those attained by VOLMER and FLOOD. Their calculated values of σ are about 10 per cent less than σ_0 at the lowest temperature of their measurements. Further, SANDER and DAMKÖHLER'S σ appears to go through a maximum with decreasing temperature in qualitative agreement with BUFF'S theory. However, in the temperature range of overlapping measurements SANDER and DAMKÖHLER'S σ is about 7 per cent less than VOLMER and FLOOD'S. Recently POUND and MADONNA[23] have repeated the experiments on nucleation of water including in their measurements the temperature range covered by VOLMER and FLOOD and SANDER and DAMKÖHLER. POUND and MADONNA'S values of σ are in good agreement with those of VOLMER and FLOOD and with σ_0 down to a nucleus containing about 70 molecules. For $i^* < 70$, $\sigma_0 > \sigma$ and at $i^* \sim 50$, $\sigma \sim 0{\cdot}9\sigma_0$.

WEGENER and SMELT[24] found that the degree of supersaturation reached in the very rapid expansion of moist air is much greater than would seem to be permitted by the V.W.B.D. theory. KANTROWITZ suggested that the supersaturation in WEGENER and SMELT'S experiments reaches a very high level because of the time lag in reaching the steady state value of I predicted by the V.W.B.D. theory. KANTROWITZ showed that these results are precisely described by the approximate theory of transient nucleation with a single disposable parameter —the accommodation coefficient, α, of water vapour molecules on a liquid water surface. However, the value of $\alpha \sim 5 \times 10^{-4}$ required for the agreement seems too low.

It appears, therefore, that the kinetics of nucleation of liquid droplets in supersaturated vapours is adequately described by the V.W.B.D. theory and it seems permissible to ignore the effect of droplet radius on its surface energy. The success of the V.W.B.D. theory encourages us to explore the consequences of the use of analogous assumptions in extending the theory and in describing the nucleation of crystalline phases.

Theory of Nucleation in Condensed System

Most of the theories of homogeneous nucleation in condensed systems start with the basic assumptions of the V.W.B.D. theory. Thus the free energy for formation of embryos in condensed systems is assumed to be given by an equation identical in form with equation (2) with the

appropriate σ and ΔG_v values and values of a and b consistent with an assumed shape of the nucleus. ΔG^* is given by

$$\Delta G^* = K\sigma^3/(\Delta G_v)^2 \quad \ldots . \quad (22)$$

where K is determined by the nucleus shape and is $16\pi/3$ for a spherical nucleus.

For nucleation in condensed systems the kinetic theory expression, z, for the rate of bombardment of the nucleus by molecules of the parent phase can, in general, no longer be used. The movement of an atom across an interface between two condensed phases usually requires a free energy of activation, ΔG_A. On this basis BECKER[25] derived a general expression for the rate of nucleation in condensed systems.

Recently TURNBULL and FISHER,[26] again following the basic assumptions of the V.W.B.D. theory, applied the absolute reaction rate formalism to nucleation in condensed systems and derived a more quantitative expression for I. In terms of the absolute rate theory the frequency, ν, for transfer of an atom across an interface is:

$$\nu = (kT/h) \exp [- \Delta G_A/kT], \quad \ldots . \quad (23)$$

where h is Planck's constant. Following the methods of BECKER and DÖRING they derived an expression for the rate of nucleation:

$$I = K_v \exp [- (\Delta G^* + \Delta G_A)/kT] \quad \ldots . \quad (24)$$

where $K_v = n^*(a\sigma/9\pi kT)^{1/2} n(kT/h)$, a (see equation (2)) is determined by the nucleus shape, n^* is the number of atoms in the surface of a nucleus of critical size, and n is the number of atoms per volume of the parent phase.

For spherically shaped nuclei,

$$(a\sigma/9\pi kT)^{1/2} = (\sigma/kT)^{1/2}(2v/9\pi)^{1/3}, \quad \ldots . \quad (25)$$

where v is volume per atom of the parent phase.

Following the treatment in the preceding section we may write for the transient rate of nucleation in condensed systems:

$$I_t \approx I \exp [- i^{*2}/4n^*\nu t], \quad \ldots . \quad (26)$$

where ν is given by equation (23). Since ν can be very small the transient period, $\tau = i^{*2}/4n^*\nu$, can be much longer for condensed systems[16] than for the nucleation of liquid from vapour.

The temperature coefficient of I, dI/dT is found, by differentiation of equation (24), to be:

$$dI/dT = I\{(\Delta G^* + \Delta G_A)/kT^2 - (1/kT)(d[\Delta G^*]/dT)\}, \quad \ldots . \quad (27)$$

where ΔG_A is assumed to be independent of T. I goes through a maximum $I = I_{\max}$ at $T = T_{\max}$. $T_{\max}$ is found by solving the equation

$$(\Delta G^*_{\max} + \Delta G_A)/T_{\max} = [d(\Delta G^*)/dT]_{T=T_{\max}}, \quad \ldots . \quad (28)$$

obtained by setting $dI/dT = 0$. From the thermodynamic identity, $(dG/dT)_p = -S$, it follows that

$$[d(\Delta G_v)/dT]_p = -\Delta S_v, \quad \ldots . \quad (29)$$

where ΔS_v is the entropy change per volume accompanying the phase transformation. Often it can be assumed to a fair approximation that ΔS_v is temperature independent whence it follows:

$$\Delta G_v = \Delta S_v(T - T_0), \quad \ldots . \quad (30)$$

where T_0 is the temperature at which the two phases of infinite mass are in equilibrium.

Substitution of equations (22) and (30) into (28) gives a general relation:

$$(T_{\max} - T_0)^3 + [K\sigma^3/\Delta G_A(\Delta S_v)^2](T_{\max} - T_0) + 2K\sigma^3 T_{\max}/\Delta G_A(\Delta S_v)^2 = 0, \quad \ldots . \quad (31)$$

that can be solved for $T_{\max}$. When $\Delta G_A = 0$, $T_{\max} = T_0/3$. If $\Delta G_A > 0$, $T_{\max} > T_0/3$.

The formal treatment of VOLMER for nucleation of liquid from vapour on catalytic surfaces has been extended to the nucleation of crystalline phases in condensed systems.[27] Consider, for example, the nucleation of a crystal in a supercooled liquid. If it is assumed that the liquid-crystal interfacial energy σ_{LC} is isotropic, equations (15) and (16) apply with the subscripts SL, SC and LC replacing SV, LS and LV, respectively, where S refers to catalyst, L to liquid and C to the forming crystal having the same composition as the liquid. By a simple extension of the treatment of TURNBULL and FISHER[26] the following expression for I_s for crystalline phases in condensed systems is obtained[28]:

$$I_s = K_s \exp[-\Delta G^* f(0)/kT], \quad \ldots . \quad (32)$$

where for nuclei shaped like spherical caps,

$$K_s = n^*(\sigma/kT)^{1/2}(2v/9\pi)^{1/3}[f(\theta)]^{1/6} n_s(kT/h) \exp[-\Delta G_A/kT], \quad \ldots . \quad (33)$$

σ is the interfacial energy per area between forming crystal and the parent phase.

As has been pointed out earlier[27] transients are as likely to be important in heterogeneous as in homogeneous nucleation. With I_s substituted for I and using the appropriate values of i^* and n^*, equation (26) gives an approximate relation for the transient in heterogeneous nucleation.

Let a phase α be in equilibrium with a phase β having the same composition at temperature T_0. Below T_0, β is more stable and above T_0, α is more stable. Suppose β is heated to a temperature in the range of α stability. It follows that at this temperature β is less stable than α

on all plane foreign surfaces for which $\theta > 0$, where θ is the contact angle between β in contact with α and the foreign surface.

However, VOLMER[29] has shown that material held in microcavities of suitable geometry in the foreign surface can be more stable as β than as α, even in the temperature range of α stability for $0 < \theta < 90°$. The basic reason is that for $\theta < 90°$ there is a net gain in surface energy for that part of the β surface that forms on the surface of a catalyst (see equation (15)). Therefore, any geometry of the β form that maximizes the fraction a_s of the β surface forming on the catalyst surface and minimizes the fraction a_α in contact with the parent α will favour the stabilization of β. Evidently negative curvature in the catalyst surface increases a_s/a_α.

A theory for the thermal history phenomenon in phase transformations, based on the concept of retention of β in microcavities, has been derived.[27] As the temperature is increased above T_0, the size of the microcavities in which β is retained is decreased and at very high temperatures becomes of the order of molecular dimensions. After an element of material in a microcavity transforms from β to α it cannot revert to β until the temperature is lowered sufficiently for the rate of nucleation of β on a flat surface to be appreciable.

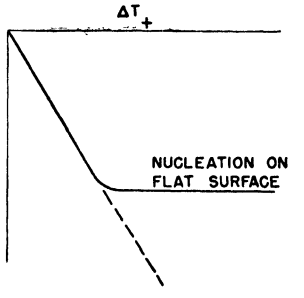


Fig. 4. Effect of superheating on the supercooling required to cause nucleation of embryos retained in cylindrical cavities[27]

Elements of β retained in microcavities nucleate the transformation from α to β. In order to serve as a nucleus the radius of the β element, and hence the microcavity in which it is retained, must equal or exceed the critical radius r_c^* for nucleation of β on a flat surface. Since r_c^* decreases with decreasing temperature it follows that the supercooling, ΔT_-, will be less the larger is the maximum size, r_m, of the microcavity in which β is retained. Since r_m decreases with increasing ΔT_+, where $\Delta T_+ = T_m - T_0$, and T_m is the maximum temperature attained in the range of α stability, it follows that ΔT_- should increase with ΔT_+. On the basis of these concepts and assuming cylindrical microcavities the undercooling should be related to the degree of superheating by the following equation[27]:

$$\tan \theta = \Delta T_-/\Delta T_+ \quad . \quad . \quad . \quad . \quad (34)$$

Fig. 4 shows ΔT_- as a function of ΔT_+ for a particular value of θ.

It is evident that equation (34) ceases to have meaning when ΔT_- exceeds the supercooling ΔT_-^* required for a perceptible rate of nucleation on a flat element of the catalyst surface. ΔT_- is a linear

function of ΔT_+ for $\Delta T_- < \Delta T_-^*$ but is independent of ΔT_- when $\Delta T_- > \Delta T_-^*$.

With suitable microcavities in its surface a substance characterized by a comparatively large value of θ may be as effective in nucleation catalysis as another substance characterized by a very small value of θ. For example, it is doubtful that substances, such as gallium, having very complex crystal structures ever crystallize from a liquid without very marked undercooling unless nucleation is effected by retained nuclei.

In heterogeneously nucleated transformations it is possible for the preferred nucleation sites to be exhausted. AVRAMI[30] (now MELVIN) recognized this possibility and developed a formal theory to describe it.

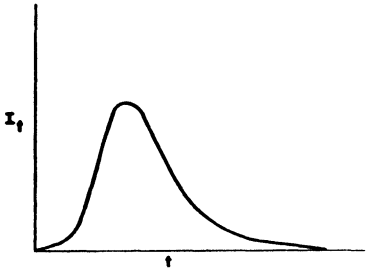


Fig. 5. Frequency of heterogeneous nucleation as a function of time showing initial build up toward steady frequency and final falling off because of exhaustion of nucleation sites[27]

Suppose there are N_0 preferred nucleation sites per volume in the system. These sites might be occupied by retained nuclei or by catalyst particles so small that the whole surface of a particular particle will be engulfed by β growing from the first β nucleus that forms on its surface. If there is equal probability that a nucleus will form at any of the preferred sites the number N per untransformed volume that have not nucleated in time, t, is:

$$N = N_0 \exp [-ft], \quad . \quad . \quad . \quad . \quad (35)$$

where f is the nucleation frequency/site. The nucleation frequency/volume is found by differentiation of (35) to be:

$$I_t = N_0 f \exp [-ft]. \quad . \quad . \quad . \quad . \quad (36)$$

AVRAMI did not consider the possibility of an initial transient and his theory requires that I_t always decrease with time. When the initial transient in I_t is taken into account a more general relation:

$$I_t = N_0 f \exp [-(ft + i^{*2}/4n^*\nu t)], \quad . \quad . \quad . \quad . \quad (37)$$

in heterogeneous nucleation is obtained that is indicated graphically in Fig. 5. I_t increases from a value of zero to a steady value and then declines as the nucleation sites are exhausted. It is evident that the entire system may be transformed to β by growth of transient β nuclei before the effects of exhaustion have become apparent.

Nucleation Phenomena in Condensed Systems

Fracture of Liquids—The breaking of liquids subjected to hydrostatic tension takes place by the nucleation and growth of bubbles. Consider the formation of a spherical cavity in the body of a liquid subjected to a

hydrostatic tensile stress, P (negative pressure). Nucleation of these cavities can be described by the theory of nucleation in condensed systems. If $P \gg P_0$, the vapour pressure of the liquid, $\Delta G_v = -P$ and $\Delta G^* = 16\pi\sigma^3/3P^2$, where σ is the liquid-vapour interfacial energy. Following FISHER[31] the frequency of bubble nucleation is:

$$I = n(kT/h) \exp [-\Delta G_A/kT] \exp [-16\pi\sigma^3/3P^2], \quad . \quad . \quad . \quad . \quad (38)$$

where n is the number of atoms per volume and ΔG_A, the free energy of moving an atom out of the liquid-vapour interface into the liquid, may be set equal to the free energy per atom for viscous flow. This formal treatment applies only to relatively fluid liquids. FISHER[31] has shown that in very viscous liquids such as glasses, fracture by the nucleation and growth of thin disc-like cavities is a much more probable occurrence than fracture by the nucleation and growth of spherical cavities.

Liquids usually prefer to pull apart at interfaces already present in the system rather than by homogeneous nucleation of cavities within the bulk of the liquid. Practically, it is very difficult to attain the negative pressure necessary to cause homogeneous nucleation of cavities before the pressure is relieved by the growth of a bubble that is nucleated heterogeneously at an existing interface. In order to minimize the likelihood of heterogeneous nucleation of cavities, foreign particles should be removed from the liquid and a container having walls that are strongly "wet" by the liquid should be selected.

TABLE 2

Comparison of Theoretical and Measured Fracture Stresses of Various Liquids. After FISHER[31] and BERNATH[32]

Liquid	*Fracture Pressure Atmospheres*	
	P_{calc}	P_{exp}
Water . .	1320	350[33]
Ethyl Alcohol .	223	39[34]
Ethyl Ether .	138	72[34]
Benzene . .	338	152[35]
Acetic Acid .	322	288[35]

The criterion for the fracture of a liquid subjected to a negative pressure is that the rate of nucleation of cavities (I) must have some perceptible value (such as 1/mole-second). In order to test the predictions of nucleation theory, negative pressures, P_{calc}, corresponding to

$I = 1$ mole^{-1} sec^{-1} calculated from equation (38) are compared with values measured experimentally, P_{exp}. In this comparison only the maximum negative pressures that have been reproduced experimentally are pertinent since it is presumed that the pressure required for homogeneous nucleation must be equal to or greater than the maximum negative pressure that is observed. The comparison of P_{calc} and P_{exp} for various fluids is shown in Table 2, on previous page.

Although the maximum negative pressure is nearly as large as the calculated pressure for acetic acid, in all other cases it is less by a factor of 2 to 6. It is probable that cavities were nucleated heterogeneously in all the latter experiments.

The heterogeneous nucleation of cavities takes place according to the general theory already described.

Crystallization of Supercooled Liquids—The supercooling of liquids was observed almost as soon as temperature began to be measured quantitatively. FAHRENHEIT (around the year 1725) probably made the first systematic study of supercooling. He discovered that water readily supercooled several degrees before beginning to crystallize and that different samples of water consistently behaved differently. Fortunately, from the standpoint of using the ice-water equilibrium temperature, as a fixed point in thermometry, ice crystals grow very rapidly so that, though different samples of water start to crystallize at widely different temperatures, the rapid release of heat associated with growing ice crystals quickly raises their temperature to very nearly the equilibrium temperature.

According to VOLMER'S[29] historical account it was realized more than one hundred years ago that the crystallization of supercooled liquids or supersaturated solutions is usually nucleated heterogeneously. These bodies are usually present accidentally and in too small an amount to permit chemical identification.

The evidence indicating that crystal nucleation in supercooled liquids is usually induced by accidental impurities has been summarized in many earlier publications[30, 27, 2]. Perhaps the most important evidence is the specificity of the behaviour of different samples of the same liquid first observed by FAHRENHEIT and the reduced nucleation frequency observed after a liquid has been passed through a fine filter.

The problem of minimizing the effect of adventitious nucleation catalysts in crystallization has been a baffling one and an apparently workable solution was not found until recently.

For more than seventy years it has been known that small metal droplets supercool several tens of degrees before solidification. The most systematic early study of the phenomenon was made by INGERSOLL and MENDENHALL.[36] They found that 50 to 100 micron diameter

droplets of metals such as gold, platinum and rhodium often supercooled over 100° when melted and cooled on a Nernst glower. There seemed to be a rough correlation between the maximum supercoolings, ΔT, and the melting temperature of the metal. For platinum the amount of supercooling was observed to be approximately 370°C. They found that ΔT was generally greater the smaller the droplet. Their results coupled with the common knowledge that large continuous masses of liquids, such as gold, that crystallize to close-packed structures rarely supercool more than 2 or 3° indicate a tremendous effect of specimen size on solidification behaviour.

Whether INGERSOLL and MENDENHALL or any of the earlier investigators appreciated the large magnitude and importance of this size effect is not entirely clear from the literature.

The recent experiments of VONNEGUT[37] on tin and of CWILONG[38] and SCHAEFER[39] on water again focused attention on the size effect in solidification. VONNEGUT studied dilatometrically the rate of solidification of a dispersion of small tin droplets, 1–10 microns in diameter, prevented from coalescing by an oxide film. In these experiments the period of crystal growth was negligible in comparison with that of nucleation. The rate of solidification was not perceptible until the dispersion had been supercooled of the order of 105 to 110°. Large continuous masses of tin have not been supercooled more than 30°C without crystallization.[49]

CWILONG[38] and SCHAEFER[39] observed that snow crystals do not form at an appreciable rate in a fog of water droplets (10–20 microns in diameter) until a "critical" temperature 39 to 40° below the normal melting point of water is reached. Large continuous masses of water have only rarely been supercooled more than 20° without crystallization.

These experiments clearly confirm the existence of a marked size effect in solidification. Thus, the frequency of nucleus formation (k) in a given specimen, is a function of the volume, v, of the specimen and the degree of supercooling, i.e. $k = f(v, \Delta T)$. We may for a given liquid define an "isokinetic" relation as follows[28]:

$$v = f(\Delta T)_{k=\text{constant}} \quad . \quad . \quad . \quad . \quad (39)$$

With tin the magnitude of the size effect is such that the value of v changes by a factor of the order of 10^{11} for a change of 80° in ΔT.

It has been demonstrated[41, 42, 43] that only the theory that partitioning the liquid isolates the effective catalyst particles in a small fraction of the droplets seems compatible with all the facts. Suppose that the liquid containing m_0 adventitious nucleation catalyst particles per volume is dispersed into m droplets per volume kept from intercommunicating by suitable barriers. If $m \gg m_0$, crystallization nuclei in most droplets either must form homogeneously or on the surface of added nucleation catalysts, e.g. a crystalline film on the droplet surface.

The probability that a particular droplet contains an accidental nucleation catalyst is $\sim m_0/m$.

The isolation theory can be tested by comparing the isokinetic relation it predicts with the isokinetic relation measured experimentally. Consider first the very simple case where all m_0 catalyst particles have the same catalytic potency per area characterized by a single value of the contact angle, θ. The isokinetic relation for droplets containing the average number of catalyst particles per volume is the curve *cd* drawn schematically in Fig. 6. Curve *ab* represents the isokinetic relation for homogeneous nucleation. For large values of $v(v = 1/m)$ $m_0 \gg m$ and the measured isokinetic relation will follow curve *cd*. In the region

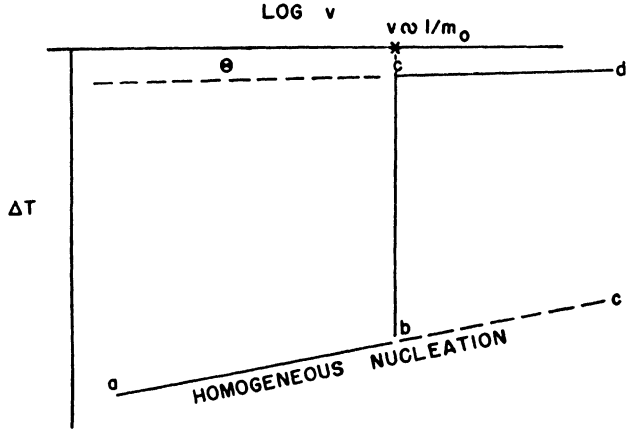


Fig. 6. Schematic isokinetic relation for solidification of liquid droplets that may contain only one variety of nucleation catalyst

$v \sim 1/m_0$ the crystallization behaviour will be very erratic because there is an average of only one catalyst particle per droplet. Since these particles will be distributed statistically some droplets will not contain any, and others will contain several times the average number of catalyst particles. Therefore, the measured ΔT will fall on the curve *ab* for some droplets and above or below the curve *cd* for others. For $1/m_0 \gg v$, the average measured ΔT will follow the curve *ab*. We have made an idealized representation (see solid curve Fig. 6) of the predicted experimental isokinetic relation. Note that, in agreement with nucleation theory, the absolute magnitude of the slope of *ab* is much greater than that of *cd*.

The experimental evidence[44] definitely indicates that there is rarely, if ever, so large a discontinuity in the ΔT-log v relation as is indicated in Fig. 6. Rather the data suggest an almost continuous decrease of ΔT with decreasing log v as shown in Fig. 7.

To account for this almost continuous variation of ΔT with log v

two additional postulates are necessary—first, the m_0 catalyst particles are characterized by a variety of θ values rather than a single one. That is, some catalysts will be more effective than others in promoting solidification. Second, the concentration of the more effective nucleation catalysts is likely to be smaller than that of catalysts that are less effective. From these postulates LEVINE[45] has derived a ΔT-log v relation that successfully describes the experimental data of DORSCH and HACKER[44] on the freezing of water droplets. LEVINE'S theory is formally identical with WEIBULL'S[46] theory for the (see also FISHER and HOLLOMON[47]) size effect in fracture. The basic postulates, due to GRIFFITH,[48] underlying WEIBULL'S theory are that fracture is initiated

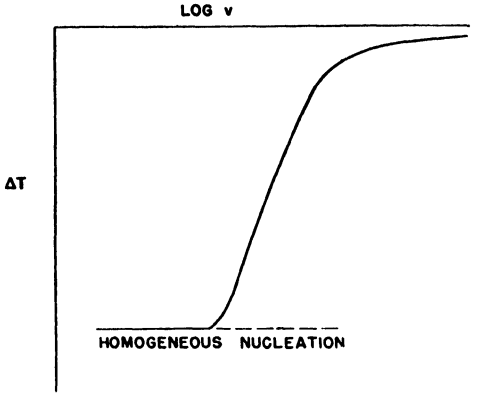


Fig. 7. General isokinetic relation for a liquid containing a wide variety of accidental nucleation catalysts

by structural defects and, that the density of these defects per area is greater the lesser their ability to initiate fracture. These postulates are analogous with the underlying postulates of the isolation theory for the size effect in solidification.

It is reasonable to suppose that a liquid may contain a wide variety of nucleation catalysts. This variety may be due to chemical differences and/or to differences in the catalytic potency of substances due to heterogeneous surface structure.[28]

The concentration of catalyst particles will decrease as their effectiveness increases only if a few substances can be potent nucleation catalysts for a given liquid-crystal transition but a wide variety of substances can be catalytic to some degree. There is experimental evidence that favours this postulate.

Probably the most quantitative support for the isolation theory and for the theories of homogeneous and heterogeneous nucleation applied to solidification are the results of TURNBULL[28] on the kinetics of solidification of dispersions of mercury droplets coated with various films. In

these experiments the temperature range of rapid solidification was first measured in continuous cooling experiments. This temperature range is relatively insensitive to cooling rate because of the large temperature coefficient of the rate of nucleation. Crystal nucleation frequencies were deduced from isothermal rates of solidifications measured for several of the dispersions. From the isothermal data and the measured droplet size distribution it was determined whether the

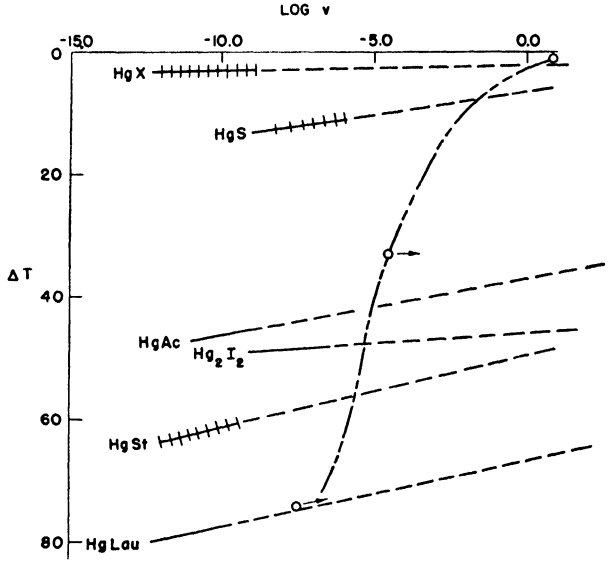


Fig. 8. Isokinetic relation for solidification of mercury droplets coated with various films. After TURNBULL[28]

nucleation frequency was proportional to the volume or to surface area of the droplet.

Fig. 8 shows isokinetic relations for the various dispersions corresponding to a nucleation frequency 10^{-4} sec^{-1}. The solid lines for droplets coated with mercury laurate (HgLau), mercury acetate (HgAc), and mercurous iodide (Hg_2I_2) films are precise. Approximate relations for droplets coated with mercury sulphide (HgS), mercury stearate (HgSt), and HgX were deduced from continuous cooling experiments and some isothermal measurements of uncertain interpretation on HgSt and HgX dispersions. Chemical evidence suggested that the HgX film was a mercury oxide. Also drawn in the figure is the assumed course of the isokinetic relation for the bulk mercury sample containing only the accidental catalysts. There is a large uncertainty in this curve, since it was inferred from only 3 points and the two lowest of them give only the upper limit of ΔT for the corresponding v value.

These results confirm two of the most important predictions of the isolation theory. First, that there is no single isokinetic relation that describes the size effect in solidification, rather the isokinetic behaviour is clearly dependent upon the nature of the film on the droplet surface. The second is that the slope of the isokinetic relation for a particular catalyst is very small. For example, the isokinetic relation for droplets coated with HgLau gives $\Delta T = 70°$ when extrapolated to $v = 1$ cc. Since 1 cc masses of mercury almost always start to crystallize for $\Delta T < 4°$, it is concluded that they must contain relatively effective catalyst particles that become isolated from most of the mass when it is dispersed.

It was established that the nucleation frequency in droplets coated with HgLau is proportional to droplet volume and that the kinetic coefficient, K_v, that describes the data is in reasonable agreement with K_v calculated from equation (24) predicted by the theory of homogeneous nucleation. Comparison of the experimental and calculated K_v is indicated in Table 3. These results are compatible with the interpretation that nuclei formed homogeneously in the HgLau coated droplets. The temperature coefficient of the nucleation frequency gives the value $\sigma = 31{\cdot}2$ ergs/cm^2 (see table) when it is assumed that the nuclei are spherical.

The disagreement (a factor of 10^7) between the experimental and calculated values of K_v is outside the experimental error. A possible reason for this disagreement is that σ and ΔS_v may not be independent of temperature as assumed in making the calculation. It was shown[28] that the disagreement can be explained on the assumption of small and apparently reasonable magnitudes for the temperature coefficients of σ and ΔS_v. It is also possible to account for the disagreement on the concept that a crystal nucleus may derive from a structure that is somewhat different from the most stable macroscopic structure, characterized by the normal melting temperature, T_m. This concept is plausible because the structure that minimizes σ is not necessarily the same structure that minimizes the Gibbs free energy of the crystal. For example, let the nucleus be strained an amount ε relative to the stable macroscopic structure α. Then σ is a function of ε, $\sigma(\varepsilon)$, and equation (1) for the free energy of forming the crystal nucleus is modified to:

$$\Delta G = 4\pi r^2\sigma(\varepsilon) + (4/3)(\pi r^3)(\Delta G_v + c\varepsilon^2) \quad . \; . \; . \; . \quad (40)$$

where c is the appropriate elastic coefficient. The macroscopic structure, β, from which the nucleus derives will be characterized by a melting temperature, T_m', which is lower than T_m and is found by solving the equation $\Delta G_v = \Delta S_v \Delta T = c\varepsilon^2$. The free energy of forming the nucleus is:

$$\Delta G^* = f(r^*, \varepsilon^*) \quad . \; . \; . \; . \quad (41)$$

where r^* and ε^* are found by solving simultaneously the equations:

$$[\partial(\Delta G)/\partial r]_\varepsilon = 0 \quad . \quad . \quad . \quad . \quad (42)$$

$$[\partial(\Delta G)/\partial \varepsilon]_r = 0 \quad . \quad . \quad . \quad . \quad (43)$$

Let the Gibbs free energy for formation of β be $\Delta G_v'$; it then follows that:

$$\Delta G_v' = \Delta S_v(T_m' - T) = \Delta S_v \Delta T''. \quad . \quad . \quad . \quad . \quad (44)$$

$\sigma(\varepsilon)$ is not known; however, we may calculate, from the measured temperature coefficient of I and substituting (44) into (24), the value of T_m' that will bring the experimental and theoretical values of K_v into agreement. The value of T_m' thus calculated is about 11–12°C below T_m. Displacements of the order of 1–2 per cent in the lattice parameter of α gives a structure that should melt 11–12° below T_m.

The nucleation frequency, k_D, of droplets coated with HgAc was proportional to the area rather than to the volume of the droplet. This result is in good agreement with the notion that the formation of nuclei in these droplets was catalyzed by the surface film. The temperature dependence of I_s is satisfactorily described by equation (32) with the values of θ and K_s given in Table 3. The experimental value

TABLE 3

Summary of Data on Isothermal Solidification of Small Droplets

Substance	*Film*	σ *ergs/cm²*	θ	*Log* K_s		*Log* K_v		$T_m - T_m'$	ΔT_m
				Expt'l	*Theor.*	*Expt'l*	*Theor.*		
Mercury .	HgLau	31·2				42·1	35·1	12	79
Mercury .	HgAc	31·2	72°	26·8	27·3			0	46
Mercury .	HgSt	31·2	87°	27·4	27·3			0	60
Mercury .	Hg_2I_2					69·6	35·5	31	48
Tin. .	Tin oxide			30·0	28·0			6	120

of K_s is in excellent agreement with the value predicted by the theory so T_m' is displaced from T_m by a negligible amount.

Other results of the investigation of mercury solidification are summarized in the table. Solidification isotherms for droplets coated with Hg_2I_2 were described on the basis that nucleation was catalyzed by particles, characterized by a single θ, suspended in the mercury and that either the mercury nucleus derives from a structure melting about

30° below T_m or that the rate controlling step is not nucleation of mercury crystals but the nucleation of a phase, possibly of Hg_2I_2, that catalyzes the nucleation of mercury.

Results for solidification of the Hg(St) and HgX coated droplets could not be described by a single value of I or I_s at constant temperature but could be described by a multiplicity (2 for HgSt; a large number for HgX) of I_s values. POUND and LA MER's[49] isotherms for the solidification of oxide-coated tin droplets can be described by a similar multiplicity in I_s. Despite this multiplicity in I_s the estimated K_s values for HgSt-coated mercury and oxide-coated tin droplets are in excellent agreement with the theoretical values as indicated in the table. As a corollary T_m' is displaced from T_m by a very small amount in comparison with the supercooling. Because of the great multiplicity in I_s it was not possible to estimate K_s for HgX-coated droplets.

In summary, the data on the solidification of supercooled mercury and tin droplets are compatible, to well within the experimental uncertainty, with the theory of homogeneous and heterogeneous nucleation in condensed systems.

A valuable feature of the theory of homogeneous nucleation in condensed systems is that K_v is calculable from the characteristic constants for the system.[26] It is thus possible on the basis of the theory to calculate the unknown parameter σ from an I value for a single temperature. We have confidence in the validity of this procedure because of the fair agreement between the theoretical and experimental K_v for mercury. TURNBULL and CECH[50] have used a microscopic technique to study the solidification behaviour of small droplets (10–50 micron diameter) of a variety of metals. The maximum supercooling ΔT_m before solidification that they obtained for the various metals is given in Table 4. Also listed in the table are ΔT_m values for other substances obtained by other investigators using a variety of techniques.* It is supposed that the magnitude of the supercooling $\Delta T'$ corresponding to a perceptible rate of homogeneous nucleation is equal to or greater than ΔT_m. Assuming $\Delta T' = \Delta T_m$[43] reasonable values of I were assumed and σ for the various substances was calculated from equation (24). These calculations have been brought up to date and the resulting σ values are listed in Table 4.

* Recently there has been some dispute on the correct ΔT_m for water. SCHAEFER[39] and CWILONG[38] agree that rapid formation of snow crystals begins when a supercooled fog reaches a temperature of − 39°. However, CWILONG,[38] SANDER and DAMKÖHLER,[22] and POUND and MADONNA[23] all report that a fog formed at − 50 to − 62°C by rapid expansion of water vapour may consist of supercooled water droplets. Their conclusion that the droplets were liquid was based on their failure to observe reflections from crystal facets that are expected from snow crystals. SCHAEFER[54] has shown, however, that a dispersion of small ice spheres can exist without any perceptible reflections. We have tentatively accepted SCHAEFER's value of ΔT_m.

TABLE 4

Summary of Liquid-crystal Interfacial Energies Calculated from Supercooling Data

Substance	*Reference*	$\Delta T_m'$	σ (ergs/cm^2)	σ_g/T_m cal/deg
Mercury	(28)	77	28·1*	1·52
Gallium	(51)	76	56·0	1·91
Tin	(49)	118	59·0	1·59
Bismuth	(28)	90	54·4	1·52
Lead	(28)	80	33·3	0·80
Antimony	(50)	135	101·0	1·59
Aluminium	(50)	130†	93·0	1·00
Germanium	(50)	227	181·0	1·71
Silver	(50)	227	126·0	1·00
Gold	(50)	230	132·0	0·99
Copper	(50)	236	177·0	1·01
Manganese	(50)	308	206·0	1·11
Nickel	(50)	319	255·0	1·08
Cobalt	(50)	330	234·0	1·02
Iron	(50)	295	204·0	0·88
Palladium	(50)	332	209·0	1·01
Platinum	(36)	370	240·0	1·05
Water	(52)	39	32·1	1·69
Benzene	(52)	48	17·2	2·58

* Calculated on basis outlined in reference 43.

† Recently FALKENHAGEN and HOFMANN[53] have, using very large cooling rates, supercooled aluminium about 195°. This value of ΔT_m gives a $\sigma_g/\Delta H_f \sim 0{\cdot}48$ and in better agreement with the other values of $\sigma_g/\Delta H_f$ for face-centred cubic metals.

The gram-atomic interfacial energy, σ_g, between liquid and crystal phases is defined as the free energy of a crystal-liquid interface containing one Avogadro number N of atoms. Assuming that the interfacial energy is localized in a layer one atom thick it follows[43]:

$$\sigma_g = N^{1/3} V^{2/3} \sigma, \qquad \ldots \quad (45)$$

where V is the gram-atomic volume of the crystalline phase. TURNBULL[43] showed that σ_g, calculated from nucleation rates, is roughly proportional to the gram-atomic heat of fusion ΔH_f. $\sigma_g/\Delta H_f$ (see Table 4) ranges from 0·32 to 0·61. σ_g for the various substances is plotted against ΔH_f in Fig. 9. Points for most of the metals fall fairly well on a common line of slope 0·46.

Large liquid masses of metals that crystallize to close-packed structures such as gold, copper or lead rarely supercool more than 5° before

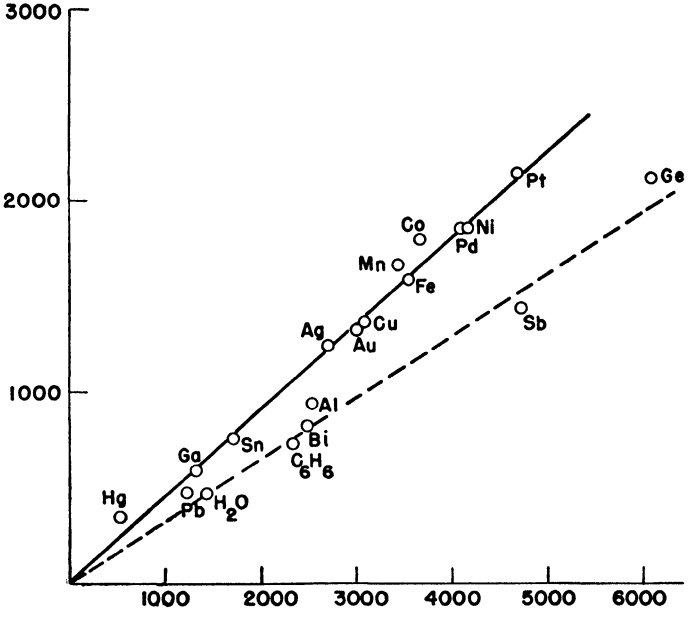


Fig. 9. Gram-atomic interfacial energy, σ_g, as a function of the heat of fusion, ΔH_f. Revision of TURNBULL[43]

beginning to crystallize. For example, DANILOV and NEUMARK, after subjecting lead to what they believed to be a satisfactory purification procedure, found that it started to crystallize at 3° supercooling. On the other hand, large liquid masses of metals that crystallize to more complex non-close-packed structures, such as tin, bismuth or gallium, often can be easily supercooled 25° or more before starting to crystallize. These results led to the concept advanced by DANILOV and NEUMARK[40] and others that the supercooling corresponding to a perceptible nucleation rate, ΔT_m, is strongly dependent on the structure of the forming crystal and for liquids crystallizing to close-packed structures $\Delta T_m \lesssim 5°$. DANILOV and NEUMARK, for example, believed that $\Delta T_m \sim 3°$ for lead. This interpretation was considered plausible on the basis of the widely held view that the arrangement of atoms in most liquid metals approximates a close-packed structure. Thus the transition liquid (semi-close

packed) → crystal (close packed) could presumably take place with a minimum of atomic displacements. These views of the liquid → crystal (close packed) transformation apparently were developed in ignorance of the results of INGERSOLL and MENDENHALL[36] and other earlier investigations of the supercooling of gold, etc. Certainly these results coupled with results summarized in Table 4 establish that $\Delta T_m/T_m$ ratios are not widely different for metals whether or not they crystallize to a close-packed structure. Also BARDENHEUR and BLECKMANN[55] have been able, under certain conditions, to supercool a 150 g mass of liquid iron 258° before the beginning of crystallization.

Although the nature of the forming structure has not a marked effect upon $\Delta T_m/T_m$ in solidification there is a rough correlation between the magnitude of the ratio σ_g/T_m (see Table 4) and the structure complexity.[43] This ratio ranges from a value of 1·9 for gallium to about 1·0 for metals that crystallize to close packed structures.

A substantial proportion of accidental catalyst particles are likely to be cubic or hexagonal in structure and may, therefore, be capable to a greater or lesser degree of catalyzing the formation of cubic or hexagonal crystals.[43] On the other hand, very few substances have a structure very similar to that of gallium and, therefore, capable of effectively catalyzing the formation of gallium crystals. In a random selection of catalyst particles that might be present accidentally in a liquid metal it is probable that a large proportion would have cubic or hexagonal structures but few, if any, would have a structure similar to that of gallium. Therefore, a liquid that crystallizes to a complex structure will probably supercool much more relative to its melting point than a liquid that crystallizes to a cubic or hexagonal structure.

Although WEBSTER[56] and others found a pronounced thermal history effect in the solidification of large masses of bismuth, no measurable thermal history dependence of supercooling in the solidification of small bismuth droplets has been found.[51] These results show clearly that the thermal history effect depends for its existence upon the presence of crystal nucleation catalyst particles in the liquid.

Crystal → Liquid Transitions—The frequency of homogeneous nucleation of liquid droplets in the body of a crystal should be given by an equation identical in form with equation (24). This theory predicts that the homogeneous formation of liquid nuclei should not become perceptible until the superheating becomes almost as great as the supercooling required for a perceptible rate of homogeneous crystal nucleation in the liquid.

Experience indicates that crystals almost invariably begin to melt at temperatures not perceptibly different from the equilibrium temperature. Also it is known that most crystals are completely wet by

liquid phase of the same composition at the melting temperature, hence:

$$\sigma_{CV} \geqslant \sigma_{LV} + \sigma_{LC} \quad \ldots \quad (46)$$

Therefore, there can be no net increase in surface energy when a liquid forms on a free solid surface of the same composition and ΔG_s^* for forming liquid nuclei on this surface is zero. For this reason crystals will not superheat when they have any part of their surface exposed to a gas phase.

Thus it is only possible to superheat solids by preventing in some way the exposure to vapour of the solid surface at the equilibrium temperature. KHAIKIN and BENET[57] have succeeded in superheating the interior of a tin single crystal 1 to 2°. In their experiment a large electric current was passed through the crystal. The energy was carried away from the free surface of the crystal so that a temperature gradient of 1–2° existed between the crystal axis and the surface. Although polycrystals and defective single crystals invariably started to melt internally sound single crystals always started to melt from the free surface, indicating that the material at the crystal axis was superheated 1–2°. This amount of superheating is only a very small fraction of what is possible theoretically ($\sim$ 100° for tin).

An experimental approach to the problem that has not yet been successfully worked out is to completely coat small crystals with a substance that strongly catalyzes their formation in supercooled liquids. For example, it should be possible to substantially superheat tiny ice crystals covered with a coherent coating of silver iodide.

STRUCTURE AND COMPOSITION

Among the common types of transformation involving a change in structure and composition are: (1) condensation of liquid from a gaseous solution; (2) the reverse process of formation of gas bubbles in liquid solution; (3) precipitation of crystalline phases from liquid solutions; (4) precipitation of crystalline phases from solid solutions. We shall ignore the first two types and treat the last two, which are of most interest in metal systems.

In the nucleation of crystals from solid or liquid solutions the important variables are structure, the strains due to the disregistry of the structures, composition and temperature. In our first approximation we will neglect effects due to strain.

Thermodynamics of Precipitation

The Gibbs free energy, G, of a solution is a function of composition x (mole fraction of one component), temperature T and pressure P. Fig. 10 (a) is a schematic representation of G as a function of x (mole

fraction of B) at constant T and P. Here, G is a continuous function of x and applies to a series of solid solutions of one crystal symmetry (e.g. face-centred cubic) or liquid solutions. Fig. 11 shows two $G = f(x)$ relations, each characteristic of phases α and β (either two solid solutions or a liquid and a solid solution). The most convenient criterion of equilibrium in these systems is that G be a minimum at constant T and P. Applying this criterion to the system represented in Fig. 10 (a) it

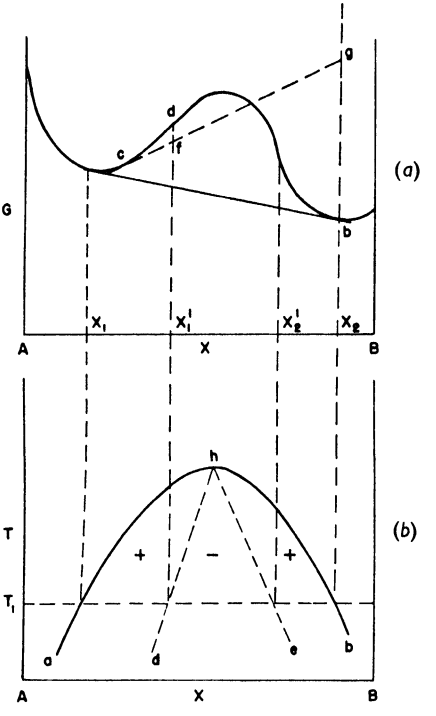


Fig. 10. (a) Free energy of a solid solution as a function of composition (b) Temperature composition equilibrium diagram for two components partially miscible in the solid state

follows that the composition of the solution ranges continuously for $x \lesssim x_1$ and for $x_2 \lesssim x$. For all $x_1 < x < x_2$, the free energy of the system is minimized by the coexistence of two phases of composition x_1 and x_2 [a and b are points of tangency of the line ab to $G = f(x)$]. Similarly it is found that the free energy of the system represented in Fig. 11 is a minimum for $x_\alpha < x < x_\beta$ when two phases of composition x_α and x_β coexist.

In the conventional representation of phase equilibrium the phase coexistence relations are projected on a T-x plane for constant P as shown in Fig. 10 (b). Also plotted as dotted lines (hd and he) in Fig. 10 (b) are the spinodal relations. The points on these curves fulfil the

condition $[\partial^2 G/\partial x^2]_{T,\,P} = 0$. In areas *hda* and *heb* $[\partial^2 G/\partial x^2]_{T,\,P} > 0$ and in area *hde* $[\partial^2 G/\partial x^2]_{T,\,P} < 0$.

Now consider fluctuations in composition in metastable solid solutions for $x_1 < x < x_2$. A cluster of atoms formed by a fluctuation may, for a given mean composition of the supersaturated solution x, be characterized by the number of atoms, i, that it contains and Δx, where $\Delta x = x' - x$, and x' = the composition of the cluster. It is assumed, following BECKER[25] and BORELIUS,[58] that the composition of the matrix immediately surrounding the cluster is x. In our discussion we shall consider only fluctuations favourable to the nucleation of x_2.

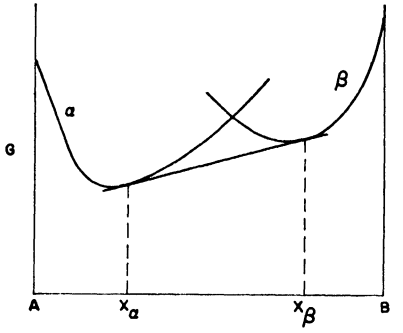


Fig. 11. Free energy—composition dependence at constant temperature of two solid solution structures, α and β

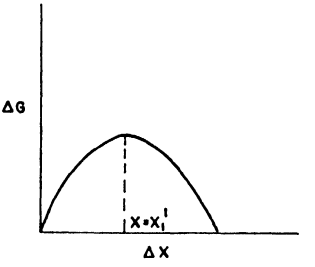


Fig. 12. Free energy of a cluster as a function of the fluctuation in composition at constant size. After BORELIUS[58]

Let ΔG be the free energy increase per gram atom, neglecting interfacial energy, for the formation of a cluster. It follows from thermodynamic principles that

$$\Delta G = G' - (G + \Delta x \partial G/\partial x) \quad . \quad . \quad . \quad . \quad (47)$$

where G' is the free energy per gram atom of solution at mole fraction x', G is the free energy per gram atom of solution at mole fraction x, $\partial G/\partial x$ is the slope of $G = f(x)$ relation at x. ΔG is the difference in G co-ordinates at a given x' of $G = f(x)$ and the tangent to $G = f(x)$ at point x. For the matrix composition corresponding to point c, ΔG is proportional to the length of the line df at $x' = x_1'$ and to minus the length of the line bg at $x' = x_2$.

Fig. 12 shows $\Delta G = f(\Delta x)$ for a given matrix composition x. The function goes through a maximum at the spinodal composition x_1'. For $x' > x_1'$ at constant i the free energy of a cluster is decreased by fluctuations increasing Δx but for $x' < x_1'$ the reverse is true.

Let ΔG_2 be the free energy per gram atom for forming clusters with $x' = x_2$, neglecting surface energy again. ΔG_2 as a function of matrix

composition x is represented schematically in Fig. 13. $\Delta G_2 = f(x)$ goes through a minimum at the spinodal composition $x = x_1'$ and a maximum at spinodal composition $x = x_2'$.

Nucleation of a New Phase

ΔG_i, the increase in free energy due to the formation of a cluster containing i atoms, is a function of T, x and Δx. For a given value of T and x we may write

$$\Delta G_i = a\sigma i^{2/3} + \Delta G(i/N) \quad . \quad . \quad . \quad . \quad (48)$$

where σ, the interfacial energy per area between the cluster and matrix, is a function of x and Δx, a is a shape factor, ΔG is given by equation

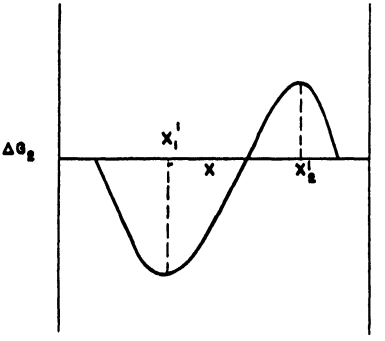


Fig. 13. Free energy per gram atom of a cluster having the equilibrium composition of the precipitate from a solid solution

(47), and N is Avogadro's number. The number, n_i, of clusters per volume in equilibrium with the matrix is

$$n_i = n_0 \exp [- \Delta G_i/kT], \quad . \quad . \quad . \quad . \quad (49)$$

where n_0 is a slowly varying function of Δx.

For a given value of x, ΔG_i can be plotted as a function of i and Δx. In order to become a nucleus for the growth of phase of composition x_2, a cluster must have attained a critical size and composition (critical cluster) such that further increase of either i or Δx will decrease ΔG_i. From the surface $\Delta G_i = f(i, \Delta x)$ the minimum, $\Delta G_i = \Delta G_i^*$, for the formation of a critical cluster can be found.

The location of the saddle point can be found formally from the following equation:

$$\Delta G_i = ai^{2/3} f(\Delta x) + (i/N) F(\Delta x) \quad . \quad . \quad . \quad . \quad (50)$$

obtained by substituting $\sigma = f(\Delta x)$ and $\Delta G = F(\Delta x)$ in equation (48), where $f(\Delta x)$ and $F(\Delta x)$ are functions of temperature and matrix

composition. By setting $[\partial(\Delta G_i)/\partial i]_{\Delta x} = 0$, we find the height of the surface ridge ΔG_i° to be:

$$\Delta G_i^\circ = A[f(\Delta x)]^3/[F(\Delta x)]^2 \quad \ldots . \quad (51)$$

where $A = 4a^3N^2/27$. The value of $\Delta x = \Delta x^*$ at the saddle point is found by solving the equation

$$3F(\Delta x)f'(\Delta x) - 2f(\Delta x)F'(\Delta x) = 0, \quad \ldots . \quad (52)$$

obtained by equating $d(\Delta G_i^\circ)/d(\Delta x)$ to zero. Substitution of Δx^* into (51) gives ΔG_i^*. The practical solution for the saddle point position from equations (51) and (52) is complicated by our lack of knowledge of the interfacial energy and its dependence upon composition. A simpler solution is obtained if, following BECKER,[25] it is assumed that the position of the saddle point is always at x_2. Then $\Delta G = \Delta G_2$ and σ is a function of x only, $\sigma = f(x)$ so that

$$\Delta G_i^* = A\sigma^3/(\Delta G_2)^2. \quad \ldots . \quad (53)$$

By a nearest neighbour approximation BECKER estimates $\sigma = (E/v^{2/3})(x - x_2)^2$, where v = volume per atom of precipitate, $E = E_{AB} - 1/2(E_{AA} + E_{BB})$ and E_{AA}, etc, are the pairwise interactions of the atoms A and B making up the solution. HOBSTETTER[59] allowed the composition of the nucleus to vary and derived a formal expression for the rate of nucleation.

From the theory of nucleation in condensed systems it follows that the steady state frequency of homogeneous nucleation of precipitate in a supersaturated solution is (see equation (24)):

$$I = (x - x_1)n^*(a\sigma/9\pi kT)^{1/2}n(kT/h) \exp[-(\Delta G_i^* + \Delta G_A)/kT], \quad \ldots . \quad (54)$$

where ΔG_i^* is given by equation (51) or (53), ΔG_A is the free energy for diffusion of component B in the solid solution and the other terms have the same significance as in equation (24).

We have assumed that σ is isotropic or that the critical clusters are bounded by crystallographically similar planes. It is easy to refine the treatment to take into account the nucleation of clusters exhibiting simple departures from spherical or cubical forms. For example, suppose that, due to the anisotropy of the interfacial energy, the critical clusters are disc-like. Let the radius of the disc be r and its thickness t, then:

$$\Delta G_i = 2\pi rt\sigma_1 + 2\pi r^2\sigma_2 + (\pi r^2 t/V)\Delta G_2 \quad \ldots . \quad (55)$$

where V is the volume per gram atom of precipitate and σ_1 and σ_2 are the interfacial energies per area of the disc edge and face

respectively. It follows that the radius r^* and thickness t^* of the critical cluster are:

$$r^* = -2\sigma_1 V/\Delta G_2 \quad \ldots\ldots (56a)$$

$$t^* = -4\sigma_2 V/\Delta G_2 \quad \ldots\ldots (56b)$$

whence it follows:

$$\Delta G_i^* = 16\pi\sigma_1^2\sigma_2 V^2/(\Delta G_2)^2 \quad \ldots\ldots (57)$$

In discussing precipitation from solid solution, KONOBEYEVSKI[60] and BORELIUS[58] have ignored the interfacial energy between the cluster and the matrix and set $\Delta G_i = (i/N)\Delta G$. Even though not implicitly stated these theories, if valid, must apply whether α and β are liquid or solid phases. Recently this theory has been further developed and advanced by BORELIUS. According to BORELIUS' theory, ΔG_i^* must lie along the ridge of the $\Delta G_i = f(i, \Delta x)$ surface corresponding to the spinodal composition x_1'. However, the height of this ridge is a minimum when $i = 1$. A literal application of the KONOBEYEVSKI–BORELIUS hypothesis, therefore, leads to the prediction that ΔG_i^* is virtually zero for any T and x. In order to obtain a partial description of the experimental data BORELIUS finds it necessary to postulate that the critical cluster crosses the ridge at $\Delta G_i^* = (i^*/N)(\Delta G)_{x_1'}$, where i^* is of the order of 100 to 200 atoms. It is presumed that a fluctuation must include 100 to 200 atoms in order to be significant but there is no theory, based on the KONOBEYEVSKI–BORELIUS postulate, that permits any prediction of the magnitude of i^*.

BORELIUS has analyzed certain data on rates of precipitation and shown that there appear to be certain singularities associated with precipitation from supersaturated solutions initially having the spinodal composition. When $\log \tau$, where τ is the time required for half-completion of precipitation, is plotted against $1/T$ the curves for various values of x, the *initial* matrix composition, are parallel at relatively low temperatures. That is, the slopes of the curves are a function of T only. It is found that the temperature T' at which the slope becomes dependent upon x as well as T corresponds to the temperature at which $x = x_1'$, the spinodal composition. BORELIUS writes for τ,

$$\tau = f_1(\Delta G_i^*) f_2(x) f_3(T) \quad \ldots\ldots (58)$$

where $f_2(x)$ and $f_3(T)$ are factors that describe the rate of growth of the precipitate. Since $\Delta G_i^* = f(x, T)$ and at $T \gtrless T'$, $\Delta G_i^* = 0$, it follows from (58) that $d\tau/dT$ varies with x for $T > T'$ and is independent of x for $T < T'$. This agreement with experience constitutes the main evidence for the KONOBEYEVSKI–BORELIUS theory. There are, however, certain facts that weaken the significance of the agreement. BORELIUS' development is intended to describe the rate of nucleation at some specified matrix composition. However, in testing his theory BORELIUS

assumed that $1/\tau$ is proportional to the nucleation frequency at zero time. Since τ must be a resultant of the rate of nucleation and growth of precipitate, neither of which were measured independently in the experiments cited, the validity of this assumption is doubtful. It should also be realized that in considering the precipitation from solid solutions strain energy induced by the transformation must be taken into account.

There are some other important characteristics of spinodal solutions that might cause singularities in their behaviour. Most noteworthy is that the driving free energy, ΔG_2, for the nucleation and growth of the precipitate of equilibrium composition, is a maximum (see Fig. 13) at the spinodal composition. Therefore, it may be possible to explain a singularity in the kinetic behaviour at the spinodal temperature (for a given x) in terms of rates of precipitate growth or nucleation of clusters of composition x_2 without recourse to the KONOBEYEVSKI–BORELIUS theory.

Comparison of Theory and Experience

In order to compare the theory with observation several conditions must be fulfilled. It must be assured that the nucleation of the new phases occurs homogeneously, and the strain energy induced by the transformation is negligible. The frequency of nucleation must be measured directly. Even more importantly for a quantitative comparison, the dependence of free energy, ΔG, and the interfacial free energy, σ, on the composition must be known. The former dependence can be measured by determining the energetics of the phases when both stable and metastable, or by studying the energetics of the stable phases and obtaining the difference of free energy by extrapolation. In principle, the dependence of the interfacial energy upon composition can be measured, but unfortunately, experimental difficulties have apparently mitigated against its successful determination in any case.

While no direct measurements of the frequency of nucleation have been made for the formation of a metallic solid solution from either a liquid or a solid solution, the temperature at which the nucleation rate becomes sensible has been determined for solidification in several binary systems.[61, 62] As indicated previously, since the rate of nucleation from fluid liquids is so temperature dependent, a nearly critical temperature of nucleation of the solid will be observed. In measuring the critical temperature for two series of liquid solutions the small particle technique already described was applied. Results were obtained in the case of the Cu-Ni system[61] in which there is complete miscibility in both the liquid and the solid states, and for the Pb-Sn system[62] in which there is miscibility in the liquid but immiscibility in the solid state. The latter is a simple eutectic system. The variation of the temperature

at which the nucleation frequency became sensible with composition are given for both these systems in Figs. 14 and 15. In the case of the Cu-Ni system the critical temperature of nucleation nearly parallels both liquidus temperatures and is continuous from the copper to the nickel side of the phase diagram. The maximum supercooling is of the order of 20 per cent the absolute melting temperature and appears to occur by essentially the same process in the alloys as in the pure elements copper and nickel. In the case of the Pb-Sn system the

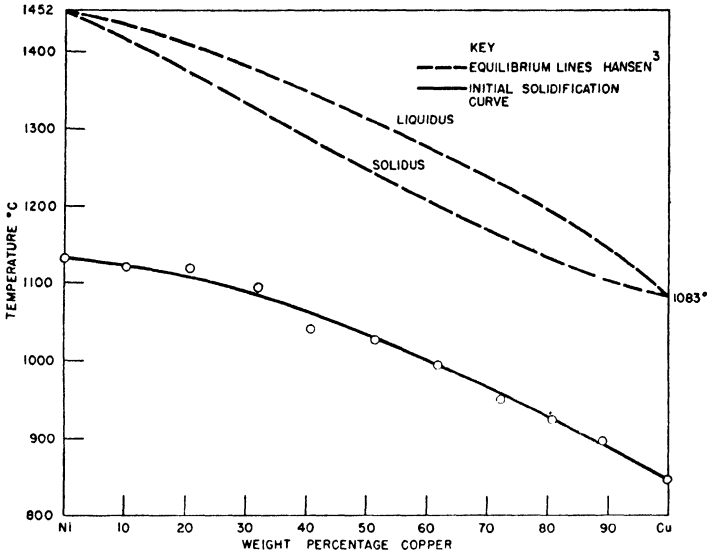


Fig. 14. Solidification temperatures of Cu-Ni alloys as a function of composition. After CECH and TURNBULL[61]

critical temperature of nucleation increases with the first small addition of tin to the lead as well as with the first small addition of lead to tin. Thus, for all compositions the magnitude of the supercooling is less than for the pure elements. These results indicate that the interfacial energy between the pure solid and liquid is decreased slightly by the first addition of both tin and lead, otherwise the results are very similar to those obtained with pure metals. However, since as already indicated, the dependence of the interfacial free energy on composition is not known for either of these systems, a quantitative comparison of these experiments with theory is impossible. Neither can the interfacial free energy be obtained from the nucleation results as was the case for the solidification of pure metals. However, the results are compatible with the general nucleation theory.

In considering precipitation from solid solution the strain energy induced by the transformation generally must be taken into account, as

indicated in the next section. Although there are no direct measurements of the frequency of nucleation during precipitation, many measurements have been made of the overall rate of the process, R. This overall rate is a composite of the rates of nucleation and growth of the precipitate. From R, it is possible, with the aid of various assumptions, to deduce something about the rate of nucleation. We have already noted that BORELIUS assumed that the value of R when precipitation is half-completed is proportional to I when precipitation

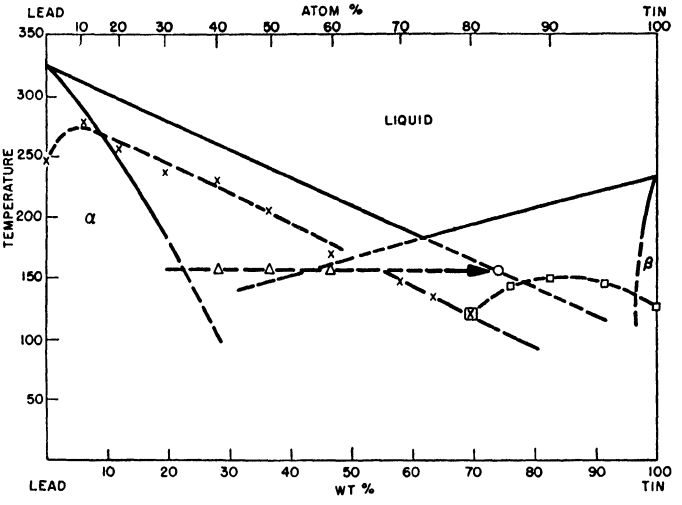


Fig. 15. Initial (X) and final (Δ) solidification temperatures of lead rich lead-tin alloy droplets as a function of their composition. Also solidification temperatures ($\square$) of tin rich lead-tin alloys as a function of composition. After HOLLOMON and TURNBULL[62]

has just begun. With the same assumption HARDY[63] showed that the data of BORELIUS *et al*[64] on precipitation of tin from lead rich solutions can be described qualitatively by the BECKER theory of nucleation. BECKER made similar assumptions when comparing the results of JOHANSSON and HAGSTEN[65] on precipitation from an Au-Pt alloy with the predictions of his theory.

A procedure far more satisfactory than using a single datum on the R vs t curve is to determine the functional relation between R and $t[R = f(t)]$ and find what assumptions about I are compatible with it. An example of this procedure is the analysis by WERT and ZENER[66] of WERT'S[67] data on $R = f(t)$ at various temperatures for the precipitation of iron carbide from α-iron. Making reasonable assumptions about the growth of the precipitate and taking impingement into account they were able to describe the $R = f(t)$ relations on the assumption that all nuclei were effectively in existence at zero time indicating

that nucleation was heterogeneous. There are in most solids non-equilibrated structural defects. These defects as well as accidental inclusions might serve as preferred nucleation sites.

Structure and Strain

Incoherent Nucleation

A change in the specific volume of the system accompanies most phase transformations. When transformation occurs within extensive gas or liquid phases of high fluidity the volume changes are accommodated by flow of the extensive phase. However, crystals and glasses can be subjected to a considerable stress without flowing at appreciable rates. It follows, therefore, that the volume change accompanying the formation of an element of a new phase in a crystalline or glassy medium may have to be accommodated by strains set up in the medium and possibly the new phase. Therefore, it is necessary to consider the strain energy as part of the free energy of the formation of an embryo. This transformation strain energy will be proportional to the number of atoms contained in the embryo and equation (2) is modified to:

$$\Delta G = a\sigma i^{2/3} + bi(\Delta G_v + E) \quad . \quad . \quad . \quad . \quad (59)$$

where E is the increase in strain energy of the system accompanying the formation of a unit volume of an embryo. Evidently nucleation will not occur unless $\Delta G_v + E < 0$.

A crystal usually begins to flow when a macrostress corresponding to an elastic strain energy $E \sim 0{\cdot}01$ cal/cc is applied. However, according to the theory of plastic deformation perfect crystals should flow only if $E \lesssim 100$ cal/cc. Real crystals flow at very low applied stresses only because they contain non-equilibrated imperfections. These imperfections (Frank–Read[68] sources) generate dislocations that lead to plastic flow upon the application of small stresses. It is believed that the concentration of these imperfections in a well-annealed crystal is so small ($\sim 10^{12}$/cc) that there is a negligible probability that a nucleus, which has a volume of the order of 10^{-20} cc, formed at random will contain one. Therefore, in the vast majority of sites available for its formation transformation stress energies of the order of 100 cal/cc can be supported. However, the ability of elements of real crystals to support such strains varies from element to element and it is expected that when the transformation stresses are large nuclei will form preferentially in elements of the crystal that do not, because of imperfections in their vicinity, support large stresses. In general, Frank–Read sources will not operate unless the stress is applied throughout an element of volume much larger than 10^{-20} cc. These sources will, therefore, play an unimportant role in relieving transformation stresses in nuclei formed in their vicinity. However, stresses in very small volume elements in the vicinity of grain

boundaries and free surfaces probably can be partially relieved by flow.* For these reasons we expect grain boundaries and free surfaces to be preferred sites for nucleation in crystalline media.†

Since the dislocations in a real crystal are not in thermal equilibrium $\Delta G_v = \Delta G_v'$ for volume elements transforming in the vicinity of dislocations is not necessarily identical with ΔG_v for the average volume element. We expect that there will be some dislocation arrays for which $\Delta G_v > \Delta G_v'$. Further the magnitude of the elastic coefficients in the neighbourhood of dislocations may be markedly different from the mean elastic coefficients of the crystal. For these reasons some arrays of dislocations in the body of the crystal will be preferential sites for nucleation.

When a nucleus forms on a grain boundary some of the surface work required for its formation is regained by destruction of the grain boundary which has a free energy per area equal to σ_g. The free energy for forming an embryo on a grain boundary is:

$$\Delta G = (a\sigma - g\sigma_g)i^{2/3} + (\Delta G_v + E)bi, \quad . \quad . \quad . \quad . \quad (60)$$

where it is assumed that σ is not dependent on crystal orientation and g is determined by the shape of the boundary region destroyed. For example, if the embryo is spherical in form, equation (60) may be rewritten:

$$\Delta G = (4\sigma - \sigma_g)\pi r^2 + (\Delta G_v + E)\tfrac{4}{3}\pi r^3. \quad . \quad . \quad . \quad . \quad (61)$$

We have noted that in liquid and gas media nucleation may be heterogeneous because nuclei form preferentially on the surface of accidental inclusions. While accidental inclusions are probably important catalysts for nucleation in crystalline media the factors summarized above probably also contribute to an important degree to the heterogeneity of nucleation in such media. On the basis of these considerations it is expected that free surfaces, grain boundaries, inclusions, and elements in the body of the crystals having a high stress energy will serve as preferred sites for the nucleation of phases in crystals. For small $|\Delta G_v|$ it is probable that nucleation in crystals is entirely heterogeneous.

However, we shall develop the theory for the nucleation of a crystalline phase in a perfect crystal. The main features of this development can be adapted easily to a formal description of heterogeneous nucleation.

NABARRO[69] has developed the theory for the relation between strain energy due to precipitate particles and their shape and coherency. His treatment will be adapted to the development of the formal theory of nucleation for phase transformations involving no composition change.

* We are indebted to E. W. HART for advice on these points.

† There are additional reasons, to be discussed, why grain boundary sites should be preferred in nucleation.

Consider the phase transformation $\alpha \rightarrow \beta$, where α and β are crystalline phases. Let a_0 be the lattice spacing of some low index plane, P_1, in the β lattice and $a_0 + \Delta a$ the lattice spacing in similar low index plane of the α lattice, then we define the disregistry, δ:

$$\delta = |\Delta a/a_0| \quad \ldots . \quad (62)$$

In general the β nucleus may be strained an amount:

$$\varepsilon = |(x - a_0)/a_0| \quad \ldots . \quad (63)$$

where x is the lattice parameter of the P_1 plane of the nucleus. When $\varepsilon = \delta$ the nucleus is said to form coherently.

Suppose that the β embryo is completely incoherent with α (i.e. $\varepsilon = 0$). The accompanying volume change is accommodated by strain in the embryo and in the α matrix surrounding it, the strain distribution depending on the elastic coefficients of the two structures. However, following NABARRO, the strain energy of the system is a minimum when incoherent β has the form of a thin disc and all the strain is in the surrounding α.

Consider a hole in the α having the form of an ellipsoid with semi-axes a, a, and b. NABARRO calculated the strain energy per volume, E, required to expand the hole for the limiting forms thin disc, sphere and needle and proposed the following general relation between strain energy and shape:

$$E = K|\Delta V|^{2/3} f(b/a), \quad \ldots . \quad (64)$$

where ΔV is the difference in specific volume of the non-stressed α and β structures, K is proportional to the rigidity modulus of α, $f(b/a) = 0$ for $b/a = 0$, and is a maximum for a sphere ($b/a = 1$).

Although the strain energy of an incoherent embryo is zero when it has the form of an infinitely thin disc it will probably have the form of an oblate spheroid so that the $\alpha - \beta$ interfacial energy will not be too large. The free energy for forming the embryo is:

$$\Delta G = \tfrac{4}{3}\pi a^3(1 - e^2)^{1/2}\{\Delta G_v + K|\Delta V|^{2/3} f(1 - e^2)^{1/2}\} + \pi a^2 \sigma \left[2 + \left(\frac{1 - e^2}{e}\right) ln \left(\frac{1 + e}{1 - e}\right)\right], \quad \ldots . \quad (65)$$

where $e = (1 - b^2/a^2)^{1/2}$. The minimum free energy for forming a β nucleus is:

$$\Delta G^* = f(a^*, e^*), \quad \ldots . \quad (66)$$

where a^* and e^* simultaneously satisfy the equations:

$$[\partial(\Delta G)/\partial a]_e = 0 \quad \ldots . \quad (67)$$

$$[\partial(\Delta G)/\partial e]_a = 0 \quad \ldots . \quad (68)$$

Coherent and Semi-coherent Nucleation

Although the increase in strain energy due to the formation of incoherent β in the form of a thin disc is very small the surface energy may be quite large. Now suppose β forms so that a plane P_1 and crystallographic direction, d, is parallel to a plane and direction of similar atomic packing in the α. It has been postulated that the interfacial energy between two phases is a minimum when they are in perfect registry ($\delta = \varepsilon$) across their boundary.

In general, $\delta > \varepsilon$ and the actual disregistry is equal to $\delta - \varepsilon$. For $\delta - \varepsilon$ not too large (e.g. $\delta - \varepsilon < 0{\cdot}20$) the α-β interface across plane P_1 can be pictured as consisting of small regions of relatively good fit separated by an array of intersecting line dislocations.† The density of dislocations in this array is $\delta - \varepsilon$. FRANK and VAN DER MERWE[72] demonstrated that in a misfitting monolayer the energy due to dislocations is directly proportional to their density, ρ. It has been assumed[20] that $\rho \propto \delta - \varepsilon$ for the formation of a β nucleus on α, so that σ between α and β may be written[20]

$$\sigma = \gamma + \eta(\delta - \varepsilon), \qquad \ldots . \quad (69)$$

where η is a constant and γ is a term due to bond type and chemistry. It is apparent that σ is a minimum when $\delta = \varepsilon$.

In the case of semi-coherent precipitation the problem of minimizing the elastic energy of the system is somewhat more complicated than in the case of a completely incoherent particle. If the moduli of the precipitating particle are in all directions less than the moduli of the matrix material, then the precipitate will be plate-like so that the strain will reside primarily in the precipitate and the elastic energy minimized. In this case the elastic energy can be evaluated in the way indicated by NABARRO,[69] who has shown that E for a disc shaped coherent precipitate is of the order of $(1/5)E$ for a spherical precipitate. Applying these arguments to the nucleation of coherent β in α we expect that β embryos will probably again have the form of oblate spheroids and the free energy required to form them is:

$$\Delta G = \tfrac{4}{3}\pi a^3(1 - e^2)^{1/2}\{\Delta G_v + K'\varepsilon^2 f(1 - e^2)^{1/2}\} + \pi a^2[\gamma + \eta(\delta - \varepsilon)]\left[2 + \left(\frac{1 - e^2}{e}\right) ln\left(\frac{1 + e}{1 - e}\right)\right], \quad \ldots . \quad (70)$$

where K' is proportional to the elastic coefficient for the coherent formation of β. The minimum free energy for forming the nucleus is:

$$\Delta G^* = f(a^*, e^*, \varepsilon^*) \qquad \ldots . \quad (71)$$

† VAN DER MERWE[70] and BROOKS[71] have developed the description of interphase boundaries in terms of a dislocation model. BROOKS has discussed in general terms the application of the theory to nucleation in crystalline media.

where a^*, e^* and ε^* simultaneously satisfy the equations

$$[\partial(\Delta G)/\partial a]_{e,\varepsilon} = 0, \quad \ldots \quad (72)$$

$$[\partial(\Delta G)/\partial e]_{a,\varepsilon} = 0, \quad \ldots \quad (73)$$

$$[\partial(\Delta G)/\partial \varepsilon]_{a,e} = 0. \quad \ldots \quad (74)$$

In the case of a precipitate having all its elastic moduli larger than that of the matrix material, the shape of the precipitate which will minimize the elastic energy is that of a sphere. If, on the other hand, the modulus in one direction of the precipitate is less than the moduli of the matrix, the shape having minimum elastic energy will be rod-like. However, for embryos and nuclei, the anisotropy of the interfacial energy between particle and matrix will probably be more important in determining the shape of the particle than will considerations of elastic energy. The exact solution for the free energy of an embryo and a nucleus having varying elastic moduli can only be established if the dependence of the elastic modulus of matrix and precipitate on orientation as well as the anisotropy of the surface energy are known.

Theory of Catalysis of Crystal Nucleation

In a preceding section a formal theory of nucleation catalysis was developed in which the catalytic potency was characterized by the contact angle θ. While this theory proved to be satisfactory for describing the rate of solidification of mercury and tin droplets coated with various surface films, it gives little information on the mechanism of nucleation catalysis.

On the basis of the postulate made in the preceding section that the interphase interfacial energy is a minimum when the phases are coherent across the boundary, a theory of nucleation catalysis has been derived.[20] According to this theory the free energy of a crystallization nucleus formed on a catalyst surface is a minimum when the nucleus is at least partially strained in the direction of better atomic matching with the catalyst across the interface.

We obtain the interfacial energy between the catalyst and the nucleus, σ_{SC}, from equation (69). Now relating σ_{SC} to the contact angle through equation (15) we obtain:

$$1 - m = \beta + \eta(\delta - \varepsilon)/\sigma_{LC}, \quad \ldots \quad (75)$$

where $m = \cos\theta$ and

$$\beta = 1 - (\sigma_{LS} - \gamma)/\sigma_{LC}. \quad \ldots \quad (76)$$

The free energy of forming a nucleus having the shape of a spherical sector of radius r on a catalyst surface is

$$\Delta G_s = V(\Delta G_V + c\varepsilon^2) + 2\pi r^2(1 - m)\sigma_{LC} + \pi r^2(1 - m^2)(\sigma_{SC} - \sigma_{LS}), \quad \ldots \quad (77)$$

where c is the appropriate coefficient of elasticity and V is the volume of the sector. ΔG_s^* for a given strain is:

$$\Delta G_s^* = 4\pi\sigma_{LC}^3(2 + m)(1 - m)^2/3(\Delta G_V + c\varepsilon^2)^2, \quad . \quad . \quad . \quad . \quad (78)$$

and the value of $\varepsilon = \varepsilon^*$ that minimizes ΔG_s^* is:

$$\varepsilon^* = (\delta + \beta\sigma_{LC}/\eta) - [(\delta + \beta\sigma_{LC}/\eta)^2 - |\Delta G_V|/c]^{1/2}. \quad . \quad . \quad . \quad . \quad (79)$$

When the nucleation rate is perceptible, say 1 cm^{-2} sec^{-1}, $\Delta G_s^* \sim 60kT$ and $|\Delta G_V| = |\Delta G_V'|$. For $c\delta^2 \gg |\Delta G_V'|$ and making the approximation $2 + m = 3$ we obtain $|\Delta G_V'| = f(\delta)$ as follows:

$$|\Delta G_V'| = (4\pi\sigma_{LC}/60kT)^{1/2}(\beta\sigma_{LC} + \eta\delta). \quad . \quad . \quad . \quad . \quad (80)$$

If $|\Delta G_V'|$ is larger than $c\delta^2$ the nucleus forms coherently and we have:

$$|\Delta G_V'| = c\delta^2. \quad . \quad . \quad . \quad . \quad (81)$$

The approximate condition for coherent nucleation is found by equating (80) and (81) to be:

$$\delta \gtrsim [\eta\mu + (\eta^2\mu^2 + 4\beta\mu c)^{1/2}]/2c = \delta_c, \quad . \quad . \quad . \quad . \quad (82)$$

where $\mu = (4\pi\sigma_{LS}/60kT)^{1/2}$.

$|\Delta G_V'|$ is proportional, equation (30), to the supercooling $\Delta T'$ at which nucleation is perceptible. For $\delta < \delta_c$, $\Delta T'$ varies as δ^2. When $\delta \gg \delta_c$ and for a given value of β, $\Delta T'$ varies as δ.

The potency, p, of a nucleation catalyst is proportional (see equation (80)) to $1/(\beta\sigma_{LC} + \eta\delta)$. Although p and δ are reciprocally related the order of p is not necessarily the same as the order of $1/\delta$ because of β. If it is assumed that[20]

$$\sigma_{LC} = \sigma_{LS} - \gamma, \quad . \quad . \quad . \quad . \quad (83)\dagger$$

$\beta = 0$ and a particularly simple theory of nucleation catalysis results wherein the order of p is identical with the order of $1/\delta$.

When equation (83) holds equations (79), (80) and (82) reduce to, respectively:

$$\varepsilon^* = \delta - (\delta - |\Delta G_V|/c)^{1/2}, \quad . \quad . \quad . \quad . \quad (84)$$

$$|\Delta G_V'| = (4\pi\sigma_{LC}/60kT)^{1/2}\eta\delta, \quad . \quad . \quad . \quad . \quad (85)$$

and

$$\delta_c = (4\pi\sigma_{LC}/60kT)^{1/2}(\eta/c). \quad . \quad . \quad . \quad . \quad (86)$$

There are a few data sufficiently quantitative to test the predictions of the crystallographic theory of nucleation catalysis. When it is assumed that ice nuclei form coherently on silver iodide ($\delta = 0{\cdot}0145$) a value of $\Delta T' \approx 3{\cdot}1°$ is calculated[20] that is in good agreement with the value $\Delta T \approx 2{\cdot}5°$ measured by VONNEGUT. Experience indicates[20] that nuclei probably do not form coherently on catalyst surfaces for

† The plausibility of equation (83) has been discussed elsewhere [20]

$\delta > 0{\cdot}005$ to $0{\cdot}015$. This result is in rough agreement with the prediction of FRANK and VAN DER MERWE'S[72] theory.

It was pointed out[20] that most of the results on the kinetics of heterogeneous nucleation of mercury and tin crystals can be derived on the assumption that the nucleus is derived from a macroscopic structure having a melting point little, if any, different from the melting point of the stable structure. These results indicate that the nuclei formed incoherently and with negligible strain. Because of the large supercooling and the marked dissimilarity in structure between the droplet surface films and the forming metal crystals it is not surprising that the nuclei formed with $\varepsilon \sim 0$ in these experiments.

Theory of Martensite-type Transformations

Theory—We have been tacitly assuming that transformation nuclei are always characterized by the minimum possible free energy of formation. Actually the net rate of transformation, dx/dt, of α to β must be some function of the product of the rate of nucleation I and the rate of increase in volume of the nucleated region. The rate of increase in volume of the nucleated region is in general proportional to R^3, where R is the linear growth rate. Therefore,

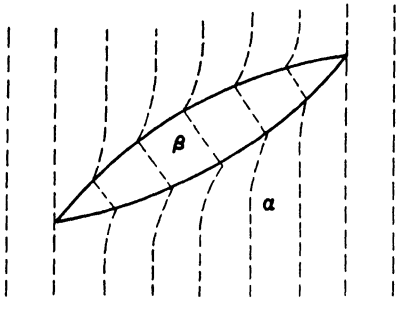


Fig. 16. Schematic representation of a martensite (β) embryo in an α matrix

$$dx/dt = f(IR^3). \quad . \quad . \quad . \quad . \quad (87)$$

Now the transformation from α to β will occur by a mode of nucleation and growth that will make dx/dt a maximum. Let the value of I corresponding to the minimum ΔG^* be I'. The structure of the nucleus corresponding to I' is not necessarily identical with the structure that leads to a maximum growth rate, R_{max}. We shall find that the martensitic mode of transformation is brought about by nuclei (S-nuclei) that require a much higher free energy for their formation than would nuclei of other possible configurations (N-nuclei). Nevertheless, transformation is effected by S-type rather than N-type nuclei because the rate of growth of S-regions is many orders of magnitude greater than that of N-regions.

It was pointed out that for the coherent formation of β in α the transformation strain energy is a minimum when β has the form of a thin disc and is strained in tension or compression to be coherent (normal or N-type coherency) with α across a particular crystallographic plane.

A different type of coherency results when the α matrix immediately surrounding the β is strained by shear to be coherent with β (shear or

S-type coherency). Fig. 16 shows schematically a β region surrounded by a region of α sheared to maintain coherency across the interface. E for a β-S embryo is larger than E for a β-N embryo. However, the atomic displacements required for the growth of S-nuclei are in most transformations much less than those required for the growth of N-nuclei. Therefore, S grains once formed grow at rates many orders of magnitude faster than N grains, so that even though the N-nuclei outnumber S-nuclei most of the transformation may be effected by S-nuclei.

FISHER, HOLLOMON and TURNBULL[73] have developed a theory for the energetics of formation of S-nuclei. They assumed that the embryos form in the shape of lenticular discs of radius r and thickness t. The volume of this disc is $\pi r^2 t/2$ and its area is $2\pi r^2$. It is assumed that the strain energy is localized in a sphere of radius r surrounding the embryo. The strain at any point in this sphere is t/r when $t \ll r$. Hence,

$$E = \tfrac{3}{4}c(t/r)^2, \quad \ldots \quad (88)$$

where c is proportional to the shear modulus of α. The free energy of formation of an S-embryo is then:

$$\Delta G = 2\pi r^2\sigma + \pi r^2 t\Delta G_v/2 + c\pi r t^2 \quad \ldots \quad (89)$$

and

$$\Delta G^* = f(r^*, t^*) \quad \ldots \quad (90)$$

where r^* and t^* simultaneously satisfy the equations

$$[\partial(\Delta G)/\partial r]_t = 0 \quad \ldots \quad (91)$$

$$[\partial(\Delta G)/\partial t]_r = 0 \quad \ldots \quad (92)$$

Carrying out the indicated operations we find:

$$\Delta G^* = 8192\pi c^2\sigma^3/27(\Delta G_v)^4, \quad \ldots \quad (93)$$

$$t^* = -16\sigma/3\Delta G_v, \quad \ldots \quad (94)$$

$$t/r = t^*/r^* = -\Delta G_v/4c. \quad \ldots \quad (95)$$

S-embryos grow and shrink rapidly so that down to relatively low temperatures α will contain an equilibrium distribution of β-S embryos characteristic of the given temperature. Therefore, this theory predicts that martensite transformation nuclei only form isothermally.

We have pointed out that nucleation in crystals may be heterogeneous to a marked degree because of non-equilibrated imperfections. That is, ΔG^* may vary from element to element in the α to give a spectrum of preferred nucleation sites characterized by different values of ΔG^* and volume, ι, of the element. Provided v is larger than the volume, v^*, of

the critical nucleus for growth of β into α of mean free energy the nucleation frequency may be written:

$$I = (K/V) \sum_{0}^{j} V_j \exp [- \Delta G_j^*/kT] \quad . \quad . \quad . \quad . \quad (96)$$

where V is the volume of the specimen and V_j is the total untransformed volume of elements characterized by ΔG_j^*. Let I_j be the nucleation frequency per volume corresponding to a particular ΔG_j^* and for simplicity assume that all the preferred sites have the same volume v and yield only one nucleus. It then follows that there will be a dependence of $I = I_t$ on time at constant temperature given by:

$$I_t = (1/V) \sum_{0}^{j} I_j V_j^\circ \exp [- I_j v t], \quad . \quad . \quad . \quad . \quad (97)$$

where $V_j = V_j^\circ$ at $t = 0$. This generalization predicts that isothermal martensite nuclei will form at preferred sites and that the nucleation frequency will decrease with increasing time.

Experience indicates that many martensite transformations are athermal—that is, the amount of transformation appears to be independent of time at temperature and is a function of temperature only. It is as though transformation nuclei were in existence at all times and that a certain number, n_0, of these become active for growth when the temperature is lowered by a unit amount—n_0 increasing with decreasing temperature.

This athermal characteristic of some martensite reactions is explained on the basis of either frozen-in composition fluctuations or by frozen-in variations in free energy from point to point in the α crystal due to local strains. Formally, the theory is very much the same from either point of view. We shall first develop the theory for the formation of athermal nuclei from composition fluctuations.

When α contains a major component A and a minor one B, where B has a strong stabilizing effect on α relative to β, we may expect that β embryos will form preferentially in regions that, because of composition fluctuations, are depleted in B. Let Δg_i be the increase in free energy associated with removing all i atoms of B from a region of α of average composition. According to the development of FISHER, HOLLOMON and TURNBULL[73] the number of these regions per volume, n_i, of pure A is:

$$n_i = n \exp [- \Delta g_i/kT] \quad . \quad . \quad . \quad . \quad (98)$$

where n is the number of lattice sites per volume. The number of these regions n_i^* exceeding a critical size corresponding to $i \geqslant i^*$ is:

$$n_i^* = \int_{i=i^*}^{\infty} n_i di = (n/i)(\Delta g_i/kT) \exp [- \Delta g_i^*/kT] \quad . \quad . \quad . \quad . \quad (99)$$

where $\Delta g_i = \Delta g_i^*$ at $i = i^*$.

Now suppose that at a given temperature the β-S nucleus would contain i^* atoms of B if it had the same composition as the α matrix. ΔG_v may be large enough so that the equilibrium concentration of S-nuclei, $n^* = n \exp [- \Delta G^*/kT]$, is negligible. However, the number of α regions, n_i^*, free of B and having dimensions equal to or larger than t^* and r^* can be appreciable. Since the volume free energy loss for formation of β in one of these $B-$ poor regions ($B-$) is much greater than for its formation in a region of uniform composition, the probability that β will nucleate in a $B-$ region is many orders of magnitude larger than the probability that it will nucleate in a region of uniform composition, $\bar{B}$.

When a β-S nucleus forms in a $B-$ element it will immediately grow to the boundaries of the element. Then if the dimensions of $B-$ element

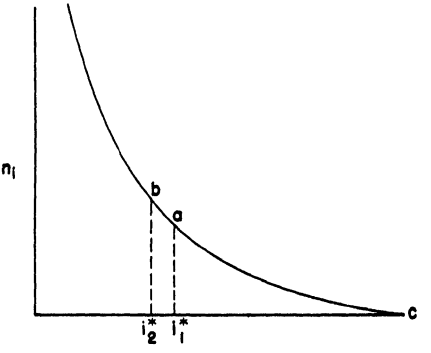


Fig. 17. Distribution of embryos, depleted in a minor solution constituent, with respect to size

equal or exceed t^* and r^* the nucleus will continue to grow into the $\bar{B}$ region providing the interfacial energy between $\bar{B}$ and $B-$ is negligible when the structure is the same on either side of the boundary. Therefore, in an alloy of the type we are considering it is possible that β is nucleated entirely by $B-$ elements formed by fluctuations.

At relatively low temperatures the composition fluctuations may be frozen in and the number and size of $B-$ regions will be the equilibrium number characteristic of a high temperature. We shall suppose that the $B-$ elements are entirely transformed to β. A schematic representation of the number of these $B-$ regions as a function of i (proportional to volume) is given in Fig. 17. At any temperature there is a critical volume, corresponding to i^*, of $B-$ that will transform the surrounding α of $\bar{B}$ composition to β. When a sample is quenched to any temperature T the number of these supercritical regions, hence β nuclei, is given by (99). A number, n_i^*, will grow immediately and the transformation will stop. When the temperature is further lowered i^* is decreased and more β nuclei are formed. Since these nuclei are brought into being not by thermal composition fluctuations at temperature but by having the

potential barrier to nucleation lowered they are called athermal nuclei. The total number that have formed upon quenching to temperature T_1 is proportional to the area $aci_1{}^*$, in Fig. 17. Upon quenching to T_2 an additional number per untransformed volume, proportional to the area $ai_1{}^*i_2{}^*b$ in Fig. 17, form.

When ΔG_v is sufficiently large so that the equilibrium concentration of β nuclei in $\bar{B}$ regions is appreciable nucleation takes place isothermally and at a steady rate. For alloys the possibility exists that β is nucleated by B– regions at temperatures too high for the formation of β nuclei in $\bar{B}$ regions. Nuclei formed in this way come into being only when the temperature is lowered and the martensite formed therefrom is athermal.

By considering the heterogeneous distribution of free energy expected in a real crystal a qualitative theory for athermal nucleation can be derived without invoking composition fluctuations. Suppose that (see equation (96)) the volume of the preferred nucleation sites is of the same order as v^*, the volume that the β nuclei must exceed in order to grow into the α matrix of *mean* free energy. Then at a preferred site having a volume, v, β may nucleate, grow to the site boundary, and be unable to grow further until the temperature is lowered sufficiently to make $v^* < v$. There will be some history-sensitive distribution of preferred sites that will be described by a function $n_v = f(v, \Delta G_v)$, where n_v is the number of sites characterized by volume v and free energy ΔG_v. The number of athermal nuclei, n, formed at a given temperature can be calculated from $n_v = f(v, \Delta G_v)$. n will be the number of high energy regions having a volume exceeding $\pi r^{*2}t^*/2$, where r^* and t^* are given by equations (94) and (95).

COHEN, MACHLIN and co-workers[74, 75] have proposed an alternative theory for the nucleation of martensite that ignores any contribution of interfacial energy to the free energy of nucleus formation. Let the angle through which α must be sheared to form β be $\theta_{\alpha-\beta}$. There will exist in the α many small embryos that are only partially sheared to β—that is, $\theta < \theta_{\alpha-\beta}$. These β embryos must be characterized by their shear angle θ and the number of atoms i that they contain.

Now according to the COHEN–MACHLIN theory, for a given value of i, the free energy of the β embryo goes through a maximum at $\theta = \theta_m$ as θ is increased. We may imagine a three dimensional surface $\Delta G = f(i, \theta)$ for constant temperature. The problem is to find the saddle point in this surface corresponding to the minimum free energy that the embryo must have before it can grow to a large size with the free energy continually decreasing.†

† It will be noted that the COHEN–MACHLIN theory of nucleation with i and θ the variables is the structure-strain analogue of the BORELIUS theory of nucleation of precipitate with i and Δx the variables.

In the $\Delta G = f(i, \theta)$ surface there is a single ridge (shear ridge), the maximum height of which is given by the relation $\Delta G_m = F(i)$, where $\Delta G = \Delta G_m$ at $\theta = \theta_m$. The value of i at the saddle point $i = i^*$ should be found by solving the equation $d(\Delta G_m)/di = 0$. The major difficulty of the COHEN–MACHLIN theory is that the functions $\Delta G = f(i, \theta)$ and $\Delta G_m = F(i)$ are only qualitatively defined. We expect that the increase in free energy of the embryo due to shearing through a given angle, e.g. $\theta = \theta_m$, should increase continuously with i and that the critical fluctuation, $\alpha \rightarrow \beta$, will, therefore, cross the shear ridge at a value $i \rightarrow 0$. Therefore, the COHEN–MACHLIN theory seems to require, as does the BORELIUS theory, an additional relation that will permit the calculation of the size of the region that undergoes the critical fluctuation.

COHEN, MACHLIN and co-workers cite the often-observed heterogeneous nature of the martensite transformation as strong supporting evidence for their theory. However, as we have seen, the surface barrier theory also predicts a heterogeneous martensite reaction if ΔG varies from one element of α volume to another due to non-equilibration.

Comparison of Theory and Experience

Non-martensite Transformations—Although there have been extensive measurements of the isothermal kinetics of various solid-solid non-martensitic N-type transformations there is little quantitative knowledge of the nucleation frequency for these transformations. It is known that the supercooling required to initiate N-type transformations in pure metals is usually very small; for example, of the order of 5–10° in the $\gamma \rightarrow \alpha$ iron transition. Orientation relations for these transformations are usually consistent with the description that the low index crystallographic planes and directions in the two structures having the most similar atomic packing and registry are parallel. Nuclei usually form preferentially in the vicinity of grain boundaries and free surfaces. More quantitative, carefully controlled measurements of nucleation frequency in N-type transformations are necessary in order to critically test the predictions of the nucleation theory.

Martensite Transformations—The data on nucleation frequency for S-type transformations are no more quantitative than for N-type but more information of value for comparison with nucleation theories can be inferred from the overall S-type transformation characteristics.

For many years it was believed that S transformations were insuppressible by quenching and almost entirely athermal in nature. Recently, however, KURDJUMOV and MAKSIMOVA[76] showed that the S transformation in certain iron alloys could be suppressed entirely by rapid quenching to liquid nitrogen temperature and upon subsequent isothermal

holding at this temperature and higher an S-type transformation occurred that was entirely isothermal in character. As much as 50 per cent of the γ-iron phase could transform to α by the S-mechanism isothermally. The essential requirement for the isothermal mode of transformation seemed to be that the alloy be characterized by an M_s temperature of the order of 300°K or less.

About the same time AVERBACH and COHEN[77, 78] showed that there was a small isothermal component of the S transformation in iron—1 per cent carbon alloys. However, the amount of alloy transformed isothermally was small relative to that transformed athermally. More recently DAS GUPTA and LEMENT[79] found that after some athermal S transformation in a high chromium steel a substantial amount of transformation took place isothermally. They found that considerable athermal transformation had occurred before the isothermal transformation began.

KULIN and COHEN[80] were unable to confirm KURDJUMOV and MAKSIMOVA'S result that athermal martensite formation in a 6 per cent manganese 0·6 per cent carbon steel could be completely suppressed by rapid quenching. However, it is significant that KURDJUMOV and MAKSIMOVA (K–M) report an M_s for this steel about 100° lower than that found by KULIN and COHEN[80] and other investigators; it seems likely, therefore, that K–M made an error in reporting the composition of the steel.

Recent results of CECH and HOLLOMON[81] on the isothermal S transformation of a 23 per cent Ni, 4 per cent Mn, 0·05 per cent carbon steel have completely confirmed K–M's main conclusions. The transformation was entirely suppressed, within the experimental uncertainty, upon rapid quenching to — 196°. Upon holding at various temperatures between — 196° and — 80° transformation occurred isothermally and as much as 50 per cent of the specimen transformed by the isothermal mode. The results were satisfactorily described by a conventional time-temperature transformation (T.T.T.) diagram (Fig. 18) with a nose temperature at — 126°C. Particular isotherms were described fairly well by straight line relations between the fraction transformed and the logarithm of the time. Metallographic examination disclosed that the rate determining step in the transformation is nucleation and not growth. K–M and DAS GUPTA and LEMENT[79] also found that nucleation is the rate-determining step in the isothermal transformation.

FISHER[82] established that CECH and HOLLOMON'S (C–H) isotherms can be described on the assumption of a constant nucleation frequency per untransformed volume. The rapid decrease in transformation rate with time is due to a partitioning effect whereby martensite plates already formed cut down the volume that can be transformed by nuclei formed at later stages of the transformation. FISHER[82] also demonstrated that

the C–H T.T.T. curves are described remarkably well by the FISHER–HOLLOMON–TURNBULL (F–H–T) theory of thermal martensite formation (equations (88) to (95)). Since the rate determining step is nucleation the temperature T_{max} corresponding to the nose of the T.T.T. curve can be calculated by substitution of equation (93) into (28). Upon performing this operation using JONES and PUMPHREY'S[83] free energy data and assuming $\Delta G_A = 0$, FISHER calculates $T_m = 147°K$, in excellent agreement with the measured $T_m = 147°K$. Further, with a single disposable parameter ($\sigma^3 c^2$) the entire T.T.T. curve was described from

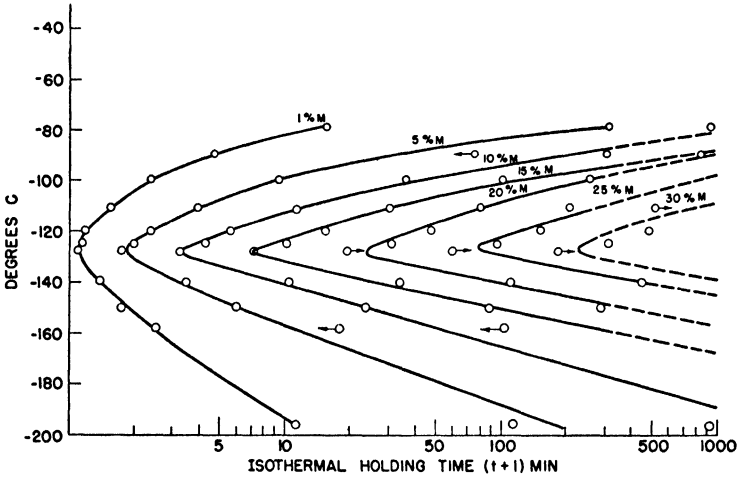


Fig. 18. T.T.T. curves for formation of martensite in Fe-Ni-Mn-C alloy. After CECH and HOLLOMON[81]

equations (24) and (93) within experimental error. K_v (see equation (24)) was estimated from the order of magnitude of I to be 10^{36} sec^{-1} cm^{-3}, which agrees very well with the value 10^{37} sec^{-1} cm^{-3} estimated from equation (24). This agreement between theory and experience is certainly not trivial; it constitutes powerful support for the F–H–T theory of nucleation without composition fluctuations in S transformations.

F–H–T have favoured the view that athermal martensite can be accounted for only by frozen-in composition fluctuations. For example, in iron base alloys the elements most effective in retarding the martensite transformation are those that decrease the free energy driving transformation. When the composition permits, fluctuations occur that lead to the formation of an appreciable concentration of regions free of alloying element and nucleation of α in these regions is much more probable than in regions of uniform composition. When an alloy is quenched to a particular temperature the total number, N, of martensite

plates that have formed per volume and when the fraction transformed is not large is:

$$N = It + n_i^* \qquad \ldots . \quad (100)$$

where I is the nucleation frequency per unit volume for regions of uniform composition $\{I = K_v \exp [- \Delta G^*/kT]$, where ΔG^* is given by equation (93)$\}$ and n_i^* given by equation (99) is the number of plates formed from nuclei produced by frozen-in composition fluctuations. The first term of equation (100) is the thermal component and the second term the athermal component of the nucleation. When the alloying content is very large (e.g. 20–30 atom per cent) the probability of forming a large region of pure parent metal is small, so that n_i^* is small and the thermal component, It, may predominate. However, if the alloying content is quite small n_i^* may become large before It becomes important and most of the transformation will take place by the athermal mode. A more quantitative theory of the relative magnitude of the thermal and athermal components is desirable.

F–H–T have demonstrated that the variation in M_s temperature with composition for the formation of athermal martensite in iron alloys containing carbon and chromium[84, 85] is described by the composition fluctuation theory with the use of only one disposable parameter which is σ, the α-γ interfacial energy per area. Also the data of HARRIS and COHEN[85] on the amount of athermal martensite formed as a function of temperature in a carbon-chromium steel were described within experimental error with c (see equation (93)) the only disposable parameter. It was shown that the values of σ (24 ergs per cm^2) and c ($0{\cdot}46 \times 10^{10}$ dynes per cm^2) chosen to give agreement between theory and experiment are plausible.

It has not yet been established whether or not composition fluctuations are a necessary condition for the athermal mode of martensite nucleation. Athermal martensite is known to form in the low temperature transition face-centred cubic → body-centred cubic in lithium[86] and in the cobalt[87] (face-centred cubic → hexagonal close packed) transformation. However, these substances were not of highest purity and no systematic study on the effect of alloying elements on the kinetics of the martensite transformation in these metals seems to have been made. In order to find whether or not composition fluctuations are a necessary concomitant of athermal martensite in these metals it will be necessary to make such a study on the effect of alloying elements.

The amount of martensite that forms, either thermally or athermally, at a given temperature T_1 may be materially decreased by decreasing the rate at which the specimen is initially cooled to T_1. This phenomenon is called stabilization. For example, CECH and HOLLOMON[81] found that the rate of isothermal transformation in the Fe-Ni-Mn alloy is substantially reduced by long holding times at room temperature

before quenching to the transformation temperature. Further if after the isothermal transformation is underway at a temperature T_1 the sample is upquenched to the ambient temperature and then quenched to T_1, the transformation rate at T_1 is very much less (by an order of magnitude) than before the upquenching.

COHEN, MACHLIN and PARANJPE[74] have suggested that stabilization results from the lowering of the stress levels at preferential nucleation sites in γ-iron by a recovery process. That is, stresses are developed in certain volume elements of the γ due to quenching, for example. α nuclei probably prefer to form at these stressed sites. These stresses may be partially relieved by recovery processes so that nucleation of α is less probable—hence, stabilization occurs.

STRUCTURE STRAIN AND COMPOSITION

Theory

An embryo for precipitation is described by its structure, composition, size, shape and state of strain. The problem of finding the most probable values of these variables for the critical cluster or nucleus is formidably complex. The free energy of formation of such a cluster, that is incoherent and has the composition of the equilibrium precipitate, is given by equation (65), except that now ΔG will be a function of x. For an oblate spheroidal form we have:

$$\Delta G_i = \tfrac{4}{3}\pi a^3(1-e^2)^{1/2}\left\{\frac{\Delta G_2}{v} + K\delta^2 f[(1-e^2)^{1/2}]\right\} + \pi a^2\sigma\left[2 + \frac{(1-e^2)}{e}\,ln\left(\frac{1+e}{1-e}\right)\right]. \quad . \quad . \quad . \quad . \quad (101)$$

The free energy of the critical cluster $\Delta G_i^* = f(e^*, a^*)$ is found in the same way as before (see equations (67) and (68)).

For reasons similar to those already outlined in the section on structure and strain, precipitation nuclei may prefer to form coherently or semi-coherently in order to minimize their surface energy. The expression for the free energy of formation of semi-coherent clusters is analogous to equation (70), if the precipitate is softer elastically than the matrix,

$$\Delta G_i = \tfrac{4}{3}\pi a^3(1-e^2)^{1/2}\{\Delta G_2/V + K'\varepsilon^2 f[(1-e^2)^{1/2}]\} + \pi a^2[\gamma + \eta(\delta-\varepsilon)]\left[2 + \frac{(1-e^2)}{e}\,ln\left(\frac{1+e}{1-e}\right)\right], \quad . \quad . \quad . \quad . \quad (102)$$

where K' is the elastic coefficient appropriate to the coherent straining of the cluster. The free energy of the critical cluster $\Delta G_i^* = f(a^*, e^*, \varepsilon^*)$ is found by the procedure indicated in the earlier equations, (72) through (74).

If δ is relatively small and γ is relatively large, clusters having a nearly spherical shape will form. Following FISHER, HOLLOMON and LESCHEN[88] the free energy of forming a partly coherent spherical cluster is written:

$$\Delta G_i = [\gamma + \eta(\delta - \varepsilon)]4\pi r^2 + (4/3)\pi r^3(\Delta G_v + c\varepsilon^2) \quad . \quad . \quad . \quad . \quad (103)$$

where $\Delta G_v = \Delta G_2/V$ and c is the appropriate elastic coefficient. It follows that the free energy of formation of the critical cluster is:

$$\Delta G_i^* = (16\pi/3)[\gamma + \eta(\delta - \varepsilon^*)]^3/(\Delta G_v + c\varepsilon^{*2})^2, \quad . \quad . \quad . \quad . \quad (104)$$

where $$\varepsilon^* = 2(\delta + \gamma/\eta) - \{4(\delta + \gamma/\eta)^2 + 3\Delta G_v/c\}^{1/2}, \quad . \quad . \quad . \quad . \quad (105)$$

and the condition for coherency is:

$$|\Delta G_v| \lesssim c\delta^2\{1 + 4\gamma/3\eta\delta\} = Z. \quad . \quad . \quad . \quad . \quad (106)$$

When the supersaturation is very large $|\Delta G_v| > Z$ and the nucleus forms coherently. When $Z \gg |\Delta G_v|$, $\delta \gg \varepsilon$.

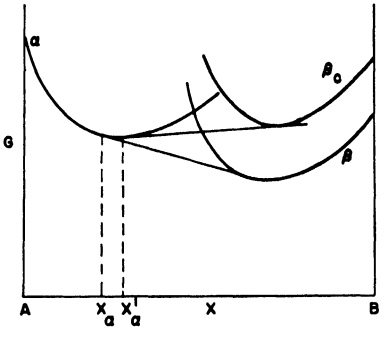


Fig. 19. Comparison of free energy-composition dependence of coherent and non-coherent β phases. After FISHER, HOLLOMON and LESCHEN[88]

One important consequence of the theory of precipitation is that the coherent precipitate can form and grow only at temperatures well below the temperature at which the completely non-coherent precipitate is in equilibrium with the solid solution. This consequence was recognized by NABARRO[69] and BARRETT[89] and has been treated in more detail by LESCHEN and FISHER.[90] Following FISHER, HOLLOMON and LESCHEN[88] the free energy relations between solid solution, α, and coherent, β, and non-coherent, β_c, phases are shown schematically in Fig. 19. The free energy of β_c at a given x exceeds that of β by an amount $c\delta^2$. The temperature at which β is in equilibrium with α is ΔT lower than the temperature at which α and β are in equilibrium.

As pointed out by NABARRO, where δ is relatively large, e.g. $\lesssim 0{\cdot}10$, it is sometimes not thermodynamically possible to form a completely coherent precipitate.

The problem becomes additionally complex when the cases of rigid and semi-rigid precipitates are taken into account. Then the shape and energy of the particle will depend on the relative elastic moduli and the anisotropy of the surface energy.

Comparison of Experience and Theory

To the best of the authors' knowledge there have been few quantitative measurements of the nucleation frequency of precipitate from solid

solutions. However, some generalizations can be made from qualitative observations. GEISLER[91, 92] has emphasized that the vicinity of grain boundaries and slip bands are preferred sites for the nucleation of precipitate when the supersaturation is not large. Thus non-equilibrated structural defects, such as grain boundaries and possibly dislocations, apparently catalyze nucleation of precipitate. The preference of nuclei for grain boundary sites is understood, since grain boundary energy is released (see equation (61)) by the formation of a cluster thereon. It is not known to what extent nucleation of precipitate is catalyzed by accidental inclusions.[92]

A second important generalization is that there exists always a definite crystallographic orientation relation between the precipitate particles and the matrix. In general, this relation is that the planes and directions in which the atomic packing is most similar are parallel in the two lattices. For example, in the precipitation of a face-centred cubic precipitate from a face-centred cubic matrix the cube edges and cube planes in the two lattices are usually parallel. This result might be explained either on the basis that only the particles with the preferred orientations nucleate or grow rapidly. However, on the basis of the theory that has been developed there is good reason for expecting the nucleation of preferred orientations.*

From the dislocation theory of interphase boundaries it is expected that their interfacial energy will be a minimum when δ is a minimum. Therefore, on the basis of surface energy considerations only, clusters having an orientation relation with the matrix that minimizes δ will form preferentially. The observed orientation relationships for precipitation are in general consistent with the principle that δ be a minimum for low index planes.

It seems probable that the dominant factor controlling the orientation relation of an incoherent nucleus with the matrix is surface energy and that strain energy will have little effect. This conclusion follows from the NABARRO theory that the major part of the strain is generally taken up by the matrix when incoherent clusters are formed. However, when the nucleus is partly coherent with the matrix it appears that the orientation relation will be controlled by the relation of the elastic coefficients and interfacial energy to crystal direction. For if the coherency strain is taken up by the cluster it is evident that the coherency plane relation corresponding to the minimum c will minimize the strain energy. The orientation relation that minimizes the strain energy of the critical cluster does not necessarily minimize the interfacial energy.

In summary the problem of nucleation of precipitates from solid

* It is certain that an "oriented growth" theory cannot account for preferred orientations of precipitates in a cubic matrix when the growth is diffusion controlled.

solution is very complex and the most useful approximations have not yet been established. Even so, the general nature of such precipitation can be described. The precipitation process that will be observed at any temperature below the solvus will be the one that occurs at the maximum rate at that temperature. At temperatures just below the solvus the precipitation will occur heterogeneously and incoherently. If the temperature is further lowered, the rate of overall precipitation can be increased either by an increase in the rate of nucleation or by an increase in the rate of growth. Such an increase can be accomplished by adjusting the composition or structure to minimize the surface energy between precipitate and matrix or by adjusting the composition and structure to allow a more rapid rate of growth of the precipitate. One way in which the surface energy is reduced is by the formation of coherent precipitates which can occur only at temperatures sufficiently below the solvus to make the free energy available for straining the precipitate. Thus at temperatures below that at which incoherent precipitation occurs coherent homogeneous nucleation will occur at a more rapid rate. At still lower temperatures, the rate of overall transformation can be further increased by allowing the precipitate to form with the same composition as the matrix thus eliminating the necessity for long-range diffusion and forming a coherent particle in such a way that transport of atoms across the interface can occur rapidly. This latter type of transformation has been given the name martensite. Thus three distinct types of precipitation are possible: first, coherently nucleated, second, incoherently nucleated, and martensitic. There are obviously intermediate types, for example, when the precipitation occurs partially coherently.

There is much valuable data on orientation relations in precipitation that can be qualitatively explained. Also there is reliable information on the coherency relations in many systems. Unfortunately, there are few data on the isothermal course of precipitation as a function of time and even less quantitative data on the nucleation frequency in precipitation from solid solution. There is a great need for such data to test the prediction of current theories of nucleation of precipitate and to aid their further development.

References

[1] Smoluchowski, R. *Phase Transformations in Solids*, pp. 149–183, Wiley 1951

[1a] Turnbull, D. and Hollomon, J. H. *The Physics of Powder Metallurgy*, pp. 109–143, McGraw-Hill 1951

[2] Hollomon, J. H. "Thermodynamics in Metallurgy," pp. 161–177, *Amer. Soc. Metals* (1950)

[3] Bradley, R. S. *Quarterly Reviews* 5 (1951) 315

[4] La Mer, V. K., Pound, G. M. and Reiss, H. *Ind. Eng. Chem.* 44 (1952) 1270–1288

[5] JOHNSON, W. A. and MEHL, R. F. *Trans. Amer. Inst. Min. Met. Engrs.* 135 (1939) 416
[6] AVRAMI, M. *J. Chem. Phys.* 7 (1939) 1103, *ibid.*, 8 (1940) 212, *ibid.* 9 (1941) 177
[7] BURKE, J. E. and TURNBULL, D. *Progress in Metal Physics*, Vol. 3 pp. 220–293
[8] VOLMER, M. and WEBER, A. *Z. phys. Chem.* 119 (1925) 277
[9] BECKER, R. and DORING, W. *Ann. Phys.* [5] 24 (1935) 719
[10] BUFF, F. P. and KIRKWOOD, J. G. *J. Chem. Phys.* 18 (1950) 991
[11] REISS, H. *J. Chem. Phys.* 20 (1952) 1216
[12] TOLMAN, R. C. *J. Chem. Phys.* 17 (1949) 333
[13] KIRKWOOD, J. G. and BUFF, F. P. *J. Chem. Phys.* 17 (1949) 338
[14] BUFF, F. P. *J. Chem. Phys.* 19 (1951) 1591
[15] ZELDOVICH, J. B. *Acta Physicochim* U.R.S.S., 18 (1943) 1
[16] TURNBULL, D. Amer. Inst. Min. Met. Engrs, T.P. No. 2365, *Metals Tech.* (1948)
[17] KANTROWITZ, A. *J. Chem. Phys.* 19 (1951) 1097
[18] PROBSTEIN, R. F. *J. Chem. Phys.* 19 (1951) 619
[19] VOLMER, M. *Zeit. Elektrochem.* 35 (1929) 555
[20] TURNBULL, D. and VONNEGUT, B. *Ind. Eng. Chem.* 44 (1952) 1292
[21] VOLMER, M. and FLOOD, H. *Z. phys. Chem.* 170A (1934) 273
[22] SANDER, A. and DAMKÖHLER, G. *Naturwiss.* 31 (1943) 460
[23] POUND, G. M. and MADONNA, to be published (1952)
[24] WEGENER, P. and SMELT, R. *Naval Ordnance Laboratory Memorandum* 10772 (Feb. 15, 1950)
[25] BECKER, R. *Ann. Phys.* 32 (1938) 128
[26] TURNBULL, D. and FISHER, J. C. *J. Chem. Phys.* 17 (1949) 71
[27] TURNBULL, D. *J. Chem. Phys.* 18 (1950) 198
[28] TURNBULL, D. *J. Chem. Phys.* 20 (1952) 411
[29] VOLMER, M. *Kinetik der Phasenbildung* Steinkopf, Leipzig 1939
[30] AVRAMI, M. *J. Chem. Phys.* 7 (1939) 1103
[31] FISHER, J. C. *J. appl. Phys.* 19 (1948) 1062
[32] BERNATH, L. *Ind. Eng. Chem.* 44 (1952) 1310
[33] URSPRUNG, A. *Ber. d. Bot. Ges.* 33 (1915) 153
[34] MEYER, J. *Zur Kenntnis des Negativem Druckes in Flussigkeiten* W. Knapp, Halle, 1911
[35] BRIGGS, L. J. *Science* 109 (1949) 440, *ibid.* 112 (1950) 427
[36] MENDENHALL, C. E. and INGERSOLL, L. R. *Phil. Mag.* 15 (1908) 205
[37] VONNEGUT, B. *J. Colloid. Science* 3 (1948) 563
[38] CWILONG, B. M. *Nature*, Lond. 155 (1945) 361
[39] SCHAEFER, V. J. *Bull. Amer. Meteorlog. Soc.* 29 (1948) 175
[40] DANILOV, V. and NEUMARK, W. *Phys. Z. Sowjet.* 12 (1937) 313
[41] TURNBULL, D. *J. appl. Phys.* 20 (1949) 817
[42] —— "Thermodynamics in Metallurgy," *Amer. Soc. Metals* (1950) 282
[43] —— *J. appl. Phys.* 21 (1950) 1022
[44] DORSCH, R. G. and HACKER, P. T. *Nat. Advisory Comm. Aeronaut. Repts. Tech. Mem. Notes* (1950) 2142
[45] LEVINE, J. *Nat. Advisory Comm. Aeronaut. Repts. Tech. Mem. Notes* (1950) 2234
[46] WEIBULL, W. *Roy. Swed. Inst. Eng. Research*, No. 151 (1939)
[47] FISHER, J. C. and HOLLOMON, J. H. *Metals Tech.* 14, No. 5 (1947)
[48] GRIFFITH, A. A. *Philos. Trans. Roy. Soc.* 221A (1921) 163
[49] POUND, G. M. and LA MER, V. K. *J. Amer. Chem. Soc.* 74 (1952) 2323
[50] TURNBULL, D. and CECH, R. E. *J. appl. Phys.* 21 (1950) 804
[51] TURNBULL, D. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 1144
[52] SCHAEFER, V. J. *Ind. Eng. Chem.* 44 (1952) 1300

[53] FALKENHAGEN, G. and HOFMANN, W. *Z. Metallk.* 43 (1952) 69
[54] SCHAEFER, V. J. Private Communication
[55] BARDENHEUR, P. and BLECKMANN, R. *Stahl u. Eisen* 61 (1941) 49
[56] WEBSTER, W. L. *Proc. Roy. Soc.* 140A (1933) 653
[57] KHAIKIN, S. E. and BENET, H. P. *Comptes Rendus* (Doklady) *de l'Academie des Sciences de l'URSS*, 23 (1939) 31
[58] BORELIUS, G. *Ann. Physik*, 28 (1937) 507, *Arkiv Mat. Astron. Fysik* 32A (No. 1) (1944) 1
[59] HOBSTETTER, J. N. *Metals Tech.* 15 (1948) T.P. 2447
[60] KONOBEYEVSKI, S. *Z. phys. Chem.* 171 (1934) 75
[61] CECH, R. E. and TURNBULL, D. *J. Metals* 3 (1951) 242
[62] HOLLOMON, J. H. and TURNBULL, D. *J. Metals* 3 (1951) 803
[63] HARDY, H. K. *J. Inst. Metals* 77 (1950) 457
[64] BORELIUS, G., LARRIS, F. and OHLSSON, E. *Arkiv. Mat. Astron. Fysik*, 31A, No. 10 (1944)
[65] JOHANSSEN, C. K. and HAGSTEN, G. *Ann. Phys.* 28 (1937) 520
[66] WERT, C. A. and ZENER, C. *J. appl. Phys.* 21 (1950) 5
[67] WERT, C. A. *J. appl. Phys.* 20 (1949) 943
[68] FRANK, F. C. and READ, W. T. *Phys. Rev.* 79 (1950) 722
[69] NABARRO, F. R. N. *Proc. Phys. Soc.* 52 (1940) 90, *Proc. Roy. Soc.* 175 (1940) 519
[70] VAN DER MERWE, J. H. *Proc. Phys. Soc.* 63A (1950) 616
[71] BROOKS, H. "Metal Interfaces" *Amer. Soc. Metals*, pp. 20–65 (1952)
[72] FRANK, F. C. and VAN DER MERWE, J. H. *Proc. Roy. Soc.* 198A (1949) 216
[73] FISHER, J. C., HOLLOMON, J. H. and TURNBULL, D. *Trans Amer. Inst. Min. Met. Engrs.* 185 (1949) 691
[74] COHEN, M., MACHLIN, E. S. and PARANJPE, V. G. "Thermodynamics in Physical Metallurgy" *Amer. Soc. Metals* (1950)
[75] MACHLIN, E. S. and COHEN, M. *J. of Metals* 4 (1952) 489
[76] KURDJUMOV, G. P. and MAKSIMOVA, O. P. *Doklady Akad. Nauk.* S.S.S.R. 61 (1948) 83, 73 (1950) 95.
[77] AVERBACH, B. L., COHEN, M. and FLETCHER, S. G. *Trans. Amer. Soc. Metals* 40 (1948) 728
[78] AVERBACH, B. L. and COHEN, M. *Trans. Amer. Soc. Metals* 41 (1949) 1024
[79] DAS GUPTA, S. C. and LEMENT, B. S. *Trans. Amer. Inst. Min. Met. Engrs.* 191 (1951) 727
[80] KULIN, S. A. and COHEN, M. *Trans. Amer. Inst. Min. Met. Engrs.* 188 (1950) 1139
[81] CECH, R. E. and HOLLOMON, J. H. Submitted to Amer. Inst. Min. Met. Engrs. (1952)
[82] FISHER, J. C. To be published
[83] JONES, F. W. and PUMPHREY, W. I. *J. Iron Steel Inst.* 163 (1949) 121
[84] LYMAN, T. and TROIANO, A. R. *Trans. Amer. Inst. Min. Met. Engrs.* 162 (1945) 196, *Trans. Amer. Soc. Metals* 37 (1946) 402
[85] HARRIS, W. J. and COHEN, M. *Trans. Amer. Inst. Min. Met. Engrs.* 180 (1949) 447
[86] BARRETT, C. S. and TRAUTZ, O. R. *Metals Tech.* T.P. 2346 (1948)
[87] TROIANO, A. R. and TOKICH, J. L. *Metals Tech.* T.P. 2348 (1948)
[88] FISHER, J. C., HOLLOMON, J. H. and LESCHEN, J. G. *Ind. Eng. Chem.* 44 (1952) 1324
[89] BARRETT, C. S. *Phase Transformations in Solids*, p. 536, Wiley 1951
[90] LESCHEN, J. G. and FISHER, J. C. *Science* 112 (1950) 448
[91] GEISLER, A. H. *Phase Transformations in Solids*, pp. 387–545, Wiley 1951
[92] —— "Metal Interfaces" pp. 269–299, *Amer. Soc. Metals* (1952)

AUTHOR INDEX

AARTS, W. H. 237, 262
ACHTER, M. R. 199, 203, 254, 264, 287, 289, 290, 330
AKEROYD, E. J. 104
ALEXANDER, B. H. 182, 189, 198, 203, 272, 314, 330, 332
ALFREY, T. 67
ALLEN, N. P. 111, 157, 168, 171
ALLIBONE, T. E. 201
AMELINCKZ, S. 260
AMES, B. N. 171
ANDERSON, P. A. 126, 169
ANDERSON, W. A. 103, 159, 162, 171, 264
ANDRADE, E. N. DA C. 245, 262, 263
ANDREW, J. H. 171
ANDREW, K. F. 102, 116, 117, 165, 166, 168
ANDREWS, K. W. 127, 169
ANDREWS, M. R. 141, 142, 170
ARCHAROW, W. J. 103
ARCHBUTT, S. L. 171
ARDLEY, G. 262
ARMBRUSTER, M. H. 172
ARNER, L. H. 172
ASARO, F. 332
AUST, K. T. 227, 262
AVRAMI, M. 334, 346, 387
AVERBACH, B. L. 380, 388
AYTEKIN, V. 252, 259, 263

BAILEY, G. L. J. 187, 203
BAILEY, R. W. 264
BAKELAAR, I. D. 331
BAKER, W. A. 171, 172
BALKE, C. W. 185, 203
BALSHIN, M. Y. 178, 186, 203
BALUFFI, R. W. 182, 203, 272, 330, 332
BANCROFT, D. 68
BARDEEN, J. 102, 277, 328, 330, 332
BARDENHEUR, P. 358, 388
BARNES, R. S. 266, 270, 273, 276, 291, 322, 330
BARRER, R. M. 111, 114, 115, 147, 168, 331
BARRETT, C. S. 230, 261, 384, 388
BARTHOLOMEYCZYK, W. 167
BASTIEN, P. 116, 168
BAUKLOH, W. 167
BAUMBACH, H. H. VON. 102
BECK, P. A. 253, 254, 264, 330
BECKER, R. 62, 64, 70, 207, 333, 335, 336, 337, 338, 339, 343, 361, 363, 367, 387
BEDWORTH, R. E. 72, 96, 103
BEILBY, G. 176, 203
BÉNARD, J. 94, 102, 132, 169
BENET, H. P. 359, 388
BENNEWITZ, K. 61, 70
BERMANN, F. 287, 330
BERNATH, L. 347, 387
BETTERIDGE, W. 264
BEVER, M. B. 171
BHAT, U. V. 116, 168
BIER, C. J. 180, 203
BILBY, B. A. 38, 68, 222, 223, 241, 261, 263
BIRCHENALL, C. E. 94, 95, 96, 102, 181, 200, 201, 203, 204, 293, 321, 330, 331, 332
BLAZEY, C. 151, 170
BLECKMANN, R. 358, 388
BLIN 266, 268, 270, 271, 330
BLOCK, J. 103
BOAS, W. 261
BOBALEK, E. G. 130, 146, 169
BOER, J. H. DE. 146, 167, 168
BOGDANOFF, J. L. 262
BORELIUS, G. 263, 361, 364, 365, 367, 378, 379, 388
BORDONI, P. G. 68, 69
BOULANGER, CH. 26, 51, 52, 65, 66, 68, 69
BOWLES, J. S. 261
BOZORTH, R. M. 62, 70
BRADASCHIA, G. 171
BRADLEY, A. J. 331
BRADLEY, R. S. 176, 203, 333, 386
BRATTAIN, W. H. 102
BRAUN, H. 110, 167
BREITFELD, H. 68
BRICK, R. M. 69
BRIESEMEISTER, S. 171
BRIGGS, L. J. 387
BRODKORB, F. 127, 169
BROWN, A. F. 211, 215, 251, 261, 263
BROWN, H. 171
BROOKS, H. 371, 388
BRÜNING, H. 167, 170
BÜCKLE, H. 266, 268, 270, 271, 330
BUFF, F. P. 337, 338, 342, 387
BUFFINGTON, F. S. 331
BURGERS, J. M. 68, 262
BURGERS, W. G. 253, 254, 264
BURGESS, H. 305, 331

Burke, J. E. 334, 387
Burns, B. D. 263
Burton, H. H. 165, 171
Burton, W. K. 260
Busk, R. S. 130, 146, 169
Butler, C. C. 188, 204
Buttner, F. H. 203

Cabarat, R. 68
Cabrera, N. 72, 97, 100, 102, 196, 197, 203, 260, 316, 332
Cady, W. D. 68
Cahn, R. W. 247, 255, 261, 264
Calnan, E. A. 263
Campbell, I. E. 143, 148, 149, 150, 170
Campbell, W. E. 97, 102
Cannon, H. S. 186, 204
Carreker, R. 264
Castellan, G. W. 81, 102
Cech, R. E. 355, 366, 380, 382, 387, 388
Chalmers, B. 47, 69, 70, 102, 217, 227, 261, 262, 264, 330, 332
Chaston, J. C. 154, 162, 163, 164, 171
Chaudron, G. 110, 131, 132, 136, 168, 169
Chen, N. K. 247, 263
Chevenard, P. 26, 68
Child, F. C. 171
Childs, B. G. 320, 332
Christensen, H. 70
Christian, J. W. 223, 261
Churchman, A. T. 151, 170, 243, 263
Clark, H. T. 130, 147, 169
Clark, P. 178, 190, 191, 192, 203
Clarke, H. 50, 68
Claus, W. 171
Coehn, A. 146, 170
Cohen, M. 261, 331, 332, 378, 379, 380, 382, 383, 388
Cook, M. 264
Cook, N. 180, 203
Cooke, W. T. 68
Coquelle, O. 94, 102
Cottrell, A. H. 38, 68, 70, 150, 151, 170, 221, 222, 223, 241, 243, 249, 252, 259, 260, 262, 263, 264
Crank, J. 332
Crussard, C. 247, 263
Cwilong, B. M. 349, 355, 387

Damköhler, G. 342, 355, 387
Daniels, F. W. 253, 263
Danilov, V. 357, 387
Danz, W. 134, 140, 170
Dardel, Y. 172
Darken, L. S. 69, 80, 102, 153, 171, 269, 270, 272, 320, 321, 322, 323, 326, 327, 328, 330
Das Gupta, S. C. 380, 388
Davies, M. H. 94, 95, 96, 102
Davis, M. 257, 264
Dawson, I. M. 260
Dawson, M. H. 189, 198, 203, 289, 220, 314, 330, 332
Dedrick, J. H. 179, 183, 203
Denbigh, K. G. 332
Delisle, L. 174, 185, 203
Dewhirst, D. W. 155, 171
Dienes, G. J. 311, 312, 331, 332
Dijkstra, L. J. 36, 37, 68, 119, 121, 122, 123, 169, 238, 262
Döring, W. 62, 64, 70, 207, 335, 336, 337, 338, 339, 343, 387
Dorsch, R. G. 351, 387
Doty, P. M. 67
Drapeau, P. 178, 203
Dravnieks, A. 142, 170
Dreyer, K. L. 145, 170
Druyvesteyn, M. J. 143, 145, 155, 156, 170, 171
Duftschmid, F. 174, 178, 204
Dugdale, R. A. 237, 262, 332
Dumbleton, M. J. 263
Dunbar, L. W. 203
Dunn, C. G. 187, 203, 227, 253, 262, 263
Dünwald, H. 102
Dushman, S. 113, 116, 167, 203
Duwez, P. 178, 188, 203

Earle, M. D. 80, 102
Eastwood, L. W. 167, 171
Eborall, R. 164, 171
Edwards, C. A. 172
Edwards, O. S. 261
Eggert, S. 104
Einstein, A. 77
Ehrlich, F. 168
Ehrlich, P. 80, 102
Engell, H. J. 72, 97, 99, 101, 102, 103
Erickson, J. L. 172
Eselgroth, T. W. 172
Eshelby, J. D. 67, 213, 214, 261, 262
Eudier, M. 185, 204
Evans, J. W. 113, 168
Evans, U. R. 102
Eyring, H. 67, 114, 168, 295, 331

Faivre, R. 80, 102
Fahrenheit, G. D. 348
Falkenhagen, G. 356, 388

FAST, J. D. 105, 107, 114, 123, 126, 130, 146, 147, 167, 168, 169
FEITKNECHT, W. 96, 102
FENSHAM, P. J. 286, 330, 331
FINE, M. E. 26, 68
FINLAY, W. L. 148, 149, 150, 170
FISCHBECK, K. 93, 102
FISHER, J. C. 187, 203, 211, 215, 251, 261, 282, 284, 286, 291, 294, 330, 331, 343, 344, 347, 351, 375, 376, 380, 381, 384, 387
FITZGERALD, J. V. 68
FITZWILLIAM, J. 134, 170
FLANAGAN, R. 289, 290, 292, 331
FLINT, G. 103
FLOOD, H. 341, 342, 387
FOOTE, F. 80, 103
FÖPPL, O. 70
FORD, G. W. 111, 168
FOREMAN, A. J. E. 231, 232, 262
FÖRSTER, F. 28, 51, 65, 68, 69, 70
FORTY, A. J. 209, 260
FOUND, G. H. 69
FRAENKEL, W. 272, 330
FRANK, F. C. 70, 73, 102, 208, 209, 210, 216, 220, 222, 223, 230, 248, 260, 261, 262, 276, 330, 368, 371, 374
FRANKLIN, A. W. 264
FRAUNFELDER, H. 294, 331
FREAR, C. L. 172
FRENKEL, J. 73, 83, 102, 188, 189, 190, 203, 212
FROELICH, K. W. 96, 102
FRÖHLICH, W. 151, 156, 170
FUKUSIMA, M. 167
FULLMAN, R. L. 219, 261
FUNK, E. R. 203
FUNK, W. 167
FUOSS, R. M. 332
FUSFELD, H. I. 52, 68, 69

GAINES, J. M. 111, 168
GALLISSOT, M. 135, 170
GEACH, G. A. 178, 182, 203
GEGG, C. C. 165, 171
GEISLER, A. H. 385, 388
GENSCH, CH. 80, 83, 84, 87, 102, 103
GERDS, A. 182, 203
GERNER, W. 172
GIBB, T. R. P. 168
GIBBONS, D. F. 263
GIBBS, W. 207
GLASSTONE, S. 331
GOLDSCHMIDT, V. M. 85
GOETZEL, C. G. 174, 203
GOLIK, V. R. 142, 170
GORDON, E. 172
GORSKY, W. S. 45, 46, 69
GRANT, N. J. 172
GREENOUGH, A. P. 181, 182, 199, 203
GREENOUGH, G. B. 68, 69, 263, 264, 331
GREESE, R. 88, 103
GRIFFIN, L. J. 209, 260
GRIFFITH, A. A. 351, 387
GROBE, A. H. 103, 162, 171
GROOT, S. R. DE 332
GROSHEIM-KRISKO, K. W. 112, 168
GRUBE, G. 185, 203
GRUENEWALD, K. 79, 80, 81, 97, 101, 104
GRUNEWALD, H. 88, 103
GUILLET, L. 49, 69
GUINIER, A. 253, 263
GULBRANSEN, E. A. 80, 93, 96, 97, 102, 103, 116, 117, 168
GULDNER, W. G. 168
GUNDERMANN, J. 75, 102
GURRY, R. W. 80, 102
GUTTMAN, L. 69
GYULAI, Z. 237, 262

HAAYMAN, P. W. 90, 104
HACKER, P. T. 351, 387
HAGEN, H. 167
HÄGG, G. 110, 167
HAGSTEN, G. 367, 388
HALL, E. O. 262, 263
HALL, M. N. A. 147, 170
HALSEY, G. 67
HAM, W. R. 114, 168
HAMMEN, H. 104
HANEMANN, H. 168
HANSON, D. 111, 167, 168, 264
HARDY, H. K. 367, 388
HARPER, S. 38, 68, 242, 262, 263
HARRINGTON, R. H. 287, 330
HARRIS, L. 185, 203
HARRIS, W. J. 382, 388
HART, E. W. 215, 261, 369
HARTLEY, G. S. 332
HARTLY, D. 237, 262
HARTMANN, W. 102
HASIGUTI, R. R. 51, 69
HAUFFE, K. 71, 72, 78, 80, 82, 83, 84, 85, 87, 88, 89, 91, 93, 96, 97, 99, 101, 102, 103
HAUSNER, H. H. 179, 184, 203
HAYWARD, C. R. 160, 161, 171
HEIDENREICH, R. D. 215, 218, 219, 234, 251, 253, 261 264
HENDERSON, R. 245, 263
HENNIG, W. 174, 203
HENRY, D. E. 124, 169
HERMANN, T. 332

HERRING, C. 199, 200, 203, 264, 277, 314, 328, 330, 332
HERTY, C. H. 111, 168
HEWITT, T. 171
HIBBARD, W. R. 245, 261
HICKMAN, J. W. 93, 102, 103
HIMMLER, W. 85, 103
HIRAI, T. 51, 69
HIRSCH, P. B. 253, 264
HOBSON, J. D. 166, 171
HOBSTETTER, J. N. 363, 388
HOFFMANN, R. E. 198, 203, 284, 285, 286, 291, 293, 294, 303, 316, 330, 331
HOFMANN, W. 168, 356, 388
HOLDEN, A. N. 215, 261, 263
HOLLOMON, J. H. 261, 333, 341, 367, 375, 376, 380, 381, 382, 384, 386, 388
HONEYCOMBE, R. W. K. 211, 244, 247, 249, 261, 264
HOTOP, W. 174, 179, 204
HOUBEN, H. 272, 330
HOWLETT, B. W. 263
HU, HSUN 253, 254, 264, 330
HUBER, R. W. 128, 169
HUGHES, L. A. 181, 200, 201, 204
HUME-ROTHERY, W. 127, 169
HUNTINGDON, H. B. 68, 69, 269, 300, 301, 304, 306, 331, 368
HUTHSTEINER, H. 203
HÜTTIG, G. F. 127, 169, 174, 178, 188

INGERSOLL, L. R. 348, 349, 358, 387
ISENBERG, I. 116
IWASÉ, K. 167

JACKSON, G. 93, 103
JACOBS, R. B. 68
JAEGER, F. M. 147, 170
JAFFEE, R. I. 143, 148, 149, 150, 170
JASWON, M. A. 70, 231, 232, 262
JAUMOT, F. E. 305, 331
JETTE, E. R. 80, 103
JEPSON, D. 167
JEFFRIES, D. 185, 203
JILLSON, D. C. 262
JIMORI, J. 93, 103
JOHANSSON 367, 388
JOHNS, I. B. 167
JOHNSON, F. M. G. 111, 167
JOHNSON, H. H. 172
JOHNSON, R. P. 310, 313, 331
JOHNSON, W. A. 103, 170, 203, 268, 272, 285, 316, 322, 323, 330, 334, 387
JONES, F. O. 178, 182, 201, 203
JONES, F. W. 381, 388
JONES, J. I. 129, 130, 169
JOOS, G. 67
JORDAN, W. A. 203
JOSEFSSON, Å. 123, 169
JOST, W. 73, 74, 103

KAHN, N. A. 171
KALÄHNE, E. 171
KAMM, R. 180, 204
KALISCHER, P. R. 185, 203
KANTROWITZ, A. 339, 342, 387
KARNOP, R. 263
KAUFMANN, A. 134, 170
KAUFMANN, A. R. 172
KAYSER, Z. 167
KÊ, T. S. 26, 36, 40, 47, 48, 56, 66, 68, 69, 118, 119, 120, 121, 122, 169, 264, 292, 293, 298, 331, 332
KEIL, A. 268, 330, 331
KELLAR, J. N. 253, 264
KELLEY, K. K. 147, 170
KEMSLEY, D. S. 262
KERSTEN, M. 70
KHAIKIN, S. E. 359, 388
KHOTKEVICH, V. I. 142, 170
KIEFFER, R. 174, 179, 204
KIKUCHI, R. 174, 204
KING, R. 47, 69, 261, 262, 264
KINGSTON, W. E. 203
KIRKENDALL, E. O. 42, 69 265, 266, 270, 271, 302, 330
KIRKWOOD, J. G. 337, 338, 342, 387
KLING, H. P. 198, 203, 314, 332
KOCH, E. 73, 83, 103, 104
KOCHENDÖRFER, A. 246, 263
KOEHLER, J. S. 54, 67, 124, 169, 230, 248, 261
KONOBEYEVSKI, S. 364, 365, 388
KONTOROVA, T. 212
KOPELMAN, B. 204
KORNETSKI, M. 70
KORNILOW, J. J. 93, 103
KÖSTER, W. 46, 51, 52, 65, 69, 70, 123, 124, 169, 298, 313, 331, 332
KOTTMANN, A. 266, 270, 271, 272, 273, 277, 323, 330, 332
KRASNIKOV, A. I. 169
KRAUSZ, S. 262
KRUMBHAAR, W. 167
KUBASCHEWSKI, O. 146, 167, 170, 331
KUCZYNSKI, G. C. 182, 183, 189, 195, 196, 197, 198, 203, 204, 315, 316. 332
KUHLMANN, D. 230, 233, 249, 252, 262, 263
KULA, E. 123, 169
KULIN, S. A. 380, 388
KURA, J. G. 171

KURDJUMOV, G. P. 379, 388

LACHER 138, 139, 140
LACOMBE, P. 224, 262
LAIDLER, K. J. 331
LALOEUF, A. 247, 263
LA MER, U. K. 333, 355, 386, 387
LANGE, H. 246, 263
LANGE, W. 172
LANGMUIR, I. 330
LARRIS, F. 388
LAURENT, A. 172
LAURENT, P. 185, 204
LAWSON, A. W. 50, 69, 70
LAZAREV, B. G. 142, 170
LAZARUS, D. 68
LEAK, G. M. 263
LE CLAIRE, A. D. 69, 265, 270, 276, 287, 293, 320, 322, 330, 331, 332
LEE, H. 166, 171
LEE, J. K. 97, 101, 103
LEIBFRIED, G. 213, 214, 261
LEMENT, B. S. 380, 388
LEPP, H. 115, 168, 172
LEROUX, J. A. 151, 170
LESCHEN, J. G. 264, 384, 385, 388
LEUTSCHER, H. 172
LEVANTO, K. I. 172
LEVINE, J. 351, 387
LIBSCH, J. F. 180, 184, 185, 204
LICHTMANN, V. I. 262
LIEMPT, J. A. M. VAN 294, 295, 331
LIKHTMAN, V. I. 180, 184
LINDE, J. O. 170
LIONETTI, F. 227, 262
LIPSON, H. 68, 261
LITTLETON, M. J. 73, 74, 103
LLOYD, H. K. 116, 168
LOBLEY, A. G. 167
LOMER, W. M. 220, 227, 228, 240, 249, 252, 253, 254, 261, 262, 290, 331
LORIG, C. H. 172
LOUWERSE, P. 255, 264
LOW, J. R. 263
LUBAHN, J. D. 264
LÜCKE, K. 246 263
LUKENS, H. C. 127, 128, 169, 172
LUMPE, W. 106, 167
LUSTMAN, B. 103
LUWEZ, P. 203
LYMAN, T. 388

MACKENZIE, J. K. 193, 194, 195, 198, 204
MACHLIN, E. S. 261, 378, 379, 383, 388
MAHN, F. 137, 170
MADDIN, R. 245, 261
MADONNA 342, 355, 387
MAKSIMOVA, O. P. 379, 388
MANINTVELD, I. 304, 331
MANN, E. H. 262
MANSFIELD, R. 103
MARINGER, R. 123, 169
MARRYAT, C. 111, 167
MARSHALL, A. L. 167
MARTENS, H. 203
MARTIN, A. B. 332
MARTIN, D. L. 158, 162, 163, 171
MARTIN, S. L. H. 147, 170
MARTIUS, U. M. 217, 261
MARX, J. 54, 67, 124, 169
MASIMA, M. 246, 263
MASING, G. 246, 263
MASLENIKOFF, V. M. 262
MASON, W. P. 261
MATANO, C. 269, 330
MATHEWSON, C. H. 111, 168, 218, 245, 247, 261, 263
MAXWELL, J. C. 5, 6, 15, 67
MAY, W. 253, 264
MCLEAN, D. 264
MCQUILLAN, A. D. 117, 168
MCSKIMIN, H. J. 261
MEHL, R. F. 42, 203, 264, 266, 272, 277, 293, 320, 321, 322, 323, 330, 331, 332, 334, 387
MEIJERING, J. L. 102, 103, 143, 145, 151, 155, 156, 170, 171
MELVIN, M. (formerly AVRAMI). 334, 346
MENDENHALL, C. E. 348, 349, 358, 387
MENTER, J. W. 262
MERICA, P. D. 111, 168
MERWE, J. VAN DER. 73, 102, 103, 226, 230, 262, 371, 374, 388
MEYER, J. 387
MICHEL, A. 80, 102, 131, 132, 133, 135, 169, 170
MICHEL, MME. 169
MILLARD, D. J. 261
MILLER, H. J. 172
MILLER, P. 103
MILLINGTON, W. E. W. 218, 261
MINORSKY, N. 67
MIYAKI, S. 103
MOLENAAR, J. 237, 262
MOORE, W. J. 81, 97, 101, 102, 103
MOREAU, L. 110, 131, 132, 136, 168, 169, 172
MORSE, R. S. 172
MOTT, N. F. 48, 69, 72, 73, 74, 97, 100, 102, 103, 134, 167, 215, 235, 236, 239, 249, 250, 251, 252, 253, 261, 262, 263, 264, 290, 331

Motz, J. 68
Muehlenkamp, G. T. 169
Müller, E. W. 294, 331
Murray, P. 198, 204
Myers, R. H. 179, 204

Nabarro, F. R. N. 69, 189, 190, 204, 213, 214, 217, 226, 227, 231, 233, 235, 261, 262, 264, 314, 330, 369, 370, 371, 384, 385
Nehlep, G. 74, 103
Nelson, B. J. 93, 103
Neményi, P. 262
Neufeld, H. 172
Neumark, W. 357, 387
Nickerson, R. A. 204, 294, 331
Nielsen, H. 68
Nikuradse, A. 167
Nix, F. C. 304, 309, 331, 332
Norton, F. J. 167
Norton, J. 68, 70
Nowick, A. S. 67, 69, 117, 122, 169, 303, 304, 313, 320, 331, 332
Nuckolls, R. 61, 70
Nye, J. F. 227, 228, 252, 253, 254, 262, 290, 330

Ogilvie, G. J. 261
Ohlsson, E. 388
O'Keefe, J. F. 180, 203
Onsager, L. 323, 332
Opie, W. R. 172
Orowan, E. 150, 176, 215, 249 259, 261, 264
Otis, W. 61, 70
Owen, E. A. 129, 130, 138, 139, 169, 170

Paneth, H. R. 302, 331
Paranjpe, V. G. 261, 383, 388
Parker, E. R. 158, 162, 163, 171, 204, 253, 262
Peach, M. 261
Pearson, S. 68
Pearson, T. F. 172
Pell-Walpole, W. T. 172
Peierls, R. 226
Pfeiffer, H. 81, 89, 91, 103
Pfeil, L. B. 164, 171
Phelps, R. T. 93, 102
Phillips, A. J. 69, 111, 143, 168, 170, 172
Phillips, D. J. 262
Pilling, N. B. 72, 96, 103, 161
Pinto, N. P. 179, 184, 203
Pittenger, J. 70
Polder, D. 332
Portevin, A. 131, 132, 133, 169
Postlethwaite, A. W. 182, 204
Potter, E. V. 127, 128, 169, 172
Pound, G. M. 333, 342, 355, 386, 387
Pratt, J. N. 166
Pratt, P. 247, 263
Probstein, R. F. 339, 387
Pry, R. H. 215, 261
Prytherch, W. E. 172
Pschera, K. 93, 103
Pugh, S. F. 180, 203
Pumphrey, W. I. 381, 388

Quarrell, A. G. 93, 102, 103
Quimby, S. 29, 68

Rachinger, W. A. 263
Rahmel, A. 82, 96, 103
Raffelsieper, J. 246, 263
Ramsay, J. A. 263
Ransley, C. E. 116, 159, 160, 161, 162, 164, 168, 171, 172
Randall, R. F. Y. 262
Randall, R. H. 62, 67, 68
Rathenau, G. W. 102, 103
Raub, E. 151, 170
Raub, H. L. 68
Rauscher, W. 313, 332
Rayleigh, Lord 67
Read, T. A. 2, 29, 53, 54, 67, 70
Read, Th. 235
Read, W. T. 210, 222, 224, 225, 252, 260, 262, 289, 330, 368
Rees, A. L. G. 147, 170
Rehbinder, P. 262
Reinacher, G. 70
Reiss, H. 333, 337, 386, 387
Rentschler, H. C. 124, 169
Rhines, F. N. 102, 103, 105, 107, 111, 115, 117, 152, 153, 154, 156, 157, 159, 162, 167, 168, 170, 171, 172, 181, 186, 200, 201, 204, 320, 330, 332
Richards, T. 264
Richardson, J. E. 69
Richter, G. 118, 169, 317, 332
Röhm, F. 246, 263
Rohn, W. J. P. 172
Romeyn, F. C. 90, 104
Röntgen, P. 110, 167
Rooksby, H. P. 110, 167
Roscoe, R. 262
Rose, F. C. 67
Rosenbohm, E. 147, 170
Rosenthal, K. 52, 69
Rötger, H. 61, 70
Rotherham, L. 69, 264, 331
Ruder, R. C. 332
Ruka, R. 102

RUNDLE, R. E. 168, 173
RUSHBROOKE, G. S. 331
RUSSELL, A. S. 167
RYMER, T. R. 188, 204

SACHS, G. 246, 263
SACHSE, H. 115, 168
SACK, H. S. 68, 69
SALZER, F. 102
SANDER, A. 342, 355, 387
SAUERWALD, F. 174, 175, 187, 204
SAVAGE, R. E. 172
SCHAEFER, V. J. 349, 355, 387, 388
SCHAFER, H. 115, 168
SCHEIL, E. 70
SCHLECHT, H. 174, 178, 185, 203, 204
SCHMAHL, N. G. 173
SCHNEIDER, A. 69
SCHOTTKY, W. 73, 74, 97, 99, 100, 102, 104
SCHPIKELMANN, A. J. 93, 103
SCHUBARDT, W. 174, 178, 204
SCHÜLTZE, W. 207, 260
SCHWARTZ, H. A. 172
SCHWARTZBART, H. 263
SCHWED, P. 196, 197, 204, 316, 332
SCRUTTON, R. F. 263
SEAGER, G. C. 331
SEELIGER, R. 106, 167
SEEMAN, A. J. 137, 170
SEITH, W. 146, 170, 266, 268, 270, 271, 272, 273, 277, 323, 330, 331, 332
SEITZ, F. 67, 69, 213, 215, 235, 236, 237, 249, 261, 262, 269, 274, 276, 300, 301, 302, 304, 306, 308, 328, 330, 331
SEYBOLT, A. U. 111, 168, 172
SHALER, A. J. 174, 176, 178, 183, 184, 187, 189, 190, 195, 204
SHIMIZU, S. 137
SHOCKLEY, W. 102, 215, 218, 219, 224, 225, 234, 251, 252, 261, 262, 289, 309, 330, 332
SHUTTLEWORTH, R. 193, 194, 195, 198, 204, 227, 262, 328, 332
SIEFERT, A. V. 69
SIEGEL, B. M. 185, 203
SIEGEL, S. 36, 68
SIEVERTS, A. 115, 132, 139, 140, 170
SILVA, L. C. C. DA 42, 67, 69, 266, 272, 277, 322, 330
SIMS, C. 172
SIMNAD 94, 95, 96, 102
SIMONS, J. H. 112, 114, 168
SKINNER, E. N., JR. 111, 168
SMART, J. S. 143, 170
SMELT, R. 342, 387
SMIGELSKAS, A. D. 69, 265, 266, 270, 302, 330
SMITH, A. A., JR. 143, 170
SMITH, A. D. N. 69, 264, 331
SMITH, C. L. 263
SMITH, C. S. 50, 68, 151, 160, 161, 170, 171, 227, 262, 330
SMITH, D. P. 106, 113, 115, 116, 126, 128, 129, 131, 134, 136, 137, 138, 139, 168, 170
SMITH, G. C. 155, 171
SMITH, R. C. 187, 204
SMITHELLS, C. J. 107, 110, 112, 114, 167, 171, 177, 204, 314, 331
SMOLUCHOWSKI, R. 197, 203, 254, 264, 287, 289, 290, 292, 305, 330, 331, 332, 333, 386
SNAVELY, C. A. 127, 128, 169
SNOEK, J. L. 33, 34, 36, 39, 65, 68, 70, 117, 118, 122, 168, 169, 238, 262, 298, 332
SNYDER, J. A. 148, 149, 150, 170
SOMIGLIANA, C. R. C. 262
SPENKE, E. 97, 101
SPERRY, P. R. 253, 254, 264, 330
SQUIRE, C. 134, 170
STANFORD, E. G. 264
STANLEY, J. K. 69, 298, 319, 320, 332
STARR, A. T. 262
STEACIE, E. W. R. 111, 167
STEARN, A. E. 114, 168
STEIGMAN, J. 309, 332
STEINBERG, M. A. 180, 204
STEINITZ, R. 180, 204
STEPHENS, H. A. 172
STEWART, M. T. 332
STOKES, G. G. 5, 67
STREET, A. C. 111, 157, 168
STROUD, E. G. 104
SWAGERMAN, A. 172
SWIFT, I. H. 69
SUITER, J. W. 263
SULLY, A. H. 264
SYKES, C. 165, 166, 171
SYLVESTROWICZ, W. 263

TAMMANN, G. 72, 104, 145, 170
TAYLOR, H. F. 172
TAYLOR, G. I. 244, 248, 250, 263
TELTOW, J. 73, 83, 84, 104
TENNEVIN, J. 253, 263
THIELE, W. 70
THOMAS, D. E. 156, 171
THOMAS, R. 332
THOMAS, U. B. 97, 102
THOMPSON, F. C. 218, 261
THOMPSON, N. 257, 261, 262, 264

THORP, J. S. 253, 264
TIEDEMA, T. J. 253, 255, 264
TOKICH, J. L. 388
TOLMAN, R. C. 337, 338, 342, 387
TOOLE, F. J. 167
TRAUTZ, O. R. 388
TROIANO, A. R. 388
TUBANDT, C. 73, 104
TURNBULL, D. 198, 203, 253, 263, 283, 284, 285, 286, 291, 293, 294, 303, 316, 330, 333, 334, 338, 341, 343, 344, 351, 352, 355, 357, 366, 367, 375, 376, 381, 387, 388
TURNER, T. 151, 170
TÚRY, P. 263
TYNDALL, E. P. T. 70
TYLECOTE, R. F. 173

UDIN, H. 187, 195, 203, 204
ULBRICH, R. 167
URSPRUNG, A. 387

VAND, V. 260
VEENSTRA, W. A. 147, 170
VERMA, A. R. 260
VERNON, W. H. J. 104
VERÖ, J. A. 172
VERWEY, E. J. W. 90, 104
VIERK, A. L. 85, 103
VOGT, E. 134, 170
VOIGT, W. 5, 6, 67
VOLMER, M. 207, 260, 333, 335, 336, 339, 341, 342, 344, 345, 348, 387
VOLTERRA, R. 180, 184, 204
VONNEGUT, B. 349, 373, 387
VOZDVIZHENSKY, G. S. 169

WAGNER, C. 71, 73, 74, 77, 78, 79, 80, 81, 82, 83, 91, 92, 96, 97, 101, 102, 103, 104, 138, 140, 170
WAIN, H. L. 150, 170, 263
WALL, T. F. 136, 170
WALTENBERG, R. G. 168
WALTHER, H. 68
WARD, L. A. 172
WARREN, D. 185, 204
WASHBURN, J. 253, 262
WATKINS, H. C. 187, 203
WAUCHOPE, K. 332
WEBBER, R. T. 142, 170
WEBER, A. 335, 336, 387
WEBSTER, W. L. 358, 388
WEGEL, R. L. 68
WEGENER, P. 342, 387
WEIBULL, W. 351, 387
WELLS, A. F. 109, 167
WELLS, C. 330
WERT, C. A. 36, 38, 39, 68, 69, 263, 295, 296, 298, 299, 305, 306, 319, 320, 331, 367
WEST, W. A. 122, 169
WHEELER, M. A. 136, 170, 264
WHITE, J. 178, 190, 191, 192, 203
WHITTENBERGER, E. J. 172
WILLIAMS, E. ST. J. 170
WILLIAMS, H. J. 70
WOOTEN, L. A. 168
WORK, R. W. J. 68
WORRELL, F. T. 49, 69
WRIGHT, P. 138, 139, 140, 141, 170
WILLIAMS, J. 198, 204
WOLF, W. 174, 204
WOOD, J. K. 262
WOOD, W. A. 70, 127, 169, 263
WOODRUFF, E. 68
WILMS, G. R. 263
WILSDORF, H. 230
WINEGARD, W. C. 332
WINKLE, D. VAN 68
WULF, J. 174, 178, 180, 184, 187, 189, 195, 203, 204
WYATT, O. H. 257, 258, 264
WYLLIE, M. R. J. 127, 169
WYMAN, L. L. 151, 158, 159, 170, 171
WYSONG, W. S. 97, 102

YENSEN, C. F. D. 117, 169

ZAGRUBSKIJ, A. M. 331
ZAPFFE, C. A. 164, 171
ZAPF, G. 167
ZAVARINE, I. N. 182, 204
ZELDOVICH, J. B. 338, 339, 387
ZENER, C. 2, 20, 38, 39, 40, 41, 50, 53, 56, 61, 62, 67, 68, 69, 70, 117, 124, 168, 169, 262, 295, 296, 298, 299, 301, 305, 306, 309, 310, 312, 313, 331, 332, 367
ZIMENS, K. E. 92, 104

SUBJECT INDEX

ACTION constant of diffusion 44, 305, 311
Activation energy for lattice and grain boundary diffusion 289, 300
— — of self-diffusion 307, 309, 313
— — of sintering 187–9, 192
— — of strain ageing 242, 243
Activity coefficient and diffusion rate 321, 322
Alkali metals, diffusion in 302
Alloying elements, effect on oxidation rate 83 *et seq*
Alloys, mechanism of oxidation 82 *et seq*
Alumina, sintering 184, 192
Aluminium, dislocations, 211, 212
— electrical conductivity 237
— internal friction, effect of annealing 52
— — — effect of cold working 51, 56, 57
— liquid-crystal surface energy 356, 357
— solubility of nitrogen in 110
— work hardening 244, 246
Aluminium alloys, internal friction 48
— — reaction with water vapour 164
Anelasticity 2, 20
Anelastic methods for measuring self-diffusion coefficients 316–20
Antimony, liquid-crystal surface energy 356, 357
Arrhenius equation 32, 295, 305, 306
Austenite, diffusion in 318

BAUSCHINGER effect 233
Beryllium, sintering 179, 184
Beryllium oxide, sintering 184
Bismuth, liquid-crystal surface energy 356, 357
— supercooling 357
Brass, annealing of cold worked 52
— diffusion in 40, 292
— internal friction peak 40
— Kirkendall effect 266–9, 271–4
— work hardening 245
— yielding 240
— yield point 241
Bubbles, growth of 346, 347

CADMIUM, mechanism of oxidation 80
— yield point 241
Carbon, effect on internal friction curves 119–23
Cation defect in semi-conductors (*see* Deficit semi-conductors)
Cerium nitride, structure 109
Chemical diffusion, mechanism 310
— — and self-diffusion, interconnections 322, 323
— potential, determining diffusion rate 321, 322
Chromic oxide, sintering 178
Chromium, effect of gases on lattice spacing 127
— solubility of nitrogen in 108, 109
Cobalt, dislocations 218
— liquid-crystal surface energy 356, 357
— martensitic transformation 382
— solubility of oxygen in 111
Cobalt alloys, diffusion in 304, 305
— — Kirkendall effect 270
Coherent nucleation 371–2, 384
Cohesive force between small particles 176
Cold work internal friction 49 *et seq*
— — — — effect of annealing 51–3
Complex modulus, definition 3, 4
Copper, deoxidation 161, 162
— diffusion in, 320
— effect of dissolved gases on electrical conductivity 143
— ΔM effect 54, 55
— electrical conductivity 237
— internal friction of annealed and cold worked single crystals 53, 54
— internal friction, effect of cold working 50–2
— mechanism of oxidation 80, 96
— pore size after sintering 181
— rational rate constant 81
— self-diffusion 197, 308, 315
— sintering 176–86, 197, 200, 201
— solubility of oxygen in 111
— surface energy 219, 356
— surface tension 187, 188
— water vapour, embrittlement 159–62
— work-function 126
— Young's modulus 51
Copper alloys, creep 257
— — diffusion in 199, 287, 288, 313, 318
— — effect of dissolved gases on electrical conductivity 144, 145
— — elastic after-effect measurements 46
— — internal oxidation 151 *et seq*
— — Kirkendall effect 270–3
— — nucleation 365
Copper-oxygen, lattice spacing 127
Coulomb micropendulum 26
Coupled relaxations 21, 48
Creep, mechanism 255, 314
— rate 256, 257
— curves of metals under small stresses at sintering temperatures 198
Crystal boundary, dislocations 224
— — energy 226, 227
— — stress field 226, 227
— growth 206 *et seq*
Crystallization of supercooled liquids 348–58
Cubic rate law of oxidation 97
Cuprous oxide, electrical conductivity 75, 80
— — lattice defects 74

DAMPING (*see* Internal friction)
— capacity 1, 12
— of free vibrations 10, 11
Defect structures of ionic crystals 73
Deficit semi-conducting layers, formation on metals 80, 89
Deficit semi-conductors, cation defect 75
— — (*see also* Electron defect semi-conductors)
Deformation twins, growth 222, 223
Density of sintered metals 177, 185, 186
Deoxidation of copper 161, 162
Deuterium, effect on magnetic properties of metals 134
Diffusion, action constant 44, 305, 311
— along grain boundaries 281 *et seq*
— application of internal friction measurements 32, 39, 44
— effect of grain size 283 *et seq*
— experimental methods 314 *et seq*
— effect of stresses 279
— governing sintering 196, 197
— heat of activation 32, 40, 297, 298
— in ionic crystals 74, 83
— in iron oxides, mechanism of 95
— in semi-conductors 72, 74, 95
— kinetic theories 328, 329
— mechanism 42, 295 *et seq*
Diffusion of gases in metals 114 *et seq*
— of vacancies 200, 274 *et seq*
Diffusion rate, effect of impurities 293
— thermodynamic theory 323–7
— vacancy mechanism 45, 274–8, 300–5
Dislocation bursts 58
— energy 234
— jogs 234–6, 249–52
Dislocations 205 *et seq*
— and internal friction 56–9
— creation 210 *et seq*
— in body centred cubic crystals 221
— in close packed hexagonal crystals 221
— in crystal boundaries 224–7
— in face centred cubic crystals 218–21
— in slip bands 228–30
— obstruction at grain boundaries 230–32
— sites for nucleation 368, 369
— sources of vacancies 45, 275–7, 303
— speed 213–15
— strain energy 225
Dispersion hardening 155
Domain theory of ferromagnetism 62, 63
Dynamic modulus, amplitude dependence 4
— — definition 3

EDDY-CURRENT method for measuring internal friction 28
Eddy currents 63, 64
Elastic after-effect 233
— — — for determining diffusion rates 318
— — — measurement 319
Elastic modulus 8
Elastic modulus (*contd.*)
— — frequency dependence 16
— range 1
— strain and stress, phase relationships 7
— substance, perfect, definition 4
Electrical conductivity after plastic strain 237
— — of metals, effect of dissolved gases on 137–46
— of semi-conductors, dependence on pressure of negative component 74, 75
— — — effect of alloying elements 85 *et seq*
— of sintered metals 179
— — of titanium dioxide 88
Electrolytic transport in metallic solutions 146
Electron-defect semi-conductors (*see* Deficit semi-conductors)
Electron diffraction, study of oxidation layers by 93
Electrostatic method for measuring internal friction 27, 28
Elongation of titanium containing gases 149
Embrittlement of metals by gases 155–8
Enthalpy of activation of diffusion 297, 298, 306, 307
Entropy of activation of diffusion 297, 298, 306–10
Equivalent linearization method 5
Excess semi-conducting layers, formation on alloys 85 *et seq*
— — — formation on metals 79
Excess semi-conductors 75, 79, 80
EYRING'S theory of reaction rates 295, 296

FRACTURE stresses of liquids 346, 347
Frank–Read sources 206–12, 215, 368
Free energy of crystal boundaries 227
— — of diffusion processes 296–8, 306, 307
— — of formation of embryos 335–8, 343–4, 362–8, 383
— — of solid solutions 360, 361
FRENKEL defects 73, 83

GALLIUM, liquid-crystal surface energy 356
— supercooling 357
Gaseous oxides, effect on oxidation of metals 89, 91
Gases in metals, effect on sintering 178
Gas evolution on solidification of metals 157, 158
Germanium, liquid-crystal surface energy 356
Glass, diffusion in 36, 197
— sintering 183, 192, 197
Gold, creep 189
— liquid-crystal surface energy 356, 357
— self-diffusion 308

Gold (*contd.*)
— sintering 185
— supercooling 349
— surface tension 187–9
— work hardening 245
— under small stresses, creep 198
Gold alloys, diffusion in 322
— — elastic after-effect measurements 46
— — Kirkendall effect 270–3
Grain boundaries 198, 199
— — migration 224, 253–5
— boundary diffusion 280–95, 312
— — diffusion, orientation dependence 287–9
— — energy 369
— — models 290
— — peak 47
— — relaxation 46 *et seq*
— — — activation energy 48
— growth 253–5, 277
— size, effect on diffusion rate 311
Growth of bubbles 346, 347
Growth spirals 209

HAFNIUM, solubility of nitrogen in 108, 109
Hardness of metal-gas solutions 131, 132, 148–54
— of metals containing gases 155
Heat of activation of diffusion 32, 40, 297, 298
— of fusion of metals and interfacial energy liquid-crystal 357
— of sublimation and energy of self-diffusion 307, 309
Heterogeneous nucleation 334–48, 369
High frequency methods for measuring internal friction 26–30
Homogeneous nucleation 334, 347, 355, 363
Hydrides 112
Hydrogen, effect on conductivity of metals 138 *et seq*
— effect on hardness of metals 148
— effect on lattice parameter of metals 127–33
— effect on magnetic properties of metals 134 *et seq*
— embrittlement 158 *et seq*
— solubility in metals 112, 113

IMPERFECT crystals growth, mechanism 207
Imperfect dislocations 216, 217
Impurities, effect on oxidation rate of metals 83 *et seq*
Incoherent nucleation 368–70, 384
Indium alloys, internal friction of twin interfaces 49
Inert gases, solubility and diffusion in metals 106
Inert marker movement 266, 267, 271–4, 322
Interfaces, sources of internal friction 46
Interfacial energies (*see also* Surface energies)
Interfacial energy as a function of the heat of fusion 357
Internal friction 1 *et seq*
— — amplitude dependence 4, 53–5
— — effect of cold working 50 *et seq*
— — discussion in terms of dislocation theory 56–9
— — frequency dependence 6, 16
— — effect of dissolved gases on 117 *et seq*
— — dependence on grain size 40
— — interfaces 46 *et seq*
— — magnetoelasticity 62–5
— — measures 8–12
— — measurements 25 *et seq*, 260, 319, 320
— — method for determining diffusion rates 32, 39, 44, 317, 318
— — peaks 33 *et seq*
— — phenomenological theory: homogeneous strain 2 *et seq*
— — — — non-homogeneous strain 12 *et seq*
— — related to plastic phenomena 49 *et seq*
— — produced by precipitate particles 66
— — effect of surface conditions 55
— — thermoelasticity 60–2
Internal oxidation 102, 151 *et seq*
Interstitial solid solutions, diffusion 295 *et seq*
Interstitial solutions, stress-induced ordering 33 *et seq*
Interstitialcy diffusion 300, 302
Interstial-ion semi-conductor (*see* Excess semi-conductors)
Intrinsic semi-conductors 75
Ionic conducting layers formed on metals 78, 83
Iron, diffusion of carbon in 39
— diffusion of nitrogen in 39
— embrittlement by gases 164, 165
— internal friction, effect of annealing 52
— — — effect of carbon 36–9
— — — effect of cold working 51
— — — effect of nitrogen 33–9
— liquid-crystal surface energy 356, 357
— mechanism of oxidation 80, 93–6
— precipitation of nitrogen and carbon 37, 38
— reaction with sulphur 82, 96
— relaxation phenomena 122, 123
— self-diffusion 292, 308
— sintering 177–80
— solubility of carbon in 38
— solubility of nitrogen in 108, 109, 238
— solubility of oxygen in 111
— yield point 239, 241
Iron alloys, diffusion in 298, 319
— — Kirkendall effect 270
— — martensitic transformations 380, 381
Iron carbide, precipitation 367

Iron-carbon, rate of strain ageing 242
Iron-hydrogen, electrical conductivity 138
— — magnetic properties 136
Iron nitride, binding energy 238
Iron oxides, formation on iron 93–6
— — mechanism of diffusion 95

KIRKENDALL effect 42, 265–79, 302, 312, 313

LANTHANUM nitride, structure 109
Lattice defects, generation by cold work 237
— — in ionic crystals 73, 74
— — in mixed crystals of ionic compounds 83
— — in mixed crystals of semi-conductors 86
— expansion of metals due to the presence of dissolved gases 126–33
Lead, diffusion of metals in 268
— liquid-crystal surface energy 356, 357
— self-diffusion 308
— sintering 175–7
— solubility of oxygen in 112
— supercooling 357
Lead alloys, nucleation 365–7
Linear rate law of oxidation 97
Lithium, dislocations 218
— martensitic transformation 382
Logarithmic rate law of oxidation 92, 97

MAGNESIA, sintering 192
Magnesium, work hardening 244
Magnesium alloys, effect of dissolved gases on electrical conductivity 146
Magnetic method for measuring internal friction 28
Magnetic properties of metals, effect of gases on 133–7
— — of sintered metals 180
Magnetoelastic internal friction 62–5
— static hysteresis 64, 65
Manganese, effect of gases on lattice spacing 127, 128
— liquid-crystal surface energy 356, 357
— solubility of nitrogen in 108, 109
Manganese alloys, internal friction of twin interfaces 49
Manganese-hydrogen, magnetic properties 136
Martensite-type transformations, theory 374–9
Mass action law, application to equilibria involving lattice defects 74, 75, 78, 79, 82–4, 90
Maxwell solid 59
— — internal friction 5–7
Mechanical equation of state 258
— hysteresis in polycrystals 232, 233
Mercury, liquid-crystal surface energy 356
— solidification of small droplets 351, 352, 354
Misfit of oxide films on metals 73
Misfitting monolayer 371
Mobility and diffusion rate 321, 322
Molybdenum, sintering 177, 178
— solubility of nitrogen in 108, 109
— yield point 241
Multilayer formation on oxidation of metals 93

NICKEL, effect of alloying elements on oxidation 90–2
— diffusion in 287, 305, 313
— effect of gases on lattice spacing 129
— liquid-crystal surface energy 356, 357
— sintering 177, 178
Nickel alloys, diffusion in 305, 313
— — internal friction, effect of cold work 51
— — Kirkendall effect 270–3
— — mechanism of oxidation 90–2
— — nucleation 365, 366
Nickel-hydrogen, electrical conductivity 138
Nickel oxide, defect structure 90
— — effect of additions on electrical conductivity 90, 91
Nitride formation of metals 108–10
Nitrogen, effect on hardness of metals 148
— effect on internal friction curves 121–3
— effect on photoelectric activity of metals 125, 126
— solubility in metals 107–10
Niobium, effect of carbon on relaxation phenomena 36
— solubility of nitrogen in 108, 109
Niobium nitride, structure 109
Non-elastic behaviour, principles 3 *et seq*
— strain and stress, phase relationships 7
Notation 201, 202
Nucleation 38, 333 *et seq*
— at grain boundaries 385
— catalysis, mechanism 372–3, 385
— coherent 371
— incoherent 368
— in condensed system, theory 342–6
— in martensite 374 *et seq*
of bubbles 346, 347
— of imperfect crystals 210
— of liquid droplets 340, 341
of liquid from vapour, theory 335 *et seq*
— of a new phase in solids 362–7
— of preferred orientations 385

ORDERING in alloys 30 *et seq*
Oxidation of metals, effect on electrical conductivity 143
— — — mechanism 71–104
— — — WAGNER'S theory 75 *et seq*
— rate, dependence on alloying elements 83 *et seq*
— — dependence on pressure of non-metal 79–82
Oxide films, effect on slip 230

Oxygen, effect on hardness of metals 148
— effect on internal friction curves 119–23
— effect on lattice dimensions of copper 127
— effect on photoelectric inactivity of metals 125, 126
— solubility in metals 110–12, 120

PALLADIUM, effect of gases on lattice spacing 129–33
— liquid-crystal surface energy 356
Palladium alloys, Kirkendall effect 270
Palladium-hydrogen, electrical conductivity 138–40
— — magnetic properties 134, 135
Parabolic rate law of oxidation 72, 97
Perfect dislocations 216, 217
Photoelectric activity, titanium and zirconium 124, 125, 126
Piezoelectric method for measuring internal friction 29
Platinum, electrical conductivity 237
— liquid-crystal surface energy 356, 357
— supercooling 349
Platinum-hydrogen, electrical conductivity 138
Poisson's equation 99
Pole mechanism of deformation twinning 223
Polygonization mechanism 251, 252
Pore size of sintered metals 181–3
Praseodymium nitride, structure 109
Precipitation, thermodynamics 359–61, 364
Pressure of non-metals, effect on reaction with metals 79–82, 96
— of oxygen, effect on electrical conductivity 74, 75
Purity of metal, effect on oxidation rate 80, 88

RATIONAL rate constant 77, 81
Reaction rate theory 295, 296
Recovery, heat of activation 52
— theory 252–9
Recrystallization 255–60, 277
References 67, 102, 167, 203, 260, 330, 386
Relaxation behaviour of metals 117 *et seq*
— strength 15–19, 35, 36, 121
— times, superposition 19, 20
Relaxed modulus 17
Resonance curve, equation 9
Rhodium, supercooling 349
Ring diffusion 301, 302
— — mechanism 308, 309

SCANDIUM nitride, structure 109
Screw dislocations 208, 212, 235
SCHOTTKY mechanism of diffusion 74, 75
Self-diffusion and chemical diffusion, interconnections 322, 323
— anelastic methods of measurement 316–20
— coefficients, estimation 310
Self-diffusion (*contd.*)
— in metals 284–6, 292, 315
— in oxides 81
— mechanism 305–10
— sintering method of measurement 315, 316
Semi-coherent precipitation 371
Semi-conductors, defects in 74, 75
— diffusion in 74, 95
— electrical conductivity 75
— electrical conductivity, effect of alloying elements 85 *et seq*
Sessile dislocation 220, 249
Silver, creep 199
— diffusion in 320
— electrical conductivity 237
— reaction with bromine 78, 83
— self-diffusion 197, 198, 284–6, 303–16
— sintering 174–83, 197
— solubility of oxygen in 111
— surface self-diffusion 294
— surface tension 187–9, 356, 357
— water vapour embrittlement 162, 163
— work hardening 245
Silver alloys, attack by halogens 83–5
— — diffusion in 318, 322
— — effect of dissolved gases on electrical conductivity 144, 145
— — internal friction 43
— — internal oxidation 154–6
— — Kirkendall effect 270–3
— — relaxation effect 40
— — relaxation time 44
— — surface self-diffusion 294
— — vacancy mechanism 45
Silver halides, formation on silver 78, 83
— — internal stresses 228
— — solubility of ions 85 *et seq*
Sintering, effect of stress 186
— mechanism 175 *et seq*
— method for measuring self-diffusion coefficients 315, 316
— surface diffusion 196, 197, 199
— volume diffusion 196 *et seq*
Slip 259, 260
— bands, dislocations 228–30
— in body centred cubic crystals 221, 222
— in close packed hexagonal crystals 221
— in face centred cubic crystals 218–21
Sodium chloride, lattice defects 74
— — work hardening 247
Solidification, size effect 349
— of droplets coated with films 352, 353
— of small droplets 349–55
Solubility in metals
— hydrogen 112, 113
— nitrogen 107–10
— of ions in nickel oxide 91
— of ions in silver halides 85 *et seq*
— oxygen 110–12
Specific heat of metals containing gases 147
Sphere of relaxation 47
Spinels, formation on metals on oxidation 92–4
Standard linear solid 15–19, 31, 45, 47

Static hysteresis 22–4
Steel, effect of hydrogen on mechanical properties 165, 166
Strain ageing 240–3
– — effect of temperature 242
– energy 306
— rate 190
Stress induced ordering, source of internal friction 30 *et seq*
— relaxation across incoherent interfaces 46 *et seq*
Structure of nitrides 109
Subscale 151 *et seq*
Substitutional solutions, anelastic behaviour 41–3
— — diffusion in 300
— — stress-induced ordering 40 *et seq*
Supercooled liquids, crystallization 348
Supercooling of metals, dependence on crystal structure 357, 358
Superheating of solids 359
Supraconductive transition of tantalum and niobium 142
Surface diffusion 280 *et seq*, 316
— — and sintering 196–9
— energies, liquid-crystal 356
— — liquid-vapour 341, 342
— energies of solids, determination 187
— self-diffusion of silver 294

TANTALUM, diffusion of carbon in 40
— effect of carbon on relaxation phenomena 36
— effect of dissolved gases on relaxation phenomena 36
— effect of dissolved gases on electrical conductivity 141, 142
— relaxation phenomena 36, 118–21
— sintering 179
— solubility of nitrogen in 108, 109
— solubility of oxygen in 120
Tantalum alloys, diffusion in 298, 319
Tarnishing reactions 71 *et seq*
Tensile strength, sintered metals 185
— — titanium containing gases 149
Thermal diffusion 61, 62
Thermoelastic internal friction 60–2
Thin film formation 97
Thorium, effect of gases on photoelectric activity 126
— solubility of nitrogen in 108, 109
Tin, liquid-crystal surface energy 356
— self-diffusion 292
— solidification of small droplets 354
— supercooling 349, 357
Tin alloys, nucleation 365–7
Titanium, effect of gases on electrical conductivity 143
— effect of gases on lattice spacing 130, 131
— effect of gases on mechanical properties 148, 149
— effect of gases on photoelectric activity 125, 126
— reactions with gases 117
— solubility of nitrogen in 108, 109
Titanium (*contd.*)
— solubility of oxygen in 112
Titanium alloys, mechanism of oxidation 88, 89
— — effect of gases on mechanical properties 149
Titanium dioxide, effect of additions on electrical conductivity 88, 89
— — defect structure 80, 88, 89
— — sintering 184
Titanium-hydrogen, magnetic properties 134
Titanium nitride, structure 109
Torsion pendulum 25, 26
Transference in metallic solutions 146
Transformation stresses in nuclei 368, 369
Transit nucleation 338, 339
Tungsten, self-diffusion 308
— sintering 177, 178
— solubility of nitrogen in 108, 109
— surface self-diffusion 294
Tungsten alloys, diffusion in 199
Twin interfaces, internal friction 49

UNIT dislocations 216–18
Unrelaxed modulus 17
Uranium, solubility of nitrogen in 108, 109

VACANCIES, annihilation 276–8
— equilibrium concentration 306
— generation 236–8, 275
Vacancy mechanism of diffusion 45, 274–8, 300, 303–5, 308
Vanadium, solubility of nitrogen in 108, 109
Vanadium nitride, structure 109
Vibrational energy, fractional decrease per cycle 11
Viscoelasticity 15, 21, 22, 55
Viscous flow of solids 188 *et seq*
Viscosity equation 6
Voigt solid, internal friction 5

WATER, liquid-crystal surface energy 356, 357
— solidification of small droplets 355
— supercooling 348, 349
— surface energies of droplets 341, 342
— vapour, embrittlement of metals by 158–64
Wolfram (*see* Tungsten)
Work-function of metals, effect of gases 124–6
Work hardening, dependence on temperature 251
— — theory 243 *et seq*
— — of cubic metals 244–8
— — of hexagonal metals 244–8

YIELD point 238–42
— — dependence on temperature 240
— — of metals 150
— — theory 241
Young's modulus 34, 63, 65

ZINC, internal friction of annealed and cold worked single crystals 53
— mechanism of oxidation 79, 98–101
— sintering 174, 175
— work hardening 244
— yield point 150, 241
Zinc alloys, mechanism of oxidation 86–8
Zinc oxide, diffusion in 98
— — electrical conductivity 75, 79
— — effect of alloying elements on electrical conductivity 85, 86
Zinc oxide (*contd.*)
— — lattice defects 75
— — sintering 184
Zirconium, effect of gases on photoelectric activity 124, 125
— solubility of nitrogen in 108, 109
— solubility of oxygen in 112
Zirconium-hydrogen, magnetic properties 134
Zirconium nitride, structure 109
Zirconium-oxygen solutions 147