Chalmers Anniversary Volume (Pergamon, 1981)
[with ERRATA through Dec. 2003]

KINETICS OF NONUNIFORM DEFORMATION*

U. F. Kocks

Materials Science Division, Argonne National Laboratory,
Argonne, Illinois 60439, U.S.A.

To Bruce Chalmers, who gave me a chance.

CONTENTS

*Work supported by the U.S. Department of Energy.

Symbols

L, A	specimen length and cross-sectional area
P; σ	load; true equivalent stress averaged over cross section
$\dot{\epsilon}$,$d\epsilon$,$\Delta\epsilon$	local true equivalent plastic strain rate and strain increments
$\bar{\dot{\epsilon}}$,$d\bar{\epsilon}$,$\Delta\bar{\epsilon}$	same, averaged along bar
Θ, H	strain-hardening rates
β, M	strain-rate sensitivities of the flow stress
δ, δ_P	deceleration parameter and constant stress or load
b, c	its derivatives with respect to strain rate and stress
μ; E	elastic modulus for slip process; for tensile bar
S; $\bar{E}$	machine-plus-specimen compliance; its inverse
v_S; v_d	velocity of sound; dislocation velocity
σ_f, σ_g	friction stress and generation stress for dislocations
σ_L, ϵ_L	Lüders front propagation stress, Lüders strain
v_F; $\dot{\epsilon}_F$	Lüders front velocity; strain rate at the front
f, F	volume fraction deforming or deformed
Y	relative area gradient along bar.

Foreword

This is not a review article. It does not summarize our current understanding in the literature of the many different phenomena associated with nonuniform deformation and unstable flow. It does not emphasize the most pressing current problems or arguments. And most of all, it makes only scant reference to available experimental evidence, though suggestions for needed measurements are occasionally made.

This article is an attempt to answer the question, what material characteristics determine which of the various kinds of nonuniform or unstable deformation are observed, and how rapidly they develop. It is an exposition of a scheme to evaluate and represent relevant material properties, and it lays the groundwork for a prediction of macroscopic behavior under various boundary conditions.

The view presented is a personal one, which has not as yet had the benefit of sufficient wide-spread scrutiny. While many of its component parts are rephrased versions of results absorbed from the literature, there are some novel suggestions. They were developed in large part through discussions with many colleagues; in particular, the contributions by J. L. Duncan, J. D. Embury, J. J. Jonas, R. B. Schwarz, and A. P. L. Turner are gratefully acknowledged.

This work was performed under the auspices of the U.S. Department of Energy and, in part, during a Visiting Professorship at McMaster University in the fall of 1978.

1. Introduction

Nonuniform deformation exerts a disturbing influence in all three areas of concern in the mechanical behavior of materials: testing, structural strength, and forming. The testing of mechanical properties is usually done on tensile specimens or other extended samples: unless the deformation is uniform, the test data do not represent (local) material properties. In structural applications, it is often important to have a margin of safety if the yield stress should indeed be exceeded: wide-spread, well-controlled plasticity is better than local, unstable behavior; and insensitivity to local stress concentrations is desirable. In forming applications, there are usually a number of degrees of freedom, both with respect to some components of strain, and with respect to the locations at which straining actually takes place; even if localized deformation does not lead to failure, it may cause disturbing surface irregularities, e.g. in sheet forming.

Nonuniform deformation in the sense used here is not deliberately caused, such as in plate bending operations; it is a path chosen by the material from all the options allowed by the boundary conditions. Thus, it is itself an expression of material behavior. The questions to be addressed in this article are, what specific material properties are responsible for the various kinds of observed nonuniformities, how they can be measured and (in principle) controlled, and what laws govern the *rate* at which nonuniformities develop.

The specific modes of nonuniform deformation that we shall treat in some detail are Lüders front propagation, jerky flow, and (tensile) necking. In all of them, rate effects play a major role, and there is a basis for useful comparisons. This leaves out a number of other, very important modes; e.g. localized necking in sheets[1–4] and shear band formation in rolling,[5] in which geometric effects predominate; and adiabatic heating,[6] which in many cases only serves to accentuate drastically an already existing tendency for localization. Further, we will not discuss any changes in mechanism, such as they may occur especially as a *consequence* of the localization; e.g. voiding.

With respect to necking in tensile bars and sheets, a considerable amount of work has been done on two aspects of the problem: the *geometric* conditions; and the possible *criteria* that determine whether or not deformation will (or at least may) turn nonuniform.[7–14] Inasmuch as these criteria attempt to provide a yes/no answer under any given set of conditions, they address the question of *stability*. Of wider concern would be an understanding of the *rate* with which nonuniformities develop, and how *severe* they are or may become. This involves material properties in a much more intimate way, and therefore requires that the kinetics of plastic deformation (even when uniform) be described realistically[6, 13] (and not only in steady state).

Nonuniform deformation by Lüders front propagation or by jerky flow, on the other hand, has been treated primarily from a metallurgical standpoint,[15–20] and discussions of the phenomenological conditions that generally may cause such conditions have been scant.[21–23] A major aim of this article is to achieve an integration of the mechanisms with the mechanics, for a variety of modes of nonuniform deformation.

The physical mechanisms of deformation need to be known in some detail, if the aim is to improve some current material properties, or when an extrapolation of the treatment beyond the limits of measurement is necessary. Our aim is more modest: we primarily wish to identify which material properties are even of concern. For such limited scope, a phenomenological description is adequate, provided it is based on realistic general assumptions about the deformation behavior. It does point the way toward which specific processes need to be understood more fully.

A description of the material properties, and a prescription for their measurement, is fraught with particular difficulties under just those conditions that favor nonuniform deformation. In fact, in some cases the properties immediately preceding the development of a nonuniformity are in practice inaccessible, because failure ensues so rapidly. In the particular cases we shall consider, however, localization is either slow or limited. Even here, though, one must carefully distinguish between local material properties, in terms of which the tendency towards localization will be expressed, and the ensuing properties of a macroscopic specimen that does deform nonuniformly.

When nonuniformities develop, different elements of the specimen evolve, almost by definition, along different *paths*. A description that is independent of path is possible only under certain conditions, which are not always satisfied in plasticity. When they are not, it is important that the relevant properties be measured along the relevant path—and that path independence is not presupposed in the entire formalism, as it usually is. For this reason, the definitions of the material parameters and the constitutive relations are discussed at some length. Changes of path in (six-dimensional) stress space, as distinct from changes of path in (equivalent) stress/strain-rate/strain space, are not discussed in detail.[24, 3]

Experiments on the rate of development of nonuniform deformation and on the physical mechanisms associated with it have not been plentiful. There is a significant body of data on Lüders front propagation,[25–28] but it has only rarely been coupled with experiments on the rate sensitivity in uniform deformation.[29, 30] While rate sensitivity has been of considerable concern in experiments on jerky flow,[18, 22, 31–35] the details in space and time of the flow localization have rarely been followed in detail.[36] In the regime of slow necking, the innate hardening (and *its* rate sensitivity) which are, in some cases, the root cause of the relative stability, have not been investigated as systematically as the effects of rate sensitivity of the flow stress.[37–41]

1.1. *Overview*

The plan of the presentation is as follows. The three specific modes of nonuniform deformation, i.e. Lüders front propagation, jerky flow, and necking, will be discussed in section 4. Sections 2 and 3 lay the groundwork common to all of them, first in terms of local properties, then as concerns the interaction of straining elements in a uniaxial test specimen.

In the section 2, we will define material properties relevant to nonuniform deformation, and describe how they can (or cannot) be measured when deformation is in fact nonuniform. A special effort is made to allow for path-dependent evolution. The most appropriate path is the deceleration (or acceleration) at constant stress (or load): the evolutionary property measured in a creep test. Some physical conditions are described under which flow tends to accelerate: negative strain-hardening (either transient or permanent), and negative rate sensitivity. The latter leads to instabilities even on a local basis.

Section 3 deals with the properties of a tensile bar: many *alternative* straining elements under the same load. Their interactions with each other and with the testing system are treated from both the geometric and dynamical points of view. The conditions are discussed under which accelerating deformation of material *elements* leads to a tendency of *specimen* deformation to localize. In a general way, nonuniform deformation is described in terms of strain and strain-rate gradients; the latter is accorded prime significance. The rate of development of these gradients is primarily governed by the (local) acceleration parameter.

In section 4, we treat in some detail the three most important modes of nonuniform deformation: Lüders front propagation, which is associated with a high generation stress for slip; jerky flow, which follows from an intrinsic negative rate sensitivity; and necking which, in its severe form, is a result of an approach to (*local*) steady-state deformation. In the context of necking, we will also investigate the influence of specimen defects on the uniformity of deformation. Nonuniform deformation in tests other than tension will be briefly discussed. It turns out to obey almost the same relations—provided the specimen is *uniaxial* in the sense that it may decompose into alternative straining elements. There are modes of sheet deformation in which this is the case; however, the existence of an added degree of freedom in the dimensionality of stress and strain poses a qualitatively new problem.

The article concludes with a summary, a statement of the results considered new, and an assessment of the current status.

2. Material Behavior

2.1. *Local Variables*

The mechanical behavior of a material is most commonly investigated in a tensile test. This test has a significant disadvantage in that it probes a large number of straining elements simultaneously: only when they all behave identically is a macroscopic measurement of displacement and displacement rate descriptive of the strain and strain rate experienced by *the material*, i.e. any one element of it.

Whether deformation is in fact uniform or nonuniform is itself, however, determined by some material characteristics. It is imperative, then, that one can execute at least a thought experiment in which the necessary parameters describing the behavior of the *material* can be measured, in order to compare these with the parameters describing the behavior of a tensile *specimen.*

Such a thought experiment can be executed as follows. Assume that a particular cross section of a tensile specimen is a material element of uniform properties. Indeed if it were not, one would have to know the average properties over the cross section in order to describe differences in deformation from one cross section to another. Now imagine a way to measure the local area; e.g. by winding a wire around it in a frictionless way and recording any changes in its free length (Fig. 1). The (local) tensile stress σ is then given in terms of the load P on the specimen by

$$\sigma = P/A \tag{1}$$

Changes in the local cross-sectional area measure local changes in tensile strain:

$$d\epsilon = -d\ln A \tag{2}$$

This may be thought of as the relative length change of the element, under volume constancy. Similary, the tensile strain rate is

$$\dot{\epsilon} = -(\partial \ln A/\partial t)_x \tag{3}$$

where t is the time, and "at constant x" is to emphasize that this is measured on a *fixed material element.**

In this way, both σ and $\dot{\epsilon}$ can be measured as functions of time or strain. They can then be used as inputs to a servo-loop that controls the load on the specimen or its length, such as to keep σ = const., or $\dot{\epsilon}$ = const., or follow some other prescription, at the measured cross section.[42, 43] This will work only so long as this cross section keeps deforming: if the entire prescribed displacement rate were supplied by another part of the specimen, or if the load carrying capacity of other elements were exceeded, control would not longer be possible. In addition, the method requires an "instantaneous" response of the servo-system, as well

*Equation (3) is the preferable expression for the connection between strain and area changes, since it explicitly precludes an interpretation of eqn. (2) in terms of *gradients* of strain and area, which are independent of each other, since area gradients can be due to machining defects.

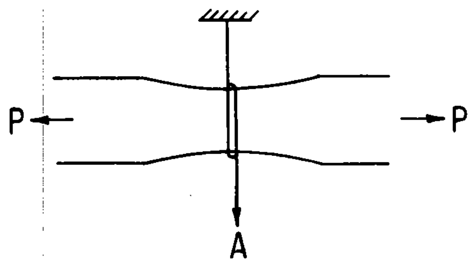

FIG. 1. By measuring the load P and a cross-sectional area A as a function of time, the local stress, strain rate, and strain increment can be determined even when deformation is nonuniform. Through servo-control of P, any of the variables can be prescribed. Their relation describes material properties.

as "instantaneous" communication along the specimen. The actual dynamic response of a tensile bar will therefore have to be considered.

The above discussion outlined an elaborate operational procedure to measure local variables. The *definition* of stress and strain rate should, of course, be independent of the particular measurement technique. In addition, it should hold for compression and when there is more than one component of stress and strain rate. What we mean, in general, by σ and $\dot{\epsilon}$ are the (convected) *equivalent stress and equivalent plastic strain rate*, respectively, so that relations between them characterize material properties.* (Elastic responses will be discussed only in connection with the testing machine, and anelastic effects will be ignored.)

In the following sections, we will define phenomenological parameters measured locally, which will be used to characterize the material, and then discuss in particular the behavior of materials with negative hardening rate or negative rate sensitivity. These regimes are associated with some forms of nonuniform deformation in a tensile specimen; however, they exhibit their most important characteristics already on a local scale.

Throughout this work, we shall regard the temperature to be constant during all tests. The influence of temperature on the material behavior is of course profound, especially with regard to some of the properties to be discussed here. We will often refer to different types of "materials" when we really mean the same material at different temperatures.

The restriction to constant temperature also implies that we shall not consider one important cause for an accentuation of nonuniform deformation: namely, adiabatic heating due to the deformation.[6]

2.2. *Phenomenological Parameters*

A fundamental problem in describing material behavior, especially with regard to mechanical properties, is that the response changes during the test. In physical terms, a material is not fully described unless its entire metallurgical structure is known, and this structure changes with the ongoing deformation. In phenomenological terms, the current properties of a material depend on its current *state*, and this state itself is undergoing evolution.

The conditions that lead to instabilities and flow localization depend very much on the current state. They are different, for example, for a material that is highly work hardened and for the same material in the annealed condition. Therefore, when we attempt to characterize

*When changes in stress or strain path are considered, the evolution may not necessarily be describable in terms of equivalent stress and equivalent plastic strain-rate only.

the "material behavior" or "material properties" that are important in this context, we must mean those in any given state. For example, in a discussion of necking, we imagine the material to have been prestrained to a point just before the region of the stress strain curve that we want to describe. When it is then reloaded, it will display its characteristics such as flow stress, rate sensitivity, etc., as they are relevant to the current state.

The most obvious mechanical properties of a material (in its current state) are[44] its flow stress

$$\sigma = \sigma(\dot{\epsilon};\ \text{state}) \tag{4}$$

or equivalently its initial creep response

$$\dot{\epsilon} = \dot{\epsilon}(\sigma;\ \text{state}) \tag{5}$$

These equations are meaningful only if the "state" is sufficiently specified to make them unique. This may take more than one parameter,[45, 46] but in practice it seems to require only a few—not as many as, say, a complete specification of the dislocation distribution, or a complete specification of the thermo-mechanical history.[44]

The *integral* properties given by eqs. (4) and (5) are, however, not responsible for the development of instabilities: what matters are their *variations*. We define first the two evolution rates: the *strain-hardening rate* Θ in a constant-strain-rate test,

$$\Theta \equiv \left.\frac{\partial\sigma}{\partial\epsilon}\right|_{\dot{\epsilon}} = \Theta(\dot{\epsilon};\text{state}) \tag{6}$$

and the *deceleration parameter* δ in a constant-stress creep test,

$$\delta \equiv -\left.\frac{\partial \ln\dot{\epsilon}}{\partial\epsilon}\right|_{\sigma} = \delta(\sigma;\text{state}). \tag{7}$$

Another important quantity is the ratio of these two parameters:

$$\beta \equiv \frac{\Theta}{\delta} = \left.\frac{\partial\sigma}{\partial \ln\dot{\epsilon}}\right|_{\text{state}}. \tag{8}$$

It is the *rate sensitivity of the flow stress at constant state*, which could have been defined directly from eqs. (4) and (5) (regardless of how many state parameters there are). We have here chosen to define it by the ratio of the two evolution parameters, which explicitly allows for (differential) strain accumulation during the tests. The equivalence of the two definitions follows from two important concepts: first, as was implied in the statement of eqs. (6) and (7), that the (current) *rate* of evolution of the state is also a unique function of the (current) state (and one of the external state variables σ and $\dot{\epsilon}$); and second, that the state does not change in the limit of a vanishing strain increment. (We will discuss changes with *time* at zero strain rate in the next section.)

For some applications, it is more convenient to express the stress in logarithmic form.*

*From a physical point of view, this is less desirable: many flow stress mechanisms are linearly additive,[47] none are multiplicative; and nondimensionality should rather be obtained by dividing the stress by an elastic modulus, especially when different deformation temperatures are encountered.

The parameter H is identical to Hart's[9] γ (previous statements to the contrary[13] notwithstanding); however, M is more general than the parameters m and ν used by Hart.[9, 48] Both H and M are generalized from our previous definitions[13] in that they are valid for an arbitrary number of state parameters.

Then, the straining-hardening rate (eq. (6)) is replaced by the work-hardening rate

$$H \equiv \left.\frac{\partial \ln \sigma}{\partial \epsilon}\right|_{\dot{\epsilon}} = H(\dot{\epsilon}\,;\text{state}) \tag{9}$$

and the dimensionless rate sensitivity is defined, in analogy to eq. (8), by

$$M \equiv \frac{H}{\delta} = \left.\frac{\partial \ln \sigma}{\partial \ln \dot{\epsilon}}\right|_{\text{state}}. \tag{10}$$

The role of strain in all of the above relations deserves some discussion. It appeared in *differential* form only, in the definitions (6), (7), and (9) of the evolution parameters. These definitions are often looked upon as partial derivatives of a *function* $\sigma(\dot{\epsilon},\epsilon)$, which would imply use of strain as a state variable, which it is not. The definitions as given are instead meant to be variations along a prescribed path, without prejudice as to whether an integrable differential equation exists between them.[48] For this reason, the rate sensitivities defined in eqs. (8) and (10) are *not* to be interpreted as partial derivatives taken at constant strain (regardless of how this strain was achieved), but rather as *instantaneous* variations, i.e. at constant state. The question of path dependence is an important one in plasticity, and will be discussed more fully in the next section. Here, we merely wish to ascertain that our definition of the phenomenological parameters did not presuppose path independence in (equivalent) stress/strain-rate/strain space.

Figure 2 illustrates the situation in a three-dimensional plot. The vertical axis is the stress, with an origin of zero. To the left, we plot the strain rate logarithmically, choosing as the origin some reference strain rate $\dot{\epsilon}_1$. To the right, we plot strain as it is accumulated from the present state on. (This is emphasized by calling it ϵ-ϵ_1.) Any test path must proceed in the positive direction along this coordinate; but different tests need not follow different paths *on the same surface.*

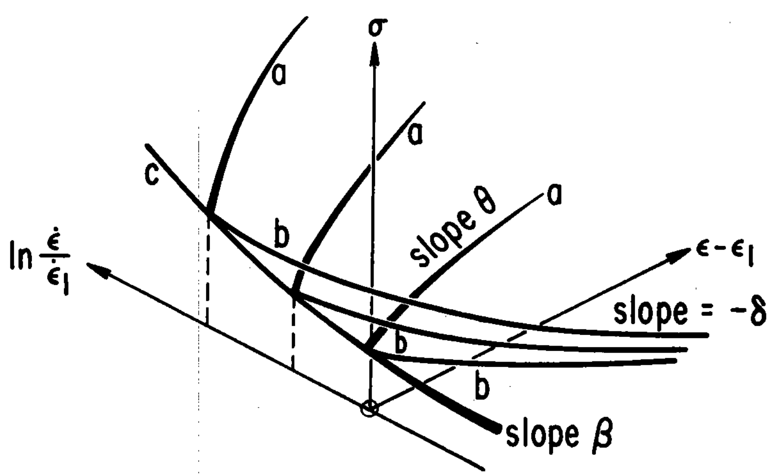

FIG. 2. Material characteristics in a particular state: stress strain curves (a) at different, fixed strain rates (strain-hardening rate Θ); creep curves (b) at different, fixed stresses (deceleration parameter δ); and instantaneous rate sensitivity β (curve *c*). When the surface formed by the *a* set is the same as that formed by the *b* set, strain can be used as a state parameter; generally, they diverge with increasing strain.

The curves labelled a are small portions of stress strain curves at three different, fixed values of strain rate. Their initial slope is $\Theta(\dot{\epsilon})$. The curves labelled b are small portions of creep curves, as they are frequently plotted, as $\ln \dot{\epsilon}$ vs ϵ, for three different, fixed stresses σ. Their initial slope is $-\delta(\sigma)$. The fact that the slope is negative, shows that creep is decelerating; in other words, this material exhibits normal primary creep, in accord with the fact that it is strain hardening. The nonunique nature of the "surface" in Fig. 2 has been emphasized by showing all the creep curves (b) overlaying, rather than intersecting, the stress strain curves (a). (They have deformed at higher strain rates and may therefore be harder.)

Finally, the curve labelled c in Fig. 2 displays the dependence of the initial flow stress (measured at the current state of the material) on the strain rate prescribed—or, *vice versa*, the initial strain rate as a function of a prescribed stress. Its slope is the rate sensitivity β. Note that only a single curve c is shown: for the beginning of the current test. Additional curves c would make sense only if an identical new state could be reached along the different paths, and if one knew the relation between the strain increments required to attain this new state along the different paths.

Looking back at the history *before* the current state was reached, there may, of course, also have been some path dependence. For example, one may have come to the present conditions by having prestrained the specimen at a constant strain rate $\dot{\epsilon}_1$, to a strain ϵ_1. (In this case, continued straining under the same conditions would reproduce the curve a shown in the plane $\dot{\epsilon} = \dot{\epsilon}_1$; an abrupt change to a higher but constant strain rate would produce the other curves a; and a change to creep conditions would be described by the curves b.) It is quite possible, however, that the same state could have been reached by prestraining at a *different* strain rate—only then the amount of prestrain would in general have been different also. It is precisely the advantage of the state parameter description to have eliminated the need for knowing anything at all about the past: its only relevance is in the current state, which reflects itself in the current conditions. This does not mean, however, that one need not know the path to be taken in the *future*, before one can describe the material response over a finite increment of strain.

The path that will be of particular interest in a discussion of local accelerations and of developing differences between different straining elements, is that at constant stress (or load). On the other hand, in many cases the material will have been prestrained at more or less constant strain rate. Thus, a path change is likely to occur as instabilities develop. We propose to deal with this by describing the material in the relevant initial state (e.g. strain hardened), and then describing its properties as relevant to the expected path. In the following, we will first discuss some special cases of path independence, and formulate the constitutive relations in a more general (and, it is believed, novel) way. The physical conditions underlying our insistence on some extra measure of generality in the present context will also be mentioned. Subsequently, diagrams of the kind shown in Fig. 2 will be presented for cases where one or the other material parameter is negative.

2.3. *Constitutive Relations*

The material parameters Θ and δ may be thought of as (reciprocal) coefficients in a differential constitutive relation of the form[48]

$$đ\epsilon = \frac{d\sigma}{\Theta} - \frac{d \ln \dot{\epsilon}}{\delta}. \tag{11}$$

It expresses a smooth variation of strain, stress and strain rate with each other, but does not

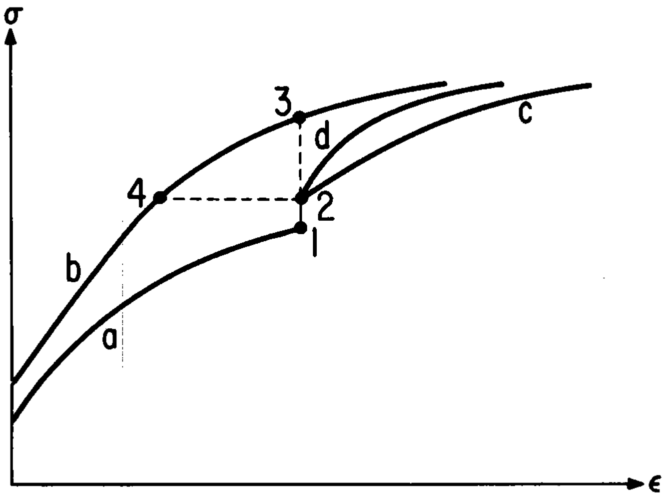

FIG. 3. Two continuous tests under conditions a and b (e.g. at different strain rates), and an abrupt change from a to b conditions at point 1. Typically, the stress jumps to 2 (not 3, as it would if strain were a state parameter), and then follows curve d (not c, as it would if there were any *single* parameter of state).

imply đϵ to be a perfect differential (as is emphasized by using đ instead of d), nor even that eq. (11) is integrable (even if multiplied by a suitable factor). As Hart has shown in an elegant paper,[48] integrability follows when Θ and δ are completely specified by σ and $\dot{\epsilon}$, and do not depend on the state of the material (or its history) in any other way. When this is the case, a *mechanical equation of state* is said to exist,[48] in the sense that the material behavior can be described by equations involving only external variables (even though the strain is generally not one of them).

A specification of both σ and $\dot{\epsilon}$ (rather than just either one of them, as in eqs. (6) and (7)) is equivalent to having specified one of the state parameters, as can be seen from eq. (4). Thus, a sufficiency of specifying σ and $\dot{\epsilon}$, but no further state parameter, is tantamount to a sufficiency of specifying a single parameter to characterize all relevant aspects of the state.[44] This may well be an unrealistic assumption for some important regimes of deformation behavior.[49] Even then, equations of state in the sense of eqs. (4) or (5) would still exist, and would be preferable over any description involving the *history* explicitly. They would require specification of at least one internal variable, not only external variables; however, the evolution coefficients Θ (or H) and δ could serve as macroscopic markers for these internal variables, by virtue of eqs. (6) and (7).

The reason for our insistence on generality is illustrated in Fig. 3. Curves a and b refer to tests along two different, continuous paths, such as at two constant strain rates. When the conditions are abruptly changed from one to the other at point 1, the flow stress will generally change abruptly to a point such as 2. Note that, if the *strain* specified the state, the jump would have to be to point 3. Now if the stress and strain-rate (or *one* appropriate state parameter) were sufficient to specify the state, points 2 and 4 should be equivalent, and curve c should be followed upon continuation of the test. (Curve c is parallel to curve b, shifted in the strain direction.) In reality, the curve followed is often of type d.[49] Thus, Θ is not a unique function of $\dot{\epsilon}$ and σ, and eq. (11) is not integrable. In physical terms, there must be a second structure parameter, which adjusts to its new steady-state value during the transient. The length of this transient is of the order of a few percent, which may be negligible in a consideration of large-strain behavior,[13] but not when momentary instabilities are of concern.

In summary, there are three situations to be distinguished:

(1) The strain can serve as a state parameter (in unidirectional tests of varying strain rates). This requires that strain hardening be independent of strain rate, which is rarely the case, as evidenced by the fact that the rate dependence of steady state (or of the UTS) is usually considerably greater than that of the yield stress. However, *near* one of these limits, the assumption may be sufficiently close to the facts: at small strains (especially at low temperatures[49]); and at very large strains *if* transients are ignored (i.e. if curve d in Fig. 3 is approximated by curve b from point 3 on). Then and only then is the conventional engineering description in terms of stress, strain, and strain rate adequate; in diagrams like Fig. 2, the material characteristics are described by a unique surface. In this article, we will never presuppose these conditions, since much more generality can be gained with little effort by assuming the existence of an equation of state without assuming the strain to be a state parameter.

(2) The material properties—strain-hardening rate Θ, deceleration parameter δ, etc.—are unique functions of stress and strain-rate. Then, nothing further need be known about the state of the material to describe its *differential* constitutive behavior. Integration is possible—but to actually perform it, one needs to know the integrating factor of the strain or, equivalently, the nature of the single state parameter and its evolution. Frequently, it has been assumed[44, 13] that Θ is an integrating factor so that the flow stress can be used as a state parameter. If one knew the state parameter, one could describe the material characteristic by a unique surface in the space of σ, $\dot{\epsilon}$, and this parameter. In stress/strain-rate/strain space, the surfaces are nonunique—but they are all tangent to each other at the current state. Thus, if it is observed that the strain-hardening rate and the deceleration parameter do not change much with strain, path independence can be assumed: either the state does not change much, or the macroscopic behavior is insensitive to it. This is what we shall assume in some applications (Lüders front propagation, jerky flow), in which the total strain of interest is not large. Then, one can use the *differential* properties to gauge macroscopic behavior.

(3) More than one state parameter is necessary as evidenced, for example, by the existence of nonnegligible transients. (The transients are evident only after abrupt changes, but the need for an additional state parameter exists, then, also when the path is smooth.) In this case, there is no way to avoid path dependence, and it is therefore important to describe the material behavior along the expected path. This is why we shall keep the symbol δ as a measure of deceleration/acceleration behavior, and use its equivalence to Θ/β only in special cases.

Equation (11) was written to obtain the strain (if possible) by integration. The material behavior we are interested in describing here is, rather, its (current) rate of evolution. This is more transparent when we reexpress eq. (11) as follows:

$$\boxed{\frac{1}{\Theta}\frac{\mathrm{d}\sigma}{\mathrm{d}\epsilon} - \frac{1}{\delta}\frac{\mathrm{d}\ln\dot{\epsilon}}{\mathrm{d}\epsilon} = 1 \qquad (12)}$$

or equivalently (eq. (9)):

$$\frac{1}{H}\frac{\mathrm{d}\ln\sigma}{\mathrm{d}\epsilon} - \frac{1}{\delta}\frac{\mathrm{d}\ln\dot{\epsilon}}{\mathrm{d}\epsilon} = 1. \qquad (13)$$

The derivatives in these equations are convected: they relate changes in stress and strain rate with progressing strain *in the same material element* to each other. In fact, for some applications,[13] it is even more transparent to express progress by an increase in time rather than strain; e.g.

$$\frac{1}{\Theta}\frac{\partial \sigma}{\partial t}\bigg|_x - \frac{1}{\delta}\frac{\partial \ln \dot{\epsilon}}{\partial t}\bigg|_x = \dot{\epsilon}$$

or

$$\frac{1}{H}\frac{\partial \ln \sigma}{\partial t}\bigg|_x - \frac{1}{\delta}\frac{\partial \ln \dot{\epsilon}}{\partial t}\bigg|_x = \dot{\epsilon}. \tag{15}$$

Whether eqs. (12) or (13) on the one hand, or eqs. (14) and (15) on the other, are more appropriate in any particular case depends on whether the limit that one wishes to include in the description is that of $\dot{\epsilon} \rightarrow 0$ (aging) or $\dot{\epsilon} \rightarrow \infty$ (athermal plasticity[44]).

In eqs. (14) and (15), we have again specified "at constant x", in order to emphasize that the evolution of a particular element is being described. This contrasts with a comparison of neighboring elements in a tensile bar, by taking differences[9, 14] or gradients.[8, 50] Since the different elements in a tensile bar may have undergone different strain increments along different paths, they will generally be in different states; then, eqn. (11) does not hold.

Equations (12) through (15) are various forms of a (differential) *constitutive relation* describing the deformation kinetics of the material. They were inspired by thinking about uniaxial, unidirectional tests. As we have indicated, other tests may be described by the same formalism, if σ and $\dot{\epsilon}$ are identified with the *equivalent* stress and strain-rate, respectively.[51] There is no physical reason, however, why the evolution parameters should be the same in different tensor "directions" even when the equivalent stress is the same.[24] For example, the rate of hardening in a torsion test may be different from that in a tension test. In particular, it is well known that, if the current state was achieved by prestraining, say, in tension, a subsequent test in torsion may exhibit substantially different hardening behavior. In some cases, strain hardening is absent or negative after such a change,[52] which may be due to some instability of the substructure against newly introduced, different types of dislocations.[53–55] In other cases, viz. reverse loading, initial hardening is especially high. Finally, under some conditions, a crystal may decompose into distinct regions some of which are subject to "geometric softening" (by lattice reorientation).[5, 12] Thus, the evolution parameters are *path dependent* in yet another way. This situation has not been extensively investigated, but it may well be of considerable importance in the development of instabilities, especially in biaxially stretched sheets.

These considerations may render the constitutive relation expressed by any one of eqs. (12) through (15) too restrictive to be of any value in tests other than under *radial* (in stress space) and unidirectional loading. Equations (6)–(10) continue to define a set of important evolution parameters, albeit incomplete.

2.4. *Acceleration. Boundary Conditions*

One of the material parameters we have defined (eq. (7)) is the deceleration in a creep test, δ. It is a most important parameter characterizing the behavior of a cross section in a tensile specimen. For when it is negative, deformation accelerates locally (and other cross sections can decrease their contribution to the strain rate). Thus, we shall be looking for conditions under which (eqs. (7)–(10))

$$\boxed{\left.\frac{\partial \ln \dot{\epsilon}}{\partial \epsilon}\right|_{\sigma} \equiv -\delta = -\frac{\Theta}{\beta} = -\frac{H}{M}} \tag{16}$$

is positive. This situation may arise either because strain hardening is negative, or because the rate sensitivity is negative. These two possibilities yield very different results in detail, in part because in the first case δ goes to negative values through zero, in the second case through infinity. We shall discuss a number of specific situations in the next three sections. Under "normal" circumstances, i.e. during strain hardening and primary creep, δ decreases from the order of 10^5 to the order of 10, (as H goes from near 10^3 in early stages of deformation to 1 at the UTS, and M increases from about 0.01 to the order of 0.1).

The word *acceleration* (of deformation) might more properly be assigned to the second derivative of strain with respect to *time*; however, this is directly related to the parameter we have been using, inasmuch as

$$\frac{\mathrm{d} \ln \dot{\epsilon}}{\mathrm{d}\epsilon} = \frac{\ddot{\epsilon}}{\dot{\epsilon}^2}. \tag{17}$$

Thus, at constant stress,

$$\ddot{\epsilon} = -\delta \cdot \dot{\epsilon}^2 \tag{18}$$

When δ is negative, any strain rate is bound to increase.

It must be pointed out that an increase in strain rate, if *possible* within the various boundary and material conditions, is not an *option* open to the material element, but a necessity. It is a direct consequence of the maximization of work dissipation under conditions of constant stress and constant temperature: a well accepted principle of irreversible thermodynamics, equivalent to a maximization of entropy production in a closed system.[56]

The accelerations discussed so far assumed that the stress σ is being kept constant. This may not always be the appropriate boundary condition. However, it is much more realistic than a prescribed constant strain rate (and consideration of softening rather than acceleration as critical material behavior). A constant strain rate cannot always be prescribed, both because of the nature of tensile specimens, and because of the nature of instabilities, as will become evident below.

An interesting boundary condition is that of constant load P. It will be particularly important when we discuss gradients along a tensile specimen. In a local cross section, it is the only variation that can be undertaken without any involvement of the servo-loop (except rapid communication along the specimen). Changes in load can be related to changes in stress by (eq. (1))

$$\frac{\mathrm{d} \ln \sigma}{\mathrm{d}\epsilon} = \frac{\mathrm{d} \ln P}{\mathrm{d}\epsilon} - \frac{\mathrm{d} \ln A}{\mathrm{d}\epsilon}. \tag{19}$$

Thus, for tension (eq. (2)),

$$\left.\frac{\partial \ln \sigma}{\partial \epsilon}\right|_{P} = 1 \quad \text{(tension)} \tag{20}$$

The constant-load condition, therefore, corresponds to a plane whose slope is 1 in $(\ln \sigma) - \epsilon$ subspace (and 0 in the $\dot{\epsilon}$-direction). Figure 4 shows how this plane intersects the (quasi-)sur-

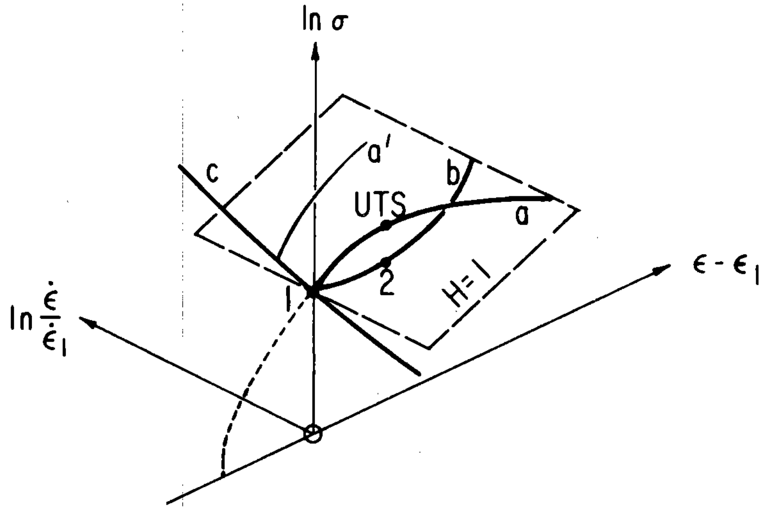

FIG. 4. Material characteristics (as in Fig. 2) being intersected by a "test characteristic", namely for constant load in tension (dashed plane). Curves *a* and *b* lie in the plane and thus follow different paths than in Fig. 2. The dotted line indicates how one might have come to the present state.

face that describes the material characteristics in the current state. This state was chosen such as to make the difference between constant stress and constant load nontrivial; i.e. near the "ultimate tensile strength" (UTS). The dotted curve at $\epsilon - \epsilon_1 < 0$ shows how one might have arrived at point 1. From this state onward, the creep rate at constant load (curve *b*) goes through a minimum at point 2 and then accelerates.

The rate of deceleration at constant load can be expressed quantitatively as (eqs. (16) and (19))*

$$\boxed{\delta_P \equiv - \left.\frac{\partial \ln \dot{\epsilon}}{\partial \epsilon}\right|_P = \delta + \frac{1}{M}\frac{\mathrm{d}\ln A}{\mathrm{d}\epsilon} \qquad (21)}$$

which, for tension (eq. (20)), is[8, 11, 50]

$$\delta_P = \frac{H-1}{M} \quad \text{(tension)}. \tag{22}$$

The difference between δ and δ_P is not important for typical hardening rates except near the UTS. When we describe *material* behavior, we will therefore continue to use δ, and regard all else as boundary conditions, which "intersect" material behavior. In the case of significant path dependence, however, it is δ or δ_P that must be measured, depending on the precise conditions—not Θ and β (or H and M); in other words, the creep test (at either constant load or constant stress) is more appropriate than the tensile test, when they yield different results.

One further boundary condition of interest is that of stress relaxation. It could, in principle, be enforced on a local cross section by keeping the area constant through the

*Hart[9] defined a stability parameter $(\delta \ln \dot{A}/\delta \ln A)_P = \delta_P + 1$. It describes constant load, absolute rather than relative area differences, in part to avoid characterizing a linear-viscous material as "unstable". Note, however, that for $M = 1$, this parameter and our acceleration parameter δ coincide.

servo-control. The more interesting effects are, however, connected with tensile bars, and will be discussed in section 3.3.

2.5. *Substantial Recovery: Negative Hardening*

At high temperatures, recovery and even recrystallization may be occurring simultaneously with the deformation.[57] By themselves, they would cause a decrease of the flow stress with time. During deformation, this decrease can be immediately observed—but only in competition with the counteracting strain hardening. One way to express the concurrent evolution process is to add a recovery term r to the right-hand side of eq. (14) (multiplied through by Θ):

$$\left.\frac{\partial\sigma}{\partial t}\right|_x - \beta \left.\frac{\partial \ln \dot{\epsilon}}{\partial t}\right|_x = h\dot{\epsilon} - r. \tag{23}$$

The strain-hardening-only parameter has here been relabelled h, in order to preserve the definition (eq. (6)) of Θ as the measurable change of flow stress with strain, regardless of what mechanism causes it. This (total, net) strain-hardening rate is therefore now

$$\Theta = h - r/\dot{\epsilon} \tag{24}$$

Equation (24) displays a particular way in which Θ can depend on $\dot{\epsilon}$ (the Bailey–Orowan relation). In fact, the strain-hardening rate depends on strain rate in many situations, notably at high strains or high temperatures. In addition to the strictly time-dependent (static) recovery and recrystallization introduced above, there is dynamic recovery, which is also described by a subtractive term from the (athermal) hardening rate.[49] The strain-rate dependence is then not as severe as indicated in eq. (24), but is nevertheless often important.[58, 59]

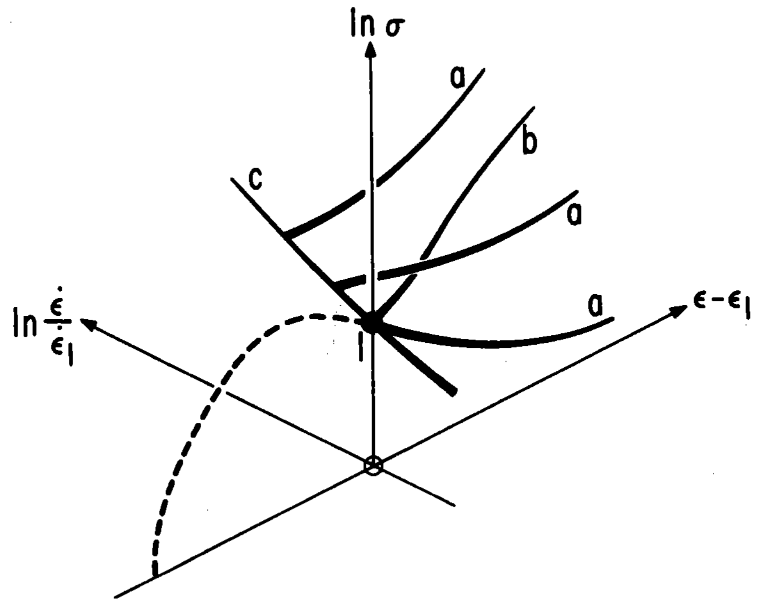

FIG. 5. A state in which hardening is negative; e.g. due to strong simultaneous recovery. The strain-hardening rate is then usually very rate sensitive (curve *a*). Creep accelerates (curve *b*).

When the (static or dynamic) recovery term outweighs the (constant) hardening term, Θ (or H) becomes negative. This situation is illustrated in Fig. 5. As in Fig. 4, the dotted stress strain curve at $\epsilon - \epsilon_1 < 0$ indicates how one might have attained the current state. If one continued with the same constant strain rate $\dot{\epsilon}_1$, one would proceed from point 1 along curve a: under decreasing stress. If, on the other hand, the stress (or the load) is held constant, one follows curve b: at increasing strain rate.

The acceleration shown in Fig. 5 is actually quite low. This is due to the fact that the strain-hardening rate Θ was made strongly strain-rate dependent. Had the curves a been made parallel, for example, the intersection of the "material surface" with the plane $H = 1$ would have curved away to higher strain rates much more rapidly. It was a cardinal point of the recent work by Kocks, Jonas, and Mecking[13] that the rate sensitivity of *strain hardening*, not only of the flow stress, is important for the rate of neck development, which is a consequence of the acceleration discussed here. It will be described in more detail later.

It must be pointed out that not all negative slopes in usual stress–strain curves correspond to local acceleration: this is true only for those in *local* stress strain curves, as they were defined in the foregoing sections. A load elongation curve, or even a true stress vs true strain curve from a tensile test may have a negative slope because of an increasing "active plastic volume", as will be discussed more fully in the next section.

2.6. *Structural Instabilities: Transient Softening*

In materials with small coherent precipitates,[60] or with small dislocation loops created by neutron irradiation,[61] it is often assumed that with the passage of each successive dislocation over the slip plane, the glide resistance decreases. This should give rise to negative strain hardening—until the process is exhausted or can no longer keep up with the simultaneously occurring hardening processes.

The stress strain curve at constant strain rate should then look like curve a in Fig. 6: exhibiting an initial decrease and a subsequent increase in flow stress. This is not what is observed macroscopically in such materials, for a very good reason. As is evident from the figure, an alternative path is open for the deformation: namely, to proceed at constant stress from point 1 along curve b—part of the creep curve. (Its endpoint 2′ is not identical to point 2 when path dependence is important.) We have assumed in this chapter that we can servo-control the local strain rate, and thus follow curve a throughout. However, this may need a really rapid response of the system, to prevent the deformation from accelerating. The actual path followed may lie somewhere in the shaded region, and will probably depend on the relation between the *magnitude* of the negative δ and the system response. When δ is strongly negative, it becomes difficult to prevent local acceleration, and the path followed will be curve b up to point 2′, and only then curve a'. A constant-stress plateau is, in fact, commonly observed in these materials; treatment of the detailed conditions will have to await a discussion of the interaction between alternative straining elements in a test specimen.

Another physical situation showing the same behaviour occurs after a change in loading path. In single crystals that were deformed first on one slip system and then another, it has been observed that the initial strain-hardening rate is sometimes quite low,[52, 54] and that there are indications of local instabilities.[53, 54] In polycrystalline materials, a change in the direction of straining has also been associated with some less-well-specified softening; in fact, it has been speculated that the generally low hardening in torsion tests may be due to this effect.[62]

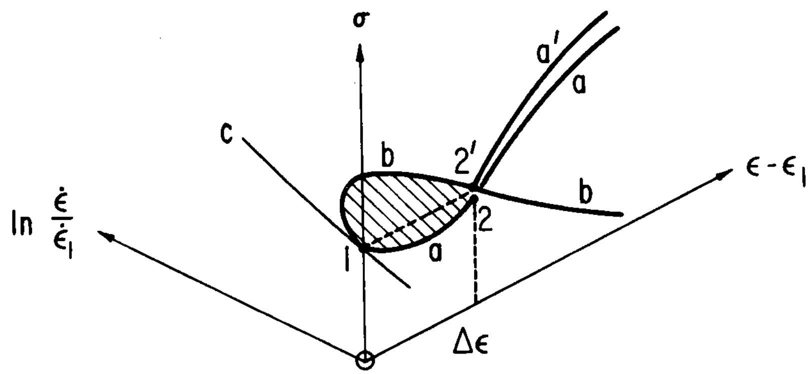

FIG. 6. Transient softening: local stress strain curve (*a*). Creep first accelerates, then decelerates (curve *b*). Unless control is very fast, path *b* will be followed from point 1 to point 2, not path *a*. (Points 2 and 2′ are distinct when deformation is path dependent.)

Finally, "work softening" is observed after a change in strain rate or increase in temperature. We will find in section 3.3 that this cannot be due to a localized structural instability.

A structural instability of the most severe kind occurs after extensive segregation of solute atoms to dislocations during previous aging. Here, one talks of a *generation stress* (to unlock the dislocations[15]) and a *propagation stress* (necessary for continued flow). In Fig. 7, an effective generation stress* is shown as point 2, a solitary point relating to the beginning of the current test. Point 1 is a hypothetical initial propagation stress, assuming that the dislocations were not locked; this propagation stress increases with strain along curve *a*.

In a real test at this (aged) initial state, the stress must first be raised to point 2. Having reached this point, the material will immediately jump to point 3 and start deforming at a rate for which the propagation stress is equal to the generation stress (which, for simplicity was assumed rate-independent—but in any case must be *less* rate sensitive than the propagation stress for the mechanism to work). Since the rate sensitivity of the (propagation) flow stress was assumed positive, there will be such a point 3 at a finite strain rate. From here on, the material could strain at constant stress along curve *b* to point 4′; or the servo-loop could become effective that attempts to hold the strain rate constant, forcing the material back down on curve *a* between points 1 and 4; however, the initial jump from 1 to 3 could not have been prevented in any test. Again, the situation more likely to occur in a tensile test is the sequence 1–2–3–4′, and then along *a*′ if constant strain rate is attempted. We shall show later that the situation described here can lead to Lüders front propagation.

The acceleration behavior exhibited by a material under the conditions described in this section differs from that in the last inasmuch as it is transient: the acceleration occurs only at the beginning of the test and automatically leads into subsequent deceleration—back to the initially attempted strain rate (or lower if creep is considered). Thus, the amount of strain that can accumulate under these conditions is essentially limited. This will be relevant to our discussion of tensile bars.

An important difference between the strain and strain-rate axes can be appreciated in Fig. 7. We have claimed that the material behavior will jump discontinuously from point 2 to

*When there is a generation stress, the material behavior is sensitive to stress concentrations; thus, the observed upper yield point is usually lower than the generation stress, and not very reproducible. (See Fig. 17.)

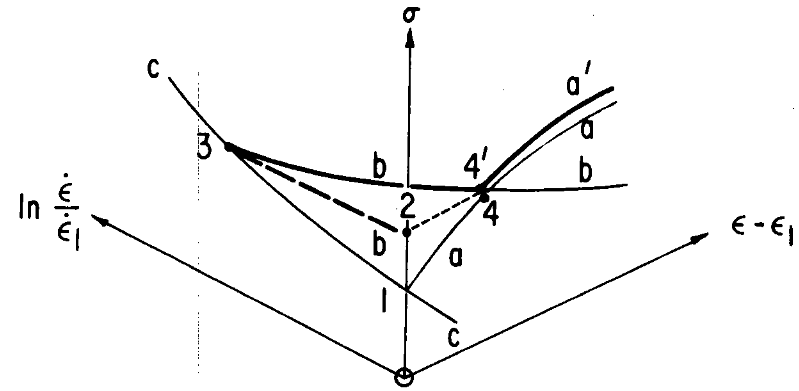

FIG. 7. If the stress was raised to a level (2) above the normal flow stress (point 1), for example because of a difficulty of dislocation generation, the strain rate will jump from point 2 to point 3, and then deformation can follow the decelerating path b to point 4′. At this point, the state is similar to that after uniform deformation (point 4).

point 3: the strain rate can change instantaneously. On the other hand, it is not possible to accumulate strain instantaneously at a finite strain rate; thus, a direct jump from point 2 to point 4 is impossible. This shows again that strain rate is a state variable, whereas strain has meaning only in incremental form.

2.7. *Dynamic Strain-aging: Negative Rate-sensitivity*

A negative rate sensitivity can arise in solution hardened alloys by *dynamic strain-aging*: the glide resistance increases with the time available for diffusion (or other rearrangement) of the solutes to or near the dislocations. The details of the proposed mechanisms vary and may in fact be different for different classes of materials. For illustration, we will use one based on the *jerky glide*(56) of dislocations: their stop-and-go motion within each slip plane, under the action of thermal activation. In this case, the macroscopic strain rate is described by

$$\dot{\epsilon} = \epsilon_0/t_w \tag{25}$$

where ϵ_0 and t_w are appropriate averages, respectively, of the strain increment after a successful activation and the waiting time between "jerks".

Figure 8 shows, in the curve labelled ACTIVATION, how the waiting time decreases as the stress is raised. In fact, there is a stress at which no thermal activation is needed for long-range slip. This *mechanical threshold*(56) depends on the density of obstacles and on the strength of their interaction with the mobile dislocations. Aging can therefore raise the mechanical threshold in the case of atomic-sized obstacles, if their mobility is high enough. This is demonstrated by the curve marked AGING in Fig. 8.

It is the contention of the most promising theories of dynamic straining(22, 36, 18, 19) that this aging process can occur during the waiting time of dislocations at relatively strong obstacle configurations. In that case, the scale of the abscissa becomes identical for activation and for aging, and both are linked to the strain rate by eq. (25).* Since the thermally

*Simultaneous static aging, on the other hand, is proportional to the total time elasped: it occurs even when $\dot{\epsilon} = 0$; during straining, it is linked to the strain rate by the *total* strain, not the small quantity ϵ_0. It affects all dislocations, even those generated during previous straining and not now mobile; and it may have a considerably larger saturation strengthening effect.(63)

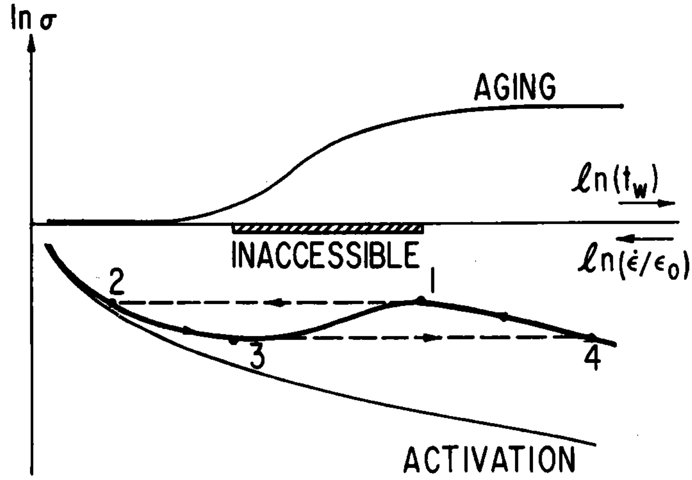

FIG. 8. Thermal ACTIVATION of dislocations past solute atoms, at a given concentration, as a function of applied stress; change of mechanical threshold by increasing the solute concentration on dislocations through AGING. The combination leads to negative rate sensitivity of the flow stress over a certain range (heavy line), and inaccessibility of certain ranges during increases (1–2) and decreases (3–4) of the strain rate.

activated flow stress (σ) is *scaled* by the mechanical threshold (τ), and the mechanical threshold was raised by aging, the ACTIVATION and AGING curves were plotted logarithmically in Fig. 8 and subtracted from each other to obtain the net effect (heavy line).

Under the conditions shown, there is a regime in which aging raises the flow stress more rapidly than thermal activation lowers it: the rate sensitivity is negative (between points 1 and 3 in Fig. 8), when

$$\left.\frac{\partial \ln \sigma}{\partial \ln \dot{\epsilon}}\right|_{\mathrm{ACT}} - \left.\frac{\partial \ln \tau}{\partial \ln t}\right|_{\mathrm{AGE}} < 0 \tag{26}$$

where we have labelled the threshold τ. (A typical time law is[64] $\tau = \tau_0 + Ct^{2/3}$, so that the second term is $\frac{2}{3}\Delta\tau/\tau$.) The region of negative rate sensitivity is, however, experimentally inaccessible in the idealized case discussed here, as has been demonstrated convincingly by Penning.[23] After the stress is raised from zero (sic) to point 1, the strain rate will immediately jump to point 2. If the stress is then immediately lowered, at point 3, the strain rate will jump back to point 4. The regime between points 1 and 3 is avoided. We shall come back to this point later, and only emphasize that *discontinuities* in strain rate should be observed when a *negative* rate sensitivity is expected on the basis of a physical mechanism. One may say that this is due to the fact that the acceleration becomes infinite before it becomes negative, since the rate sensitivity goes through zero (eq. (16)).

We must now incorporate the effects of strain hardening. This is done in Fig. 9. It is seen that the same type of test, in which the stress is prescribed, will now proceed from point 1 to point 2 and thence, accumulating some strain, to point 3′: in the "valley" of minimum stress, for which $\beta = 0$. From here, if the stress is kept constant, the strain rate must jump back to point 4′, to then proceed along the creep curve b. Thus, the strain vs. time curve would have to have kinks in it under these conditions: this is called *jerky flow*—a much more macroscopic phenomenon than the *jerky glide* of dislocations in a slip plane mentioned above.

A test in which one attempts to keep the (local) strain rate constant would show a more complicated response. Let us first assume that this *prescribed* strain rate corresponds to $\dot{\epsilon}_1$, as at point 1. Then, after initial strain hardening to get the stress to point 1, the strain rate would have to jump to point 2. Now, if rapid control is possible, the path can go back to point 3; at constant stress it would go to point 3′. In either case, continued straining can now occur in only two ways: very slowly, by following the creep path from point 4 or 4′ along curve *b*; or still quite rapidly, at increasing stress in the "valley", at $\beta = 0$. The only way the strain rate can be kept *constant* at $\dot{\epsilon}_1$ is in a time-average sense: sometimes "creeping", sometimes "jerking". This is jerky flow.

Note that, neglecting creep, the amount of strain per jerk is limited. It is

$$\boxed{\Delta\epsilon = \frac{\sigma - \sigma_{(\beta=0)}}{\Theta} \equiv \frac{\Delta\sigma_-}{\Theta} \qquad (27)}$$

where $\sigma_{(\beta=0)}$ is the stress at which the rate sensitivity is zero in the initial state (Fig. 9). This would be an interesting value to measure for any quantitative interpretation.

If the "prescribed" strain rate is chosen between those corresponding to points 1 and 3, the forbidden region can be approached, inasmuch as it may be at a maximum in the strain direction. It is in this way that a negative rate sensitivity can in fact be measured: not really instantaneously, but comparing approaches at different $\dot{\epsilon}$ to the jerky-flow condition (in the same state).[34, 65]

Finally, it must be emphasized that the valley indicated in Fig. 9 may not lie parallel to the strain axis. It often moves "to the right" in the coordinate system we have chosen to display Fig. 9. This leads to the observation that jerky flow is often observed only after some critical strain has been attained. This could well be due to an increase in the mobile dislocation density with strain, which enters into the scaling factor ϵ_0 between time and

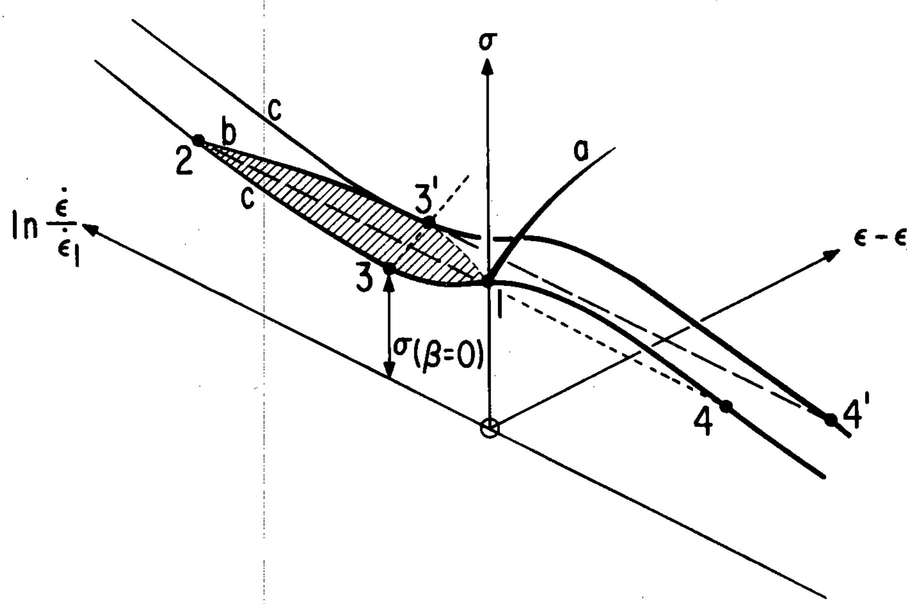

FIG. 9. Negative rate sensitivity at constant state (between points 1 and 3), combined with positive strain hardening (points 3–3′) leads to jerky flow (point 1–point 2) of limited strain (to point 3′), followed by very slow creep (starting at point 4′).

strain rate (eq. (25)).[18, 19, 66] Additional effects are expected due to the generation of vacancies during plastic flow,[19] and due to an influence of solution hardening on strain hardening.[34]

3. The Tensile Bar

3.1. *Strain Gradients and Strain-rate Gradients*

Nonuniform deformation may be described by (infinitesimal) *differences* in behavior between two discrete volume elements, or by *gradients* along the bar. For example, the strain gradient is

$$\epsilon' \equiv \left.\frac{\partial\epsilon}{\partial x}\right|_t = \epsilon'(x, t). \tag{28}$$

Here, x was introduced as the coodinate along the bar. We shall generally imply that the location relative to the origin of x was chosen such that "localization" is characterized by $\epsilon' > 0$.

What is of foremost interest, however, is not the strain gradient itself, but whether it will accentuate or attenuate. Its change with time at any one cross section is

$$\left.\frac{\partial\epsilon'}{\partial t}\right|_x = \left.\frac{\partial\dot{\epsilon}}{\partial x}\right|_t \equiv \dot{\epsilon}'. \tag{29}$$

In words: it is the strain-*rate* gradient that controls whether strain gradients increase or decrease; whether nonuniformities tend to grow or disappear.[8]

A strain-rate gradient may, in fact, have been the cause of developing nonuniformities in the first place.[11, 13] For example, it could be the result of a gradient in applied stress (itself due to a *machining defect*,[50] or to some stress concentrator). At constant stress, a strain-rate gradient could be the result of some heterogeneity in the material properties, such as the glide resistance.[56] More generally, differences in strain rate can arise by all sorts of fluctuations.* As we have seen in the last chapter, the strain rate is a quantity that can change instantaneously (as distinct from the strain); it may be forced to change abruptly by a negative rate sensitivity.

The most important quantity needed to describe the development of nonuniform deformation is, in fact, the next derivative: the *change* in the strain-rate gradient.[13] This is easily seen as follows. If the strain-rate gradient decreased fast enough, the strain gradient would stop growing: the nonuniformity would reach some saturation. On the other hand, when the strain-rate gradient itself increases (or stays constant), then the growth continues and may be catastrophic. The change in the strain-rate gradient is equal to the gradient in the strain acceleration:

$$\left.\frac{\partial\dot{\epsilon}'}{\partial t}\right|_x = \left.\frac{\partial\ddot{\epsilon}}{\partial x}\right|_t \equiv \ddot{\epsilon}'. \tag{30}$$

Using eqs. (17) and (21), this is

$$\ddot{\epsilon}' = -\frac{\partial}{\partial x}(\delta_P \cdot \dot{\epsilon}^2). \tag{31}$$

*The equation of state, such as eq. (4), characterizes only the *average* behavior of the volume elements.

Here, we have incorporated the equilibrium condition along the bar, namely that the load (rather than the stress, eq. (18)) is constant. For *constant* δ_P, we would get

$$\ddot{\epsilon}' = -(2\,\delta_P\,\dot{\epsilon})\,\dot{\epsilon}' \qquad (\delta'_P = 0). \tag{32}$$

The important result is that the sign of δ_P determines whether localization is bounded or not. When it is positive (i.e. the slope in a $\ln\dot{\epsilon}$ vs. ϵ diagram at constant load is negative), strain-rate gradients decrease and strain gradients may saturate; when it is negative, local deformation accelerates and deformation tends to become more and more nonuniform.

For a quantitative description, we will first execute the differentiation of eq. (31) and divide by $\dot{\epsilon}^2$:

$$\frac{\ddot{\epsilon}'}{\dot{\epsilon}^2} + 2\delta_P\,\frac{\dot{\epsilon}'}{\dot{\epsilon}} = -\,\delta'_P.$$

This can be simplified by introducing the gradient in the *logarithm* of the strain rate, which is also physically more meaningful:

$$\boxed{\frac{\partial}{\partial\epsilon}\,\frac{\partial\ln\dot{\epsilon}}{\partial x} + \delta_P\,\frac{\partial\ln\dot{\epsilon}}{\partial x} = -\,\delta'_P.} \tag{33}$$

Note that "progress" is here more conveniently described by an increase in strain (at that cross section), rather than an increase in time.

We will use eq. (33) in detail in our discussion of necking in a later chapter; in fact, it was originally introduced for this purpose, in a slightly different form.[13] For a first analysis,

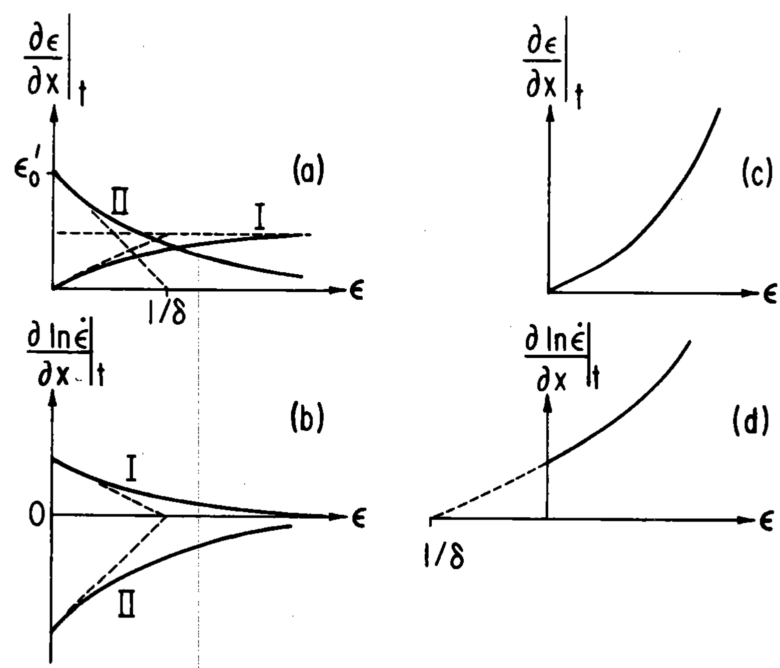

FIG. 10. Development of strain gradient and relative strain-rate gradient with strain for three cases: (a) and (b), positive deceleration ($\delta > 0$); (c) and (d), accelerating deformation ($\delta < 0$). Curves I are for positive strain-rate gradient, such as in a geometric or physical defect or fluctuation; curves II for negative strain-rate gradient as in a deformation defect.

however, let us here assume that δ_P is constant during the localization process; this is reasonable for small strain increments and (note!) when strain-*rate* differences are small. (See the discussion on negative rate sensitivity in the previous chapter.) In that case, eq. (33) has a ready solution:

$$\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t = \left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_0 \cdot \exp(-\delta_P \epsilon) \quad (\delta'_P = 0). \tag{34}$$

Evidently, strain-rate gradients attenuate with a strain constant $1/\delta_p$—or increase exponentially when it is negative. The strain gradient itself can be obtained by noting that

$$\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t = \frac{1}{\dot{\epsilon}}\frac{\partial}{\partial t}\frac{\partial \epsilon}{\partial x} = \left.\frac{\partial \epsilon'}{\partial \epsilon}\right|_x \tag{35}$$

which gives, integrating eq. (34):

$$\epsilon' = \epsilon'_0 + \frac{1}{\delta_P} \cdot \left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_0 \cdot [1 - \exp(-\delta_P \epsilon)] \quad (\delta'_P = 0). \tag{36}$$

The saturation in strain gradient for constant positive δ_P can here be seen quantitatively.

Figure 10 displays eqs. (34) and (36) graphically for two special values of the initial strain gradient and strain-rate gradient that were introduced as integration constants. (Others will be discussed in section 4.3.) Figs. 10(a) and (b) assume $\delta_P > 0$, i.e. decelerating creep: an initial positive strain-rate gradient (curve I in (b)) leads to an increasing but saturating strain gradient (curve I in (a), for $\epsilon'_0 = 0$). Similarly, when the initial strain-rate gradient is negative (curves II), any initial strain gradient *decreases* to a saturation value (usually zero). A negative value of δ_P, on the other hand (Figs. 10(c) and (d)), makes both the strain-rate gradient and the strain gradient increase without limit.

Figure 10 is instructive in that it shows how the *initial* behavior (at zero strain in the current test) provides very incomplete information; for example, the curves labelled I in Fig. 10(a) would indicate *instability*[9] in the current state—yet it is a very controlled kind of instability. (We have avoided using the word "instability" for this reason among others and prefer to describe the various kinds of behavior in more detail.)

3.2. *The Creep Test*

In the last section, we have described an elaborate procedure that would, in principle, allow the measurement of local strain-rates and strain increments. Common tests on tensile (or compressive) bars measure *average* values of these variables:

$$\bar{\dot{\epsilon}} \equiv \frac{1}{L}\int_0^L \dot{\epsilon}\, \mathrm{d}x \tag{37}$$

for the strain rate, and

$$\Delta\bar{\epsilon} \equiv \frac{1}{L}\int_0^L \Delta\epsilon\, \mathrm{d}x \tag{38}$$

for the strain increment. Only when deformation is uniform is

$$\left.\frac{\partial \ln \dot{\bar{\epsilon}}}{\partial \bar{\epsilon}}\right|_P = -\delta_P \quad \text{(uniform deformation)}. \tag{39}$$

Translating this into the actually measured quantities

$$\dot{L} = L\dot{\bar{\epsilon}} \quad \text{(creep)*} \tag{40}$$

and

$$\mathrm{d}\ln L = \mathrm{d}\bar{\epsilon} \tag{41}$$

we get, for uniform deformation,[9]

$$\left.\frac{\partial \ln \dot{L}}{\partial \ln L}\right|_P = \left.\frac{\partial \ln \dot{\bar{\epsilon}}}{\partial \bar{\epsilon}}\right|_P + 1 = -\delta_P + 1. \tag{42}$$

The above relations were written in terms of the *current* length L of the specimen, and thus refer to "true" strain rates and strains. These are nevertheless average quantities that reflect *material properties* only when the deformation is uniform. Similarly, the "true" stress may be measured, or kept constant, by using the current cross-sectional area (eqs. (1) and (2)), assuming this to be uniform. Only then, δ may be inserted for every δ_P above, for constant-stress creep tests, and only then can one use the test characteristic described by eq. (19) to intersect the material characteristic, as was done for local deformation in Fig. 4.

What does one measure when the deformation is nonuniform? We shall discuss a number of specific cases in the sections to follow, but will give a general physical discussion here. Assume, for example, that there is some variation in glide resistance along the sample at time zero, but that the cross section is uniform (and there are no stress concentrators). Now raise the load slowly from zero. The first detectable plastic deformation will occur where the glide resistance is lowest. If there were no (or negative) hardening in this locality, deformation would accelerate here—and would never start anywhere else unless the stress is raised some more. Conversely, when $\delta_P > 0$, the local deformation will tend to cease. When one raises the stress, it will go again—but so it may at other locations: $\delta_P > 0$ tends to spread the flow.

The flow stress mechanism implied by the illustration above was one in which a fixed (athermal) threshold must be overcome before any deformation can be observed. In most cases, there is a more gradual relation between flow stress and strain rate, as in fact we have implied by using a finite rate sensitivity. Then, deformation does not really *start* at one locality; it is only more rapid here. Thus, when it decelerates, the strain rate may become more comparable to that elsewhere, and deformation will become more uniform. We will now give a detailed description of how one can calculate this process quantitatively.

Let us assume that the initial strain-rate distribution $\dot{\epsilon}(x)$ is known. Then the local acceleration is (eq. (18), at constant load):

$$\ddot{\epsilon}(x) = -\dot{\epsilon}^2(x)\cdot\delta_P. \tag{43}$$

By integrating twice over a certain time increment Δt, one can calculate the new distribution $\dot{\epsilon}(x)$ (for the next step), and the local strain increments $\Delta\epsilon(x)$; integrating these over x

*This equation holds for creep tests only, not in a hard machine (for which see the next section). For a general definition of $\dot{\bar{\epsilon}}$ we retain eq. (37).

gives the change in length, ΔL, according to eqs. (38) and (41). Finally, $\dot{L} = \Delta L/\Delta t$. The change in local stress follows from $\Delta\epsilon$ (eqs. (1) and (2)). The procedure is complicated by the fact that the calculation must be done in constant *time* increments along the bar, whereas the local material properties change more nearly with the local *strain* increment. For this reason, it is not instructive to perform one of these triple integrations in closed form. In a numerical integration, it is of course easy, in addition, to incorporate any known dependence of δ_P on strain rate and stress.

We have seen that it is only the *sign* of δ_P that determines whether flow will localize or spread, not any dependence of δ_P on strain rate or stress. A negative value of the deceleration parameter can, however, have very different effects depending on whether it is due to negative hardening or negative rate sensitivity. In the case of negative hardening, the macroscopic creep curve should, in fact, follow the microscopic curves in a qualitative way (curves *b* in Figs. 5 and 6). Negative rate sensitivity, on the other hand, will lead to jerky flow on a macroscopic as well as on the local scale—only that now successive jerks may come from different regions before the same locality deforms again. Macroscopic strain-time curves thus should have kinks in them when $\beta < 0$ (or $M < 0$).

3.3. *Constant Extension Rate and Stress Relaxation*

We now turn to *hard* testing machines, in which there is an interaction between load and extension: a change in load causes elastic changes in length of both the specimen and the components of the machine that are between the points at which length is controlled or measured. The total change in length is therefore

$$\dot{L} = \int_0^L \dot{\epsilon}\,\mathrm{d}x + S\dot{P} \tag{44}$$

where S is the actually measured slope in a P vs. L diagram in the elastic regime. (Recall that $\dot{\epsilon}$ was defined as the *plastic* strain rate.) It is the combined compliance of specimen and machine:

$$S = L/AE + 1/M_M \tag{45}$$

where E is the Young's modulus of the specimen and M_M is the machine stiffness. The machine compliance is typically about 5 times the specimen compliance in long tensile specimens, but can be much larger by comparison in compression tests. In any case, the change in length represented by the second term in eq. (44) is generally small compared to the first except in stress relaxation tests.

For typical stress strain tests, in which $\dot{L} = \text{const.}$ is prescribed, we can therefore use eq. (40) rather than eq. (44) and write

$$\left.\frac{\mathrm{k}\ln\dot{\bar{\epsilon}}}{\partial\bar{\epsilon}}\right|_{\dot{L}} = -1. \tag{46}$$

When deformation is uniform, this geometric condition represents a plane in $\ln\dot{\epsilon} - \epsilon$ subspace, which may be thought of as intersecting the material characteristic. This construction is shown dashed in Fig. 11. It is of the same kind as the cut at constant load shown before in Fig. 4 and expressed analytically by eq. (19).

For the stress relaxation, we must set $\dot{L} = 0$ in eq. (44). Expressing load changes by stress changes (since dimensional changes can here be neglected), we can write

$$\dot{\bar{\epsilon}} = -\dot{\sigma}/\bar{E} \tag{47}$$

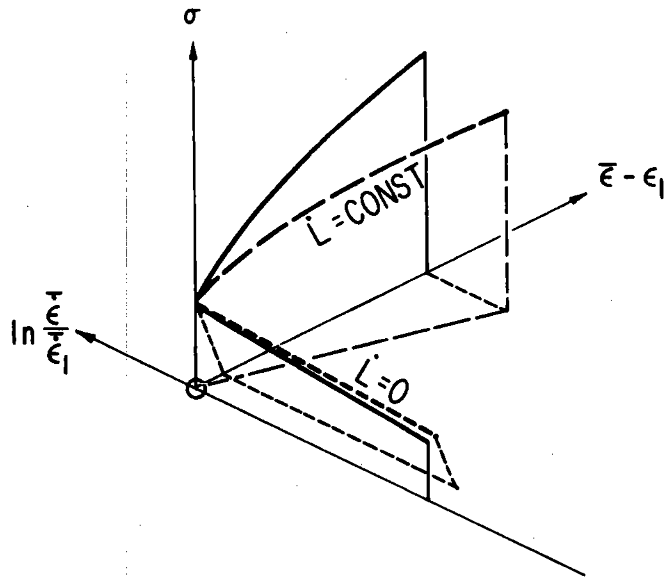

FIG. 11. Test characteristics for constant extension rate $\dot{L}$ (dashed) and for stress relaxation (dotted). The machine response is negligible in the first case, geometric changes are in the second.

where we have used an effective Young's modulus defined by

$$1/\bar{E} \equiv SA/L. \tag{48}$$

This machine characteristic can be represented as a plane in $\sigma - \epsilon$ subspace:

$$\left.\frac{\partial \sigma}{\partial \epsilon}\right|_L = -\bar{E}. \tag{49}$$

Its intersections with the plane $\dot{\epsilon} = \dot{\epsilon}_1$ and with the material characteristic are shown dotted in Fig. 11. Since $\Theta \ll \bar{E}$ under all common circumstances, the plane is very steeply inclined in the figure: stress relaxation measures properties virtually at constant state.[67, 68] (This is not necessarily true when time rather than strain effects are important, such as static aging or recovery.[69]) The slope of the stress relaxation curve is therefore quite closely related to the strain-rate sensitivity. The quantitative relation follows from eqs. (6)–(8) and (47):

$$\left.\frac{\partial \ln \dot{\epsilon}}{\partial \sigma}\right|_L = \frac{1}{\beta}(1 + \Theta/\bar{E}). \tag{50}$$

All the above relations concerned *average* strain rates and strains. They give insight into material behavior only when deformation is uniform. Nonuniform deformation exhibits effects in constant-elongation-rate tests and in stress relaxation that are different in some respects from those discussed in the previous section on creep.

When a constant elongation rate (or a constant true *average* strain rate) is prescribed, there is an intimate coupling between accelerating deformation and its localization; and between decelerating deformation and its spreading.[70] Let us reconsider, for example, the case in which a structural instability is supposed to have given rise to transient local softening (Fig. 6, replotted in Fig. 12). The tendency of a local element at point 1 will be to

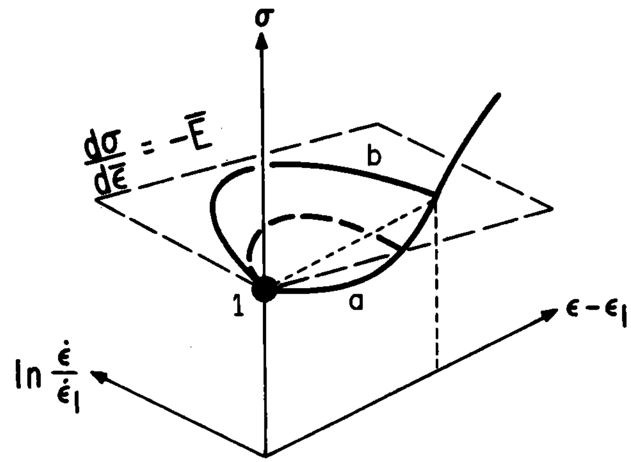

FIG. 12. When local strain hardening is negative, the actual path taken (heavy dashed line) depends on the response of the machine (light dashed plane; characterized by an effective modulus $\bar{E}$ of machine and specimen): when the deforming region is narrow, $d\bar{\epsilon}/d\epsilon \ll 1$ and the stress stays nearly constant (curve *b*).

accelerate, along curve *b*. If *all* elements did this simultaneously, the average strain rate would exceed the prescribed one, and the stress would drop according to eq. (47). As was indicated in Fig. 11, this response is quite "steep"; thus, the applied stress would decrease more rapidly than the plastic resistance plotted in Fig. 12, and curve *a* is followed. If, on the other hand, only one small volume element started its acceleration from point 1, the macroscopic strain, and thus the drop in applied stress, would be negligible, and curve *b* is followed. An intermediate case is illustrated in Fig. 12, where the material characteristic is intersected by a plane that characterizes the effective machine stiffness $\bar{E}$ (eq. (49)) converted to *local* strain, as it is plotted to the right in the figure. If the length of the deforming element is w, $d\bar{\epsilon}/d\epsilon = w/L$. Inasmuch as the local elements act more or less independently of each other, the normally expected path should be curve *b*. When a slow decrease in stress is actually observed macroscopically, this must indicate a very homogeneous material behavior: an insensitivity to fluctuations. Classical *work softening* apparently falls into this category, and can thus not be explained on the basis of a strain-triggered, local structural instability.

Just as negative local hardening cannot be observed in a macroscopic stress strain test, negative local rate sensitivity cannot be observed in a stress relaxation test. What should in fact happen can be derived by referring to Fig. 9. There we have seen that steady deformation (when the prescribed strain rate is in a region of negative rate sensitivity) is possible only in the "valley" corresponding to zero rate sensitivity. Since this is at a higher strain rate than $\bar{\dot{\epsilon}}$, it must be localized. If the cross-head motion is now stopped, only the regions that are actually deforming will contribute to the stress relaxation—but these can lower their strain rate only discontinuously (from point 3′ to point 4′ in Fig. 9). Thus, the initial strain rate derived from a stress-relaxation plot must here be lower than the prescribed strain rate in the previous test. This is in contradistinction to all usual cases where there is continuity of plastic strain rates as the machine is switched off.[71] Thus, the kinks that would be observed in creep curves, should be compressed into a single one located before the *beginning* of a stress-relaxation curve.

An *apparent* negative rate sensitivity might be observed in stress relaxation tests in a *very* soft machine: when $\bar{E} < |\Theta|$ and $\Theta < 0$ (eq. (50)). Then, the stress vs. time plot will be convex upward (over the initial region in which $\Theta < 0$). This would be interpreted as a negative rate sensitivity if the correction term in eq. (50) were neglected. The condition $\bar{E} = |\Theta|$ may be considered the demarcation line between hard (tensile) and soft (creep) machines.

Localization seems unlikely during stress relaxation in a hard machine, since the stress decreases so rapidly. For a quantitative calculation (when $\beta = M\sigma > 0$), one must proceed along precisely the same cumbersome route as for any finite prescribed $\dot{L}$, as was discussed above.

A quantitative calculation of the spreading or localization again demands a triple integration over time and length, preferably performed numerically as outlined in detail in the preceding section. The only change due to the changed boundary conditions is that now one must introduce an iterative procedure to correct the total extension rate (if it came out different from the prescribed $\dot{L}$) by adjusting the load.

3.4. *Dynamic Response to a Local Instability*

In the two preceding sections, we have implicitly assumed that the phenomena occurring on a local scale must be in *equilibrium* with the boundary conditions; more correctly put: that there is *communication* between local events and boundary conditions. The information transfer required does, of course, take a finite time, governed by the acoustic wave propagation speed, v_S. Under usual conditions, this time turns out to be negligible except at strain rates in excess of about $10^2 s^{-1}$. In jerky flow, high strain rates are indeed attained locally; we must investigate whether and how this may contribute to the macroscopic behavior.

An acoustic wave is generated at a certain cross section if it marks the interface between regions of different strain rates. Let us assume that there is a region w wide (*along* the specimen, Fig. 13) that deforms faster than the rest by $\Delta\dot{\epsilon}$. Then it displaces the interface between it and the "bulk" at the rate

$$\Delta v \sim w\,\Delta\dot{\epsilon}. \tag{51}$$

Since we are only concerned with orders of magnitude in this section, we will neglect geometric and various other factors and use the symbol $\sim$ for "proportional to and of the same order".

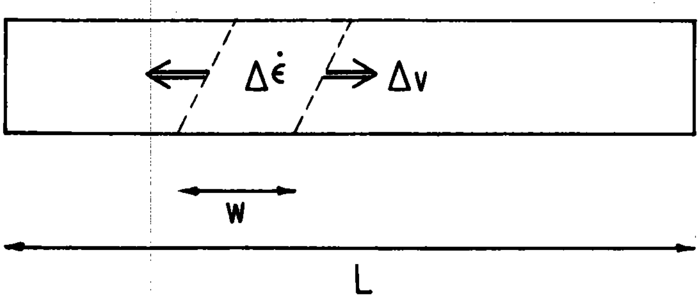

FIG. 13. When a region w wide suddenly increases its strain rate by $\Delta\dot{\epsilon}$, its boundaries are displaced by $\Delta v = w\Delta\dot{\epsilon}$, and acoustic waves are generated.

As a result of the wave travelling away, the stress drops behind it by[72]

$$\Delta\sigma = -\frac{E}{v_S}\Delta v \tag{52}$$

where E is an appropriate modulus. The term E/v_S is the acoustic impedance (usually labelled "ρc", where ρ is the mass density of the material and $c \equiv v_S$). It is apparent from eq. (52) that for noticeable unloading to occur, $\Delta\sigma/E$ and thus $\Delta v/v_S$ must be a substantial fraction of the elastic strain; say, 10^{-4}. We will find that this condition is apparently fulfilled in the regime of jerky flow.

Inserting eq. (51) into eq. (52) gives us another test characteristic:

$$\left.\frac{\partial\sigma}{\partial\dot{\epsilon}}\right|_t = -E\frac{w}{v_S}. \tag{53}$$

It is a line of negative slope in the σ-$\dot{\epsilon}$-plane at the present state ($\epsilon - \epsilon_1 = 0$). To illustrate its effect, we repeat Fig. 8 in Fig. 14, with linear axes and two lines drawn in dashed which have the slope given by eq. (53). They delineate the inaccessible region of the material characteristic. The important result is that for jerky flow to occur, the (true) rate sensitivity of the material must now be *sufficiently* negative. Modifying eq. (26) by eq. (53), the condition reads

$$\boxed{\left.\frac{\partial \ln\sigma}{\partial \ln\dot{\epsilon}}\right|_{\mathrm{ACT}} - \left.\frac{\partial \ln\tau}{\partial \ln t}\right|_{\mathrm{AGE}} < -\frac{E}{\sigma}\cdot\frac{w\bar{\dot{\epsilon}}}{v_S}.} \tag{54}$$

Note that the stress and the strain rate appear on the right-hand side only because we have chosen to express the *left*-hand side by the *logarithmic* derivatives. In particular, the strain rate is the one at which the material characteristic is being diagnosed (say, the average specimen strain rate $\bar{\dot{\epsilon}}$), and not the fast strain rate expected in the jerk.

The condition of negative (total) rate sensitivity is usually approached gradually: the aging term increases with temperature and often with strain, and decreases with strain rate. The point at which jerky flow can begin to occur is reached when all three terms in eq. (54) balance out to zero. The importance of the dynamic response in this balance can be judged by comparing the right-hand side of eq. (54) with either *one* of the terms on the left. The first term, the relative rate sensitivity due to thermal activation is of order 0.01 in most cases. Thus, the influence of the dynamic response on the balance of positive and negative rate sensitivities is trivial when

$$\frac{E}{\sigma}\frac{w\bar{\dot{\epsilon}}}{v_S} \ll 10^{-2}. \tag{55}$$

Since $E/\sigma < 10^3$, $w < L$, and $L/v_S \sim 10^{-5}s^{-1}$ (see eq. (56) below), this condition is always fulfilled for $\bar{\dot{\epsilon}} < 1\ s^{-1}$. In practice, the correction should be trivial even at higher strain rates, since it is in the very nature of the phenomenon that it should be local, i.e. $w \ll L$.

A final consequence of dynamic effects to be considered is this: a finite amount of strain will occur in the unstable region before interaction with the boundary conditions has taken place. This communication is complete after the acoustic wave has travelled to the ends of the specimen (where it is assumed to be reflected by the massive grips) and back: in a time of the order

$$t_D \sim L/v_s \sim O(10^{-5}s) \tag{56}$$

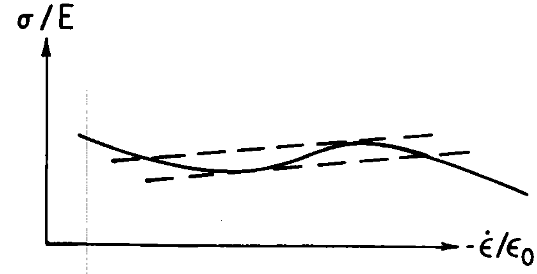

FIG. 14. The dynamic response of the specimen (dashed) alters the inaccessible region of a partially negative stress vs. strain-rate characteristic, and thus alters the criterion for jerky flow. The effect is derived to be negligible under most circumstances.

The extra local strain increment during this time (if the extra local strain rate $\Delta\dot{\epsilon}$ remains approximately constant) is

$$\Delta\epsilon_D \sim \Delta\dot{\epsilon}\cdot t_D \tag{57}$$

The average strain increment in the specimen comes out to be (eqs. (51), (52), (56), (57))

$$\Delta\bar{\epsilon}_D = \frac{w}{L}\,\Delta\epsilon_D = -\Delta\sigma/E. \tag{58}$$

Equation (58) is precisely the same as that for stress relaxation at fixed length—in the absence of dynamic effects. Not surprisingly, then, dynamic effects are negligible at times large compared to t_D (eq. (56)). The stress-relaxation equation (47) given in the last section did differ in one respect: it allowed machine relaxation, so that $\bar{E}$ (eq. (48)) appears instead of E. Clearly, this involves additional time.

It is sometimes supposed[72] that the returning stress wave can *cause* the end of straining during this particular jerk, so that eq. (58) (or one in which E has been replaced by $\bar{E}$) gives the entire strain per jerk. This could be true only if the second local stress drop were to move the material below the flow characteristic, while the first did not. (In the average, this effect would double the dynamic correction on the right-hand side of eq. (54).) The strain per jerk could, alternatively, be limited by strain hardening, as given in eq. (27), which is not limited to be less than the elastic strains.

In conclusion, dynamic effects seem to be of no consequence on the criterion for jerky flow (unless the prescribed strain rate $\bar{\dot{\epsilon}} \gg 1\ s^{-1}$). They may determine the magnitude of the stress drop, in response to the magnitude of the strain-rate jump, if the width of the deforming zone is sufficiently large. Then the strain increment associated with the strain-*rate* discontinuity equals the elastic strain released in the entire specimen during the stress drop.

3.5. *Alternative Straining Elements*

The most important feature of a tensile bar in connection with nonuniform deformation is the fact that it contains alternative straining elements: the average strain and strain rate measured at the ends can be made up of deformation in any or all of them.* In this section, we shall assume that a *fraction* of the volume deforms, the rest does not. Even the further

*Since the *strains* (or compliances) from the different volume elements add, rather than the forces (or resistances), such an arrangement is often referred to as *in parallel*, even though the actual volume elements are in series.

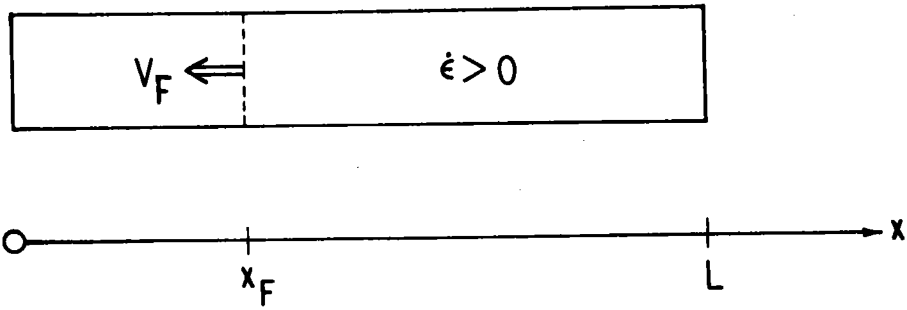

FIG. 15. Increasing the deformed volume fraction: a deformation front travelling at speed v_F into the undeformed region at $x < x_F$.

specialization that the deforming part behaves *uniformly* covers two realistic cases adequately.

One way of splitting up a tensile bar into two regions is to assume, in one part, a constant strain increment $\Delta\epsilon$. This may be an adequate description of the spatial nonuniformities generated by jerky flow (see eq. (27) and Fig. 10). If the fraction of the specimen that *has* deformed is F, the average strain increment is

$$\Delta\bar{\epsilon} = F\,\Delta\epsilon. \tag{59}$$

This is a special form of eq. (38) for $\Delta\epsilon$ independent of x. (Since the deformed regions may not be contiguous, eq. (59) is preferable over eq. (38) with an upper limit set at FL instead of L.) If $\Delta\epsilon$ is assumed to be independent also of time (or accumulated strain), the average strain rate and acceleration are, respectively,

$$\bar{\dot{\epsilon}} = \dot{F}\,\Delta\epsilon \tag{60}$$

and

$$\bar{\ddot{\epsilon}} = \ddot{F}\,\Delta\epsilon. \tag{61}$$

The deformation kinetics is governed by the rate of spreading.

An entirely different way of splitting up a tensile bar into deforming and nondeforming regions is to say that the strain *rate* is zero for $x < x_F$ (Fig. 15). Then the average-strain-rate equation (37) specializes to

$$\bar{\dot{\epsilon}} = \frac{1}{L}\int_{x_F}^{L} \dot{\epsilon}\,\mathrm{d}x \tag{62*}$$

and the average acceleration becomes

$$\bar{\ddot{\epsilon}} = \frac{v_F}{L}\dot{\epsilon}_F + \frac{1}{L}\int_{x_F}^{L} \ddot{\epsilon}\,\mathrm{d}x\,. \tag{63}$$

Here, we have introduced the *front velocity* $v_F = -\dot{x}_F$, and the strain rate at the front, $\dot{\epsilon}_F$ (to be exact: the strain-rate *discontinuity* at the front).

The appearance of two terms in eq. (63) indicates that it is possible to have steady-state deformation in the bar as a whole ($\bar{\ddot{\epsilon}} = 0$) when the local $\ddot{\epsilon} < 0$, i.e. for *decelerating* deforma-

*We use $\bar{\dot{\epsilon}}$ instead of $\dot{L}$ in order to avoid carrying the load-change term along (eq. (44)). In practive, either $\dot{L}$ or $S\dot{P}$ is usually negligible, the other equal to $L\bar{\dot{\epsilon}}$ (see Fig. 11).

tion. This is the case of Lüders front propagation, which will be dealt with in more detail in section 4.1.

Since deceleration is an essential part of front propagation, it would be an inappropriate oversimplification[70] to set the strain rate *constant* in the deforming region. It turns out that a much better approximation is to set the total *strain* in the deformed region constant ($\Delta\epsilon = \epsilon_L$, the Lüders strain), and to use eq. (59) with $\dot{F} = v_F/L$:

$$\bar{\dot{\epsilon}} = \epsilon_L \, v_F/L. \tag{64}$$

A final case, which we shall treat in detail in this section, assumes again that the strain rate is zero over part of the volume, but now that it is *constant* over the remainder; say, a fraction f. (Note the difference to F, eq. (59).) Then,

$$\bar{\dot{\epsilon}} = f\dot{\epsilon} \tag{65}$$

and

$$\bar{\ddot{\epsilon}} = \dot{f}\,\dot{\epsilon} + f\,\ddot{\epsilon}. \tag{66}$$

Here again, it is possible to have zero *average* acceleration when the local deformation *de*celerates: by spreading of deformation ($\dot{f} > 0$).

The physical situation described by this formalism is that of an increasing mobile dislocation density (setting $f = b^2\rho_m$), and a decreasing dislocation velocity (setting $\dot{\epsilon} = v_d/b$). The result (which is well known[73, 66]) is especially interesting in the present context: creep (Fig. 16(a)) accelerates during an "anomalous" transient, goes through a stationary state ($\bar{\ddot{\epsilon}} = 0$), and finally decelerates normally under the dominating action of hardening processes ($\delta > 0$). Equivalently, in a constant-$\bar{\dot{\epsilon}}$-test (Fig. 16(b)), the stress first decreases (during a *slow* yield drop) and then increases again.

The transients just described are associated with completely stable deformation and are, in fact, a *consequence* of locally decelerating deformation: only then is there any reason for additional volume elements to deform (mobile dislocations to be generated). Note that the new *volume elements* that deformation spreads to must deform at the *then prevailing* strain rate (it was assumed independent of x); they must have been affected by the strain hardening that went on before. This could be true when the density of mobile dislocations increases *everywhere* in the specimen; it would be hard to understand, however, if deformation spread into entirely new regions. (That is why we studiously avoided talking about a *length* fraction that deforms, in this last case.) We see, again, that a *macroscopically* observed (finite) negative hardening rate is evidence for *homogeneous* deformation.

4. Modes of Nonuniform Deformation

4.1. *Lüders Front Propagation*

Many solution hardened alloys, tested in a regime of relatively low temperatures, display the initial yield behavior illustrated in Fig. 17. Deformation begins, usually at a grip end, at the upper yield stress σ_U. The stress then immediately drops to the lower yield stress σ_L, and deformation spreads along the whole specimen at essentially constant stress. (Under some circumstances, the plateau is serrated, as in Fig. 17(b); see the next section.) Finally, a

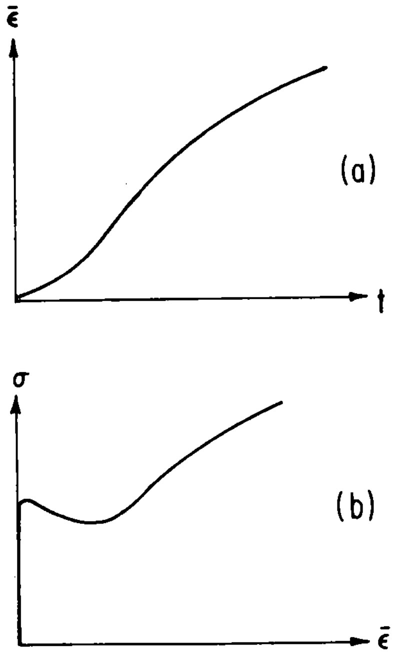

FIG. 16. Increasing deforming volume fraction: anomalous creep (a) and negative hardening (b) are observed in macroscopic tests, even though local material elements are continuously decelerating/hardening.

strain-hardening curve is followed that appears unaffected by the yield phenomenon: it is the same, for example, for specimens in which the yield behavior was deliberately altered.[16]

The common interpretation of this phenomenon is that an initial *generation stress* σ_g is required to break dislocations loose from solute atmospheres acquired during annealing.[15] The macroscopically measured upper yield stress may be lower than σ_g because of stress concentrators; in fact, it is not very reproducible. The lower yield stress (which is much

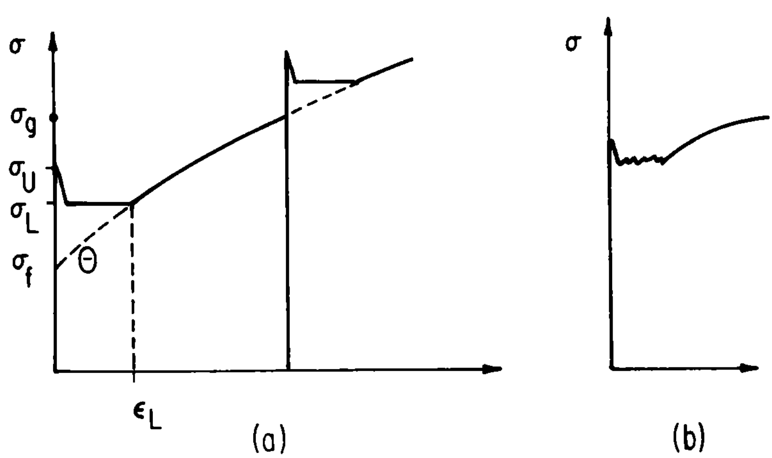

FIG. 17. A sharp yield point, and its recurrence after straining and aging (a); (b) another form of the phenomenon, not treated here.

more reproducible) corresponds to the stress necessary to propagate the deformation front through the specimen. Subsequently, the specimen deforms uniformly.

The presumption in this model must be that, had there been no difficulty in getting the dislocations started initially, deformation would have been possible at an initial friction stress σ_f, and would then have taken place uniformly throughout the specimen. This would suggest that a back-extrapolation of the stress strain curve from the homogeneous regime to strain zero would give the value of σ_f. Back-extrapolating is relatively easy when yielding is followed by strong (and more or less linear) strain hardening. It would be problematical when the stress strain diagram is strongly curved, such as that shown in Fig. 17(b). Figure 17(a), at larger strains, shows an elegant way[74] to establish σ_f: as the endpoint of some predeformation into the homogeneous regime, followed by aging (at a temperature where solute dislocation locking saturates, but dislocations don't recover), and retesting.

The fact that deformation during the yield plateau occurs at a stress higher than the friction stress, must mean that it occurs faster than the prescribed average strain rate—and this is consistent with the observation that it occurs in only part of the specimen. The situation may thus be well described by Fig. 7, for each deforming cross section: when it is initiated into deformation at an "overstress" (corresponding to point 2 in the figure), it must deform at a high strain rate, and then decelerate continuously, until it (more or less) rejoins the stress strain curve for the homogeneous material.

The stress σ_L at which this process takes place is lower than the generation stress. Its derivation is a major task facing a quantitative treatment of this phenomenon, and the problem does not appear to have been solved satisfactorily. In the following, we shall derive a number of interrelations between the various quantities that are measurable, and we shall find them to be one too few.[21] Thus, but a single physical assumption need be made to describe the phenomenon fully.

Two relations are obvious from the figures already presented. The excess of the lower yield stress over the (hypothetical) friction stress must be related to the strain ϵ_L associated with the plateau by the strain-hardening rate Θ:

$$\boxed{\sigma_L - \sigma_f = \Theta\,\epsilon_L} \tag{67}$$

and to the initial (maximum) strain rate $\dot{\epsilon}_F$ by the rate sensitivity β:

$$\boxed{\sigma_L - \sigma_f = \beta \ln (\dot{\epsilon}_F/\bar{\dot{\epsilon}}).} \tag{68}$$

These are two equations between three unknowns. (σ_L, ϵ_L, $\dot{\epsilon}_F$). While the strain ϵ_L at the end of the plateau must clearly be a *consequence* of the stress level, there is no such obvious cause-and-effect relation between the excess stress and the excess strain rate.

We have seen that the strain rate is not uniform. Let us first concentrate on any one cross section that does deform. If the strain rate begins at a value $\dot{\epsilon}_F$, it must decelerate at constant stress (or load) because of local hardening. Simple integration of eq. (18) gives the relation between strain rate and time at any one cross section:

$$\frac{1}{\dot{\epsilon}} - \frac{1}{\dot{\epsilon}_F} = \delta \cdot t. \tag{69}$$

This is plotted in Fig. 18(a) (thin line). After some time, the strain rate has decreased to the prescribed average strain rate $\bar{\dot{\epsilon}}$, and from then on straining will continue even more slowly. It is common to label deformation at $\dot{\epsilon} < \bar{\dot{\epsilon}}$ *creep* and neglect it, even though it is merely the continuation of the ongoing process. The heavy line shows this approximation graphically.

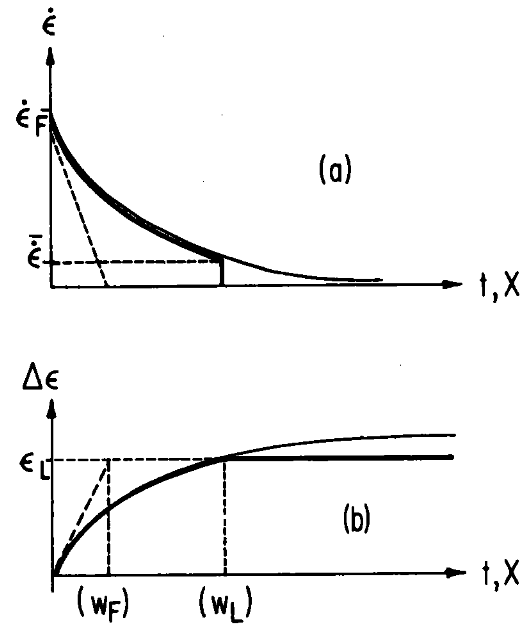

FIG. 18. Decrease of strain rate (a), and increase of strain increment (b) with time when δ = const. (thin lines). Only when strain rates below a certain limit (such as the average strain rate $\bar{\dot{\epsilon}}$) are neglected is the strain limited (heavy lines). Then, steady-state front propagation is possible, and these figures also represent the spatial distribution as a function of the distance from the Lüders front, X.

It must be emphasized that Fig. 18(a) is not drawn to scale: for the typical values of $(\sigma_L - \sigma_f)/\sigma \sim 0.1$ and $\beta/\sigma \equiv M \sim 0.01$, one would get $\dot{\epsilon}_F/\bar{\dot{\epsilon}} \sim 10^4$ (eq. (68)).

A further integration of eq. (69) over time gives

$$\Delta\epsilon = \frac{1}{\delta} \cdot \ln(1 + \delta\dot{\epsilon}_F t). \tag{70}$$

This is classical *logarithmic creep*[75]: a direct consequence of the assumed constancy of δ, i.e. of both the strain-hardening rate and the rate sensitivity (eq. (16)). It is illustrated in Fig. 18(b): the thin line. The strain increases without limit. When the strain rate has decreased to $\bar{\dot{\epsilon}}$, $\Delta\epsilon$ has reached a certain value, whose identification with the previously defined ϵ_L will need to be discussed below. Neglecting "creep" would be expressed by the heavy line.

Realistic creep laws are generally faster than logarithmic creep; e.g. the curve labelled ANDRADE[76] in Fig. 19. At the low strains we are concerned with here, logarithmic creep (which plots as a straight line on this diagram) is in fact a realistic approximation. Using it for the "overstress" regime, but neglecting deformation in the range $\dot{\epsilon} < \bar{\dot{\epsilon}}$, gives the heavy line. Finally, Fig. 19 shows "exhaustion creep": it truly saturates at a finite strain. In Figure 18(a), it would have plotted as an *exponential*, rather than inverse, decay with time, and in Fig. 18(b) as exponential saturation to ϵ_L. This type of law has, in fact, been used to describe deformation in Lüders bands since Hart;[21] it corresponds to assuming $(\partial\dot{\epsilon}/\partial\epsilon)$ constant, rather than the *logarithmic* derivative, and is quite unrealistic.*

*In a previous treatment,[56] we inadvertently made the same assumption by an inappropriate approximation in integrating eq. (64e). For this reason, eqs. (64f) and (64g)[56] are wrong: ϵ_L should there be divided by $\ln(\dot{\epsilon}_F/\bar{\dot{\epsilon}}) \simeq 10$.

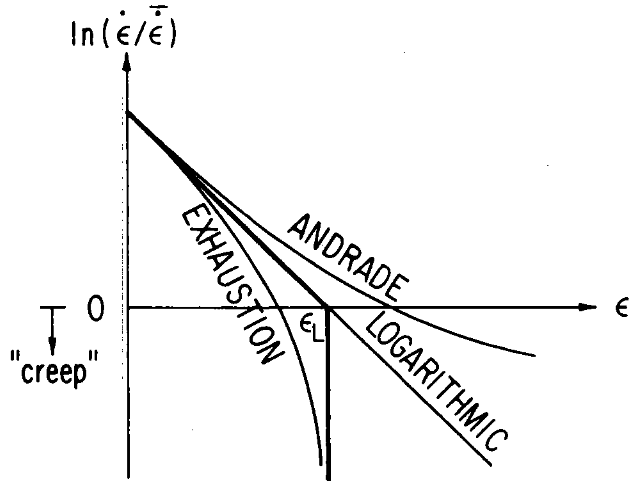

FIG. 19. Creep laws. Normal primary creep decelerates qualitatively as in the curve labelled ANDRADE. Constant δ is LOGARITHMIC creep: a good approximation for small strains. The heavy line represents the relation assumed here. EXHAUSTION creep is physically limited in strain and usually not realistic.

Since deformation decelerates in each cross section, it must spread to other regions of the specimen to maintain a constant average strain rate. A particular mode of spreading is observed to be always associated with this form of yield point: namely, the propagation of a Lüders front. (The term "Lüders band" is sometimes used for the deformed region behind the front, but sometimes also for a small region of deforming material limited on both sides.) It is not clear what material characteristics lead to this particular form of nonuniform deformation, as we will discuss in somewhat more detail below.

For now, we shall assume steady-state propagation. (We have shown in eq. (63) that such a solution exists.) If the current position of the front is x_F and the propagation speed is v_F, the entire behavior can then be described in terms of the parameter

$$X \equiv x - x_F + v_F\, t \tag{71}$$

and there is, at every point, a connection between the strain rate and the strain gradient ϵ':

$$\dot{\epsilon} = \epsilon'\, v_F. \tag{72}$$

The time profiles of the strain rate and the strain shown in Fig. 18(a) and (b) now become distance profiles also. If we assume the heavy lines (neglecting creep), we find that there is a finite zone in which deformation is taking place at any one time, and at its end, $\Delta\epsilon = \epsilon_L$. The width w_L of this zone can be derived from eqs (67), (68), (70), and (72) to be

$$\boxed{w_L/L = \ln(\bar{\dot{\epsilon}}/\dot{\epsilon}_F). \qquad (73)}$$

The average strain rate follows simply from eq. (60) (see eq. (64)):

$$\boxed{\bar{\dot{\epsilon}} = \epsilon_L\, v_F/L. \qquad (74)}$$

For typical values of $\bar{\dot{\epsilon}}$ ($\sim 10^{-3}\ s^{-1}$), ϵ_L ($\sim 10^{-2}$), and L ($\gtrsim 10$ mm), this gives front velocities of order 0.1 to 1 mm/sec.

Let us now consider the effects of *creep*, i.e. the more realistic material behavior illustrated by the thin (solid) lines in Figs. 18 and 19, at $\dot{\epsilon} < \bar{\dot{\epsilon}}$. Inasmuch as straining is still

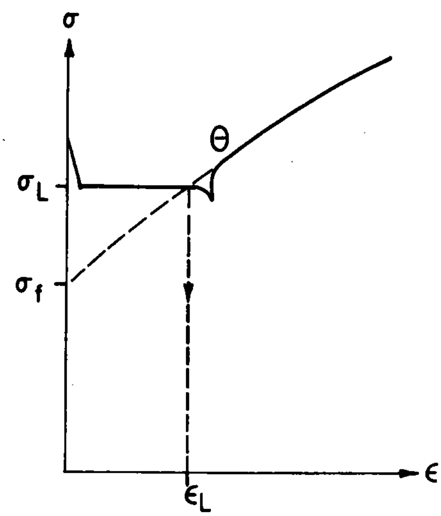

FIG. 20. Definition of the Lüders strain when the assumptions used here are consistently applied. The final drop in load is probably due to an end effect; the sharp raise may be due to creep in the deformed region.

going on everywhere behind the front, the average strain rate would have to continuously increase if (a) the front propagated at constant speed, and (b) the stress remained constant. If $\bar{\dot{\epsilon}}$ = const. is considered a boundary condition, then either the stress must decrease, or v_F. Neither effect is observed, presumably because it is so small. However, one effect that is observed may possibly be due to this material characteristic. It is illustrated in Fig. 20. At the end of the Lüders strain regime, there is a fairly abrupt, small increase* in flow stress, which we have neglected in Fig. 17, before the "normal" stress strain curve is rejoined. This might correspond to the fact that the whole specimen was deforming at a low (creep) strain rate, and is now suddenly forced to proceed at the prescribed $\bar{\dot{\epsilon}}$. If this is so, back-extrapolating the *normal* stress strain curve to the stress plateau would now operationally define the Lüders strain ϵ_L as shown in Fig. 18(b), and the small amount of strain accumulated *after* ϵ_L in the plateau would be due to slow creep.

Finally, we will address the question, what material characteristics may be responsible for the Lüders mode of nonuniform deformation. It is clear that a high generation stress forces a high initial strain rate, and that this implies either localized deformation or severe unloading; since generation is likely to be a heterogeneous event in the first place, localized deformation can be easily rationalized. Further, the decelerating characteristics of the material must lead to spreading of the deformation along the specimen—and they leave the material in an (almost) homogeneous state (Fig. 7). But why does the spreading occur in such an orderly fashion as by the propagation of a front; and why does it occur at a particular stress which, while being higher than the friction stress is also lower than the generation stress?

These questions have to be answered on the basis of some physical mechanism. Many authors imply that there is an innate *propagation stress* (for *front* propagation, not dislocation propagation), possibly connected with the front velocity; but no structural reason for such a material property appears to have been given. In general, any one further relation between the various unknowns introduced above would make the problem determinate†—but it should also answer the qualitative question of why there is orderly front propagation,

***This follows on a small decrease, which can be explained by the front running into the far grip.**

†By 2002, significant progress has been made in this area [L.P. Kubin, C. Fressengeas, and G. Ananthakrishna, "Collective Behavior of dislocations in plasticity": in *Dislocations in Solids* **11**, Edited by F.R.N. Nabarro and M.S. Duesberry (North-Holland, 2002) pp.101-192]

and not successive generation at random places. Why is it easier to generate new dislocations near a region that has already deformed?

One approach to this question is based on the presumed existence of stress concentrations at the front. For example, a fairly sharp change in area must establish itself at the front, and this could lead to stress concentrations at the two ends of the current band. This hypothesis has been disproven by experiments in which the specimen was remachined after the front had passed along half its length: after reloading, the front continued where it had left off before. On the other hand, a similarly half-deformed specimen that went through an intermediate anneal, started a new Lüders front at the grip end upon reloading (even if it had not been remachined).[16]

The latter experiment clearly indicates that it must be a structural feature that defines the front. If it were in the nature of an *internal* stress concentrator, it would have to have a "wavelength" that is structurally determined (since the *average* internal stress across the cross section must be zero). The grain size might be a suitable candidate for such a parameter—if it weren't for the fact that single crystals also show Lüders front propagation under similar circumstances.[77]*

Another approach to the problem may be the most promising: to think of the front as a place where (unlocked) dislocations are injected into the as yet undeformed region.[78] Even locked dislocations may be freed by the high internal stress peaks associated with passing dislocations. In this model, there would have to be a fairly close correspondence between the front velocity v_F and the dislocation velocity v_d. Taking them both to be of order 1 mm/sec, and $\dot{\epsilon} \sim 1\ \text{sec}^{-1}$, would require a mobile dislocation density ρ_m of order 10^9 cm^{-2}, using

$$\dot{\epsilon} = \rho_m\, b\, v_d. \tag{75}$$

If ρ_m were a given material property, eq. (75) would provide the missing relation. This hypothesis would predict that a specimen which was heavily cold-worked and then aged should, all other parameters being equal, exhibit a lower front velocity and smaller excess stress. Most significantly, this theory would suggest an intimate connection between the kinetics of front propagation (v_F) and of later uniform straining ($\dot{\bar{\epsilon}}$), since both are proportional to the dislocation velocity. Such a similarity in kinetic behavior has often been observed.[79]

Until now, we have assumed that the steady-state solution of front propagation is actually realized. A check on this assumption can be made by deliberately disturbing an assumed steady state and ascertaining whether at least the system will then return to steady state. This check turns out affirmative; for if the propagation velocity is raised (keeping everything else the same), the average strain rate increases (eq. (74)); thus, the load must drop (eq. (44)), which will decrease all rates, including that of front propagation, if indeed it is controlled by the dislocation velocity. This feedback mechanism relies on a positive rate sensitivity and works the better the larger it is. Solution hardened alloys are (next to those controlled by lattice resistance) the most rate sensitive of all materials, in this low-temperature regime.

Lastly, let us discuss the possible influence of a dynamic response in the specimen, due to the strain-rate discontinuity when deformation starts. In principle, this could determine σ_L

*Another possibility is that the *strain gradient* cannot exceed some critical value, such as ϵ_L/d, where d is the grain size or the slip plane spacing. This would establish, with eq. (72), an additional relation between the strain rate at the front and the propagation velocity. In general, if the strain gradient entered into the material constitutive laws in any way, v_F could be derived from eq. (72).

by being a certain $\Delta\sigma$ below σ_U (even though this difference is not observed to be very reproducible). The order of magnitude of this effect can be estimated from eqs. (51) and (52):

$$\Delta\sigma = -\frac{E}{v_S} \cdot w\dot{\epsilon}_F. \tag{76}$$

The width w to be used here should be of order of the *width of the front* defined by

$$w_F \equiv \epsilon_L/\epsilon'_F \tag{77}$$

which, by virtue of eqs. (72) and (74), equals $\dot{\mathrm{L}}/\dot{\epsilon}_F$. Thus,

$$\Delta\sigma/E \sim -\dot{\mathrm{L}}/v_S \tag{78}$$

which is of order 10^{-6} and thus much too small. Of course, even if this explanation held for the stress, it would not be able to rationalize the orderly progression of the front.

4.2. *Jerky Flow*

In an intermediate temperature range, some solution hardened alloys exhibit a phenomenon variously called jerky flow, serrated flow, the Portevin–LeChatelier effect or, by the assumed mechanism, dynamic strain-aging. Stress strain curves (at constant extension rate) show repeated load drops of various kinds, which are illustrated in Fig. 21. They have been classified[17] into the types *A*, *B*, and *C*, all of which are fairly regular; in addition, there is an irregular kind of slow load-drop behavior (Fig. 21(d)), which is sometimes called "jerky flow" in a narrower sense. We will refer to all these phenomena as *jerky flow.*

As indicated in Fig. 21, jerky flow often commences after a finite amount of prestrain; also, it often ceases at large strains. The strain range for jerky flow depends on temperature and strain rate. Under some conditions, jerky flow occurs simultaneously with (initial) Lüders front propagation; then it is called "serrated yielding" (Fig. 17(b)). pointed out

It has been ~~postulated~~[23] that a negative rate sensitivity of the flow stress $\boxed{\text{is}}$ always should be associated with jerky flow. Measurements of negative rate sensitivity are difficult, as we have indicated above, and are therefore scarce. The most fruitful approach is to investigate the rate sensitivity as a function of strain: one finds that it decreases toward zero as the starting point of jerky flow is approached, and it becomes more positive after jerky flow has ceased. Figure 22 shows the rate sensitivity β as a function of the flow stress reached during

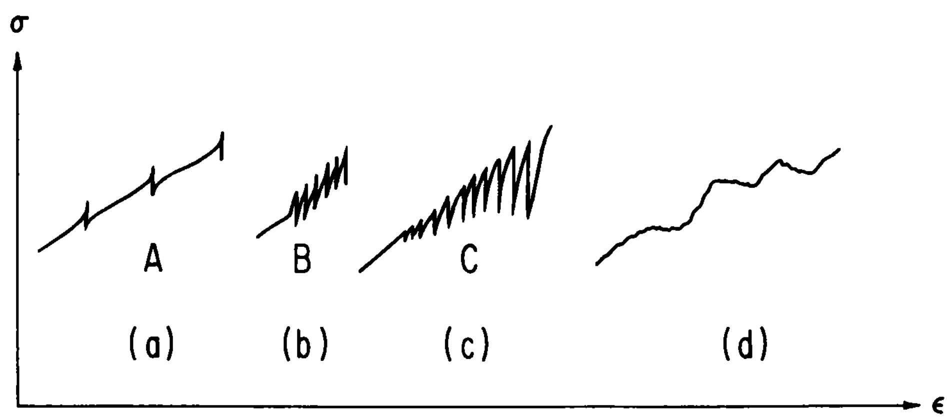

FIG. 21. Types of jerky flow. We primarily address type *C*.

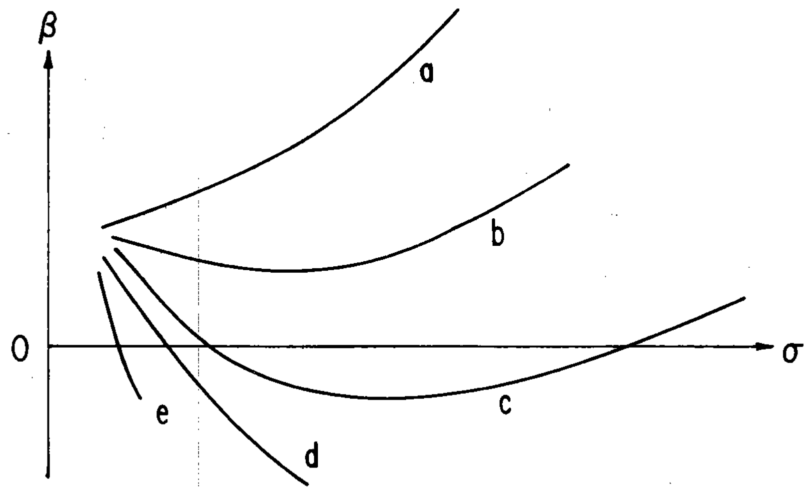

FIG. 22. Strain-rate sensitivity β and flow stress σ as a function of prestrain. Normal solution hardened allows follow curve a; all others exhibit dynamic strain-aging, although jerky flow is only observed when the (total) rate sensitivity is negative. This occurs after a critical strain in curves c and d, which is virtually zero in curve e. Dynamic recovery is proposed to be responsible for the eventual rise in rate sensitivity, and thus for the ceasing of jerky flow.

the prestraining.[34] Curve c exhibits the behavior just discussed. It is obvious that even if the portion of the curve at $\beta < 0$ were not actually observed, it could be inferred by interpolation. Changing temperature, strain rate, or solute concentration can change the behavior either towards curves d and e, or towards curves b and a. Curve a is typical for a "normal" solution hardened alloy, i.e. outside the regime of dynamic strain aging:[56] the initial rate sensitivity is due to the friction stress, the (mild) subsequent increase is due to dislocation forest intersection, and the stronger eventual rise is due to dynamic recovery.[49] If all the other curves in Fig. 22 are treated as departures from this "normal" behavior, the difference is a *negative* contribution to the rate sensitivity that increases with prestrain. Such a negative contribution is associated with dynamic strain-aging by one mechanism or another.[19, 34] In this interpretation, curve b would then show evidence of dynamic strain aging, even though the *total* rate sensitivity is not negative and jerky flow is never observed.[34]

In summary, then, the physical mechanism underlying the observation of jerky flow (Fig. 21) is presumed to be as follows: dynamic strain-aging, which is effective over a certain range of temperatures and strain rates, makes a negative contribution to the rate sensitivity, which increases with prestrain; when it has become dominating (at the "critical strain"), jerky flow begins; jerky flow ceases if and when, in the given temperature and strain-rate range, dynamic recovery becomes sufficiently important at large strains.

We have seen in section 2.7 that a negative rate sensitivity is bound to cause abrupt acceleration. The question is now whether, or under what conditions, such time-discontinuities in strain rate, which we had already labelled "jerky flow", lead to space-discontinuities, i.e. localized flow, and to load drops in a hard machine.

Figure 9 showed the material characteristic for dynamic strain-aging on a local basis. Figure 23 presents a section through it for a particular stress path (not necessarily constant). It illustrates that jerky flow in each material element consists of three stages: an initial (positive) strain-rate discontinuity (point 1 to point 2); followed by decelerating accumu-

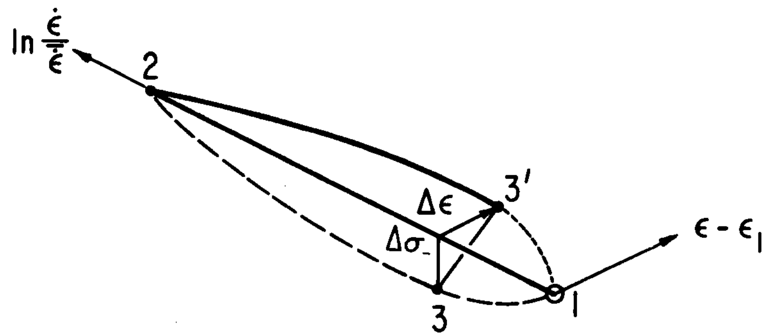

FIG. 23. Appropriate cut through the material characteristic for jerky flow. The strain rate will change dicontinuously (from point 1 to point 2); strain is accumulated during decelerating deformation (to point 3′), and then ceases. The final strain rate exceeds the average strain rate (at point 1).

lation of plastic strain until the condition $\beta = 0$ is reached (point 3′); and then an abrupt arrest of deformation or, more exactly, a discontinuous change to slow creep, which is here neglected.

In a creep machine, at constant stress, the local strain increment is immediately given by eq. (27): it is dictated by strain hardening and the depth of the valley, $\Delta\sigma_-$, in the stress vs strain-rate characteristic. In principle, this strain increment could occur in the whole specimen at the same time. In practice, however, there are likely to be small variations along the specimen length so that the fluctuation will occur in one small volume element first; e.g. near a grip. Then, the stress may have to be raised slightly to activate the next element, or alternatively a deformation front may propagate along the specimen at constant stress. When the entire volume has experienced the strain $\Delta\epsilon$, no further deformation is possible until the stress is raised and the process starts over.[36] It is the essential difference between Lüders front propagation and jerky flow that in the former, completed band propagation leaves the material free to deform homogeneously, whereas after jerky flow the material is returned to the jerky-flow condition.

Why and under what conditions deformation should progress orderly by front propagation is an open question, just as it was for Lüders bands. For the subsequent fronts in jerky flow, the unidirectional progression may be a consequence of a strain gradient built up by the first front.[23]

In a hard machine, the stress is not necessarily constant: it may be driven down if, for some physical reason, the actual plastic extension rate exceeds the prescribed extension rate (eq. (44)) or, dynamically, if there is a strong local strain-rate discontinuity (eq. (52)). One or the other of these effects has been deemed to be responsible for limiting the local strain increment.[23, 72] We hold this to be unlikely, inasmuch as the stress drop due to either effect is proportional to the width of the deforming region and this would have to be a substantial fraction of the specimen length to be of any importance. It is not clear why such a large region should deform uniformly, and what determines its size. We will now present an alternative analysis that assumes deformation to start in a very thin sliver and spread from there. In this case, it will turn out that the interaction with the machine determines the distance the front will travel, which is the effective band width w.

The amount of deformation in a small volume element (or even a single slip plane) is not limited by the machine response but by strain hardening (eq. (27)). If deformation now spreads into neighboring elements with a propagation speed v_F (this is the *ad hoc* assumption), the applied stress will drop at a rate $\dot{\sigma}$ given by the difference between the plastic extension rate $\Delta\epsilon \cdot v_F$ (eqs. (37) and (64)) and the "prescribed" one, $\dot{L}$. After the front has progressed a distance w, the stress will have dropped by (eqs. (44) and (48))

$$\Delta\sigma = \frac{w}{v_F} \cdot \dot{\sigma} = -\bar{E} \cdot \left(\Delta\epsilon - \frac{\dot{L}}{v_F}\right) \cdot \frac{w}{L}. \tag{79}$$

It is seen that at a particular front velocity $v_F = \dot{L}/\Delta\epsilon$ there would be no stress drop. On the other hand, when v_F is very large, the stress drop is equivalent to that in a stress relaxation experiment. The crucial observation now is that when the stress has dropped to the "bottom of the valley", the front must stop:

$$|\Delta\sigma| \leqslant \Delta\sigma_- \tag{80}$$

(This condition could be violated by the effect of a stress concentration at the front.) In fact, since the stress drops in proportion to the distance the front has travelled, the overstress decreases linearly, and $\Delta\epsilon$ decreases linearly to zero. Thus, the average strain in the band is half $\Delta\epsilon$, and the band width w follows by insertion into eq. (79):

$$\boxed{\Delta\epsilon \approx \frac{1}{2} \cdot \frac{\Delta\sigma_-}{\Theta}; \quad \frac{w}{L} \approx \frac{2\Theta}{\bar{E}}.} \tag{81}$$

Since typically $\Theta < E/50$ (which it is at small strains and low temperatures) and $\bar{E} \simeq E/5$, the band width is less than about 1/5 the specimen length. Values of order 1/10 are typically observed.[36]

After the band has thus been arrested, another one can start either adjacent to it[36] ("hopping band"[23]), or randomly along the length of the specimen.[17] In either case, no significant stress increase is expected until volume elements must be strained a further time. Such plateaus have sometimes been observed; more often they seem to be smoothed out, possibly through the introduction of a strain gradient along the specimen.[23] This sequence of events appears to characterize the observed type-C bands quite well (Fig. 21). Bands of type A and B may well be due to a smaller intrinsic front velocity, according to eq. (79).

The question remains, what controls the front velocity. As in the case of Lüders front propagation, it seems plausible to postulate that it has some connection to the local strain rate (or to the dislocation velocity[78, 80]). In that case, v_F ought to be the larger the more severe is dynamic strain aging. This concept is in accord with the observation that, as the most critical strain and temperature ranges are approached, the serration type progresses from type A to type C. (Type B is classified as an intermediate stage, but somewhat differently by different authors.[17, 36])

Finallly, a word about dynamic effects. In all of the above discussion, we have assumed that communication along the specimen is rapid enough for dynamic effects to be ignored (section 3.4). This means, first, that $v_F \ll v_s$, which is certainly expected. More stringent is the assumption that the time required to strain an individual cross section is longer than the time t_D for the round trip of elastic waves to the end of the specimen: $\Delta\epsilon/\Delta\dot{\epsilon} > t_D$. With a strain increment of order 1%, and t_D of order 10^{-5} s (eq. (56)), this would require local strain rates less than $10^3\ \mathrm{s}^{-1}$, which sounds reasonable. However, on a scale of individual slip bands, the condition may be violated and lead to a significant influence of dynamics.[72]

4.3. *Necking*

The most common form of flow localization in all ductile materials is that occurring near the end of a tensile test: the development of a neck. In first approximation, it is the result of a *geometric* instability (eq. (19)): when the rate of work hardening becomes less than the relative rate of area decrease with strain in a tensile specimen, H (eq. (9)).

$$H < 1 \tag{82}$$

deformation accelerates at constant load (eq. (22)):

$$\delta_P < 0. \tag{83}$$

It accelerates more wherever the strain rate is already larger (eq. (43)); thus, strain-rate gradients (and therefore strain gradients) tend to accentuate (eq. (32)). Equation (82) is called the Considère criterion.[7] Equation (83) is more general in that it also holds for compression and other cases of area-change vs. strain relations.

In a creep test at constant load, the critical condition $\delta_P = 0$ (eq. (83)) coincides, by definition, with the minimum strain rate. Even in a creep test "at constant true stress" (assuming uniform deformation, that is), the conditions (82) and (83) can be reached; however, the value of $(-\delta)$, the slope of the $\ln \dot{\epsilon}$ vs. ϵ diagram, is still negative at this point.

In a tensile test "at constant true strain rate" (if the deformation were uniform), the equilibrium condition of constant load along the specimen coincides with the condition of zero load change with strain: the point of maximum load or the ultimate tensile strength (UTS). At constant *extension* rate, the condition is modified slightly.[9]*

In conventional treatments, the Considère criterion (with or without Hart's[9] modifications) is considered to be the demarcation line between "stable" and "unstable" deformation. This is strictly true only, when a particular (and, we think, unrealistic) *defect* is assumed to be present in the specimen. We shall show below that the particular formulation we have given above (eq. (83)) in fact allows for (realistic) instabilities long before the Considère criterion is reached—but that, on the other hand, the condition $\delta_P \simeq 0$ has a much broader significance: only as it is approached can localization become *severe*. Following this discussion, we shall address the question of the *rate* of neck development.

As distinct from the other modes of nonuniform deformation, which start abruptly and then decay (at least temporarily), necking develops slowly, with sometimes considerable accumulation of strain. It is for this reason that area changes, and the resulting stress changes, become important. The distinction between δ and δ_P was already a consequence of this effect. We shall now account for it in a more explicit fashion. Following our earlier treatment,[13] we define the relative gradient of the cross-sectional area, which is always (eq. (1)) linked to the relative gradient in stress, by

$$Y \equiv \left.\frac{\partial \ln A}{\partial x}\right|_t = -\left.\frac{\partial \ln \sigma}{\partial x}\right|_t \tag{84}$$

For tension (eq. (2)), it is also equal to the negative strain gradient, plus an initial area defect Y_0:

$$Y = -\epsilon' + Y_0 \qquad \text{(tension)} \tag{85}$$

*Another small modification appears in Hart's[9] paper, because he considered the accentuation of *absolute* area differences. (See footnote to eq. (21).)

Its change with strain is then related to the strain-*rate* gradient (eq. (35)) by

$$\left.\frac{\partial Y}{\partial \epsilon}\right|_x = -\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t \qquad \text{(tension)} \tag{86}$$

Thus, the differential equation for the development of strain-rate gradients (eq. (33)) can now be expressed as

$$\frac{\partial^2 Y}{\partial \epsilon^2} + \delta_p \frac{\partial Y}{\partial \epsilon} = \delta'_p \qquad \text{(tension)} \tag{87}$$

with the integral (eq. (36))

$$Y = Y_0 + \frac{1}{\delta_p} \cdot \left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_0 \cdot [1 - \exp(-\delta_p \epsilon)] \qquad (\delta'_p = 0) \tag{88}$$

This is different from eq. (36) in one important respect: it allows one to specify as one of the integration constants an initial *area* gradient Y_0 (instead of a strain gradient). We will discuss this type of initial condition as the first in a number of "defects".

A *geometric defect* (or *machining defect*) is characterized by an initial (relative) area gradient (eq. (84))

$$Y_0 = \frac{\mathrm{d} \ln A_0}{\mathrm{d}x} < 0. \tag{89}$$

(The location of the defect with respect to the origin of x is again chosen such as to make localization correspond to *positive* strain and strain-rate gradients, as in eq. (28)—see eq. (85).) This area gradient causes an initial stress gradient $-Y_0$ (eq. (84)) and therefore (if the material is homogeneous) an initial strain-rate gradient (eq. (10))

$$\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_0 = -Y_0/M > 0. \tag{90}$$

Inasmuch as this strain-rate gradient is positive, it will *accentuate* the relative area gradient even when $\delta_P > 0$, i.e. before the Considère criterion.[50] If one looks at small differences only and therefore neglects gradients in the material properties (the right-hand side of eq. (87)), the development is described by eq. (88): while the strain gradient does increase, it saturates exponentially. Thus, deformation is "unstable" against area variations—even though it is *bounded.*

A very similar situation obtains when the area is completely uniform, but the material is heterogeneous (a *physical defect*): if the glide resistance is lower somewhere, then the strain rate will be greater there at the same stress. Equivalently, one may assume a *fluctuation* in strain rate. In this case, $Y_0 = 0$, and the initial strain-rate gradient is an independent constant; however, the development is then the same as for the geometric defect. Thus, we have plotted them together in Fig. 24, as ϵ' vs. strain. This is similar to curve I in Fig. 10(a), but is now shown for four different values of the deceleration parameter δ_P: 10, 1, 0, and -1. (Note that typical values at small strains are of order 10^3.)

Figure 24 demonstrates that the *severity* of the instability increases drastically as δ_P decreases. In fact, the precise condition $\delta_P = 0$ becomes almost academic: even if the geometric area change could be neglected (or were opposite, as in a compression test) will strain-rate gradients persist for a long time, and strain gradients grow, when the hardening

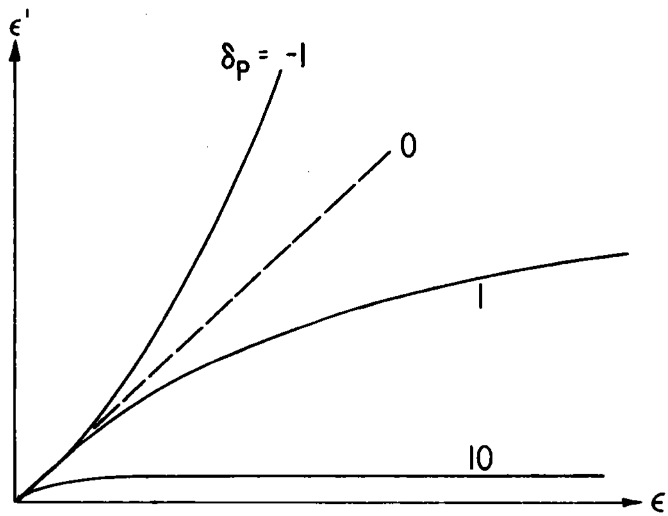

FIG. 24. Development of the strain gradient ϵ' with strain, for a constant initial fluctuation in strain rate (initial slope in this diagram) which could have been due to a negative initial area gradient. Only when the deceleration parameter is negative is localization unbounded; but small positive values cause considerable strain gradients to develop over large strains.

rate becomes low. A *general condition for high localization resistance* may thus be written as

$$\delta \gg 1 \tag{91}$$

where the boundary-independent material parameter δ serves just about as well as the (more proper) quantity δ_P. (Note again that this is the *local* property, not necessarily the slope of the macroscopic creep curve, say, when dislocation generation is important.)

In the above examples, the Considère criterion signalled the demarcation line between bounded and unbounded growth,[15] not between stable and unstable deformation.[9] The latter can be obtained in a different type of defect; namely, when the initial strain-rate gradient is *negative*. In particular, this is the case when one assumes that a certain cross section has had more *deformation*: while this makes its area smaller, it has work hardened enough to make the total effect on the strain-rate gradient negative:[50, 13]

$$Y_0 = -\epsilon_0'; \quad \left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_0 = -\delta_P \epsilon_0'. \tag{92}$$

The resulting behavior was illustrated in Fig. 10 (curves II), for $\delta_P = \text{const}$. It is also reproduced in Fig. 25, for four different values of δ_P: 10, 1, 0, and -1. Again, the behavior for $\delta_P \lesssim 1$ is of a long-term nature in any case, but for $\delta \gg 1$, it is strongly damped (eq. (91)).

This *deformation defect* cannot properly be considered the result of a fluctuation: the strain (as opposed to the strain rate) cannot change instantaneously. It has been called a "hammer blow"[50]— but it would correspond to a very special hammer blow indeed: one that causes uniaxial extension at the same rate as the previous uniform deformation. In conclusion, we hold this defect to be an especially unrealistic one; moreover, it is especially safe. Thus, the Considère condition is not an adequate criterion for *(in)stability*.

The most important result of these considerations, which dealt with the rate of development of nonuniformities, not just with yes/no criteria, was that much strain gets accumu-

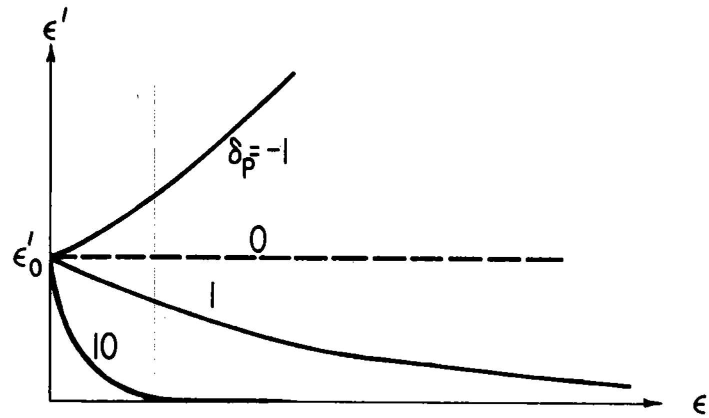

FIG. 25. Development of the strain gradient ϵ' with strain for a deformation defect (characterized by a constant initial strain gradient ϵ'_o and an initial relative strain-rate gradient $-\delta \cdot \epsilon'_o$). Growth occurs only for $\delta_P < 0$ but decay is slow for small positive values of the deceleration parameter.

lated during changes when δ_P is small. For this reason, it was here actually *not* justified to assume the material properties constant. For example, if one begins at a prestrain just before the Considère criterion has been reached, one will exceed it soon. Thus, one would have to pass from the bounded-growth case to the unbounded-growth case; yet all the curves shown in Figs. 24 and 25 are monotonic. Figure 26 illustrates qualitatively what should happen, for both a geometric defect and a deformation defect. The curves have two branches:[13, 81] they are integrals of a *second*-order differential equation. This is precisely the equation we have derived (eq. (87)), when the right-hand side is *not* neglected, as will now be demonstrated in detail.

We shall assume δ_P to be a function of $\dot{\epsilon}$ and σ. As was discussed at length before, this would be a *unique* function, if but one parameter were necessary to specify the mechanical state of the material (if an equation of state existed). Otherwise, the function is path dependent—but the parameter δ_P was chosen precisely because it is specified along the path of

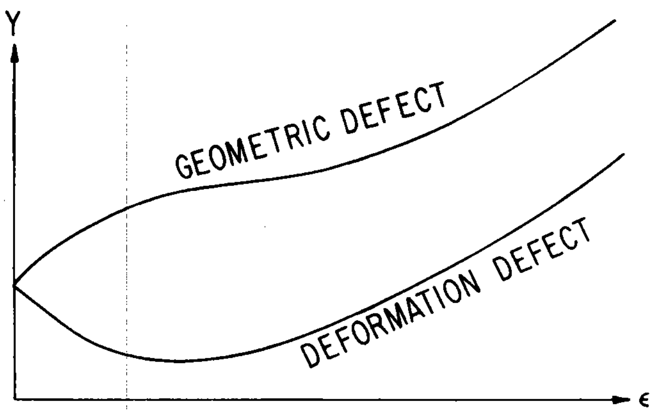

FIG. 26. Qualitative development of the relative area gradient for a geometric defect and a deformation defect, when the initial value of δ_P is positive but small, e.g. near the UTS. The two branches are described as solutions of a second-order differential equation, which has been derived on the basis of realistic constitutive laws.

interest. (It can only *momentarily* be replaced by $(H - 1)/M$, eq. (22).) The gradient of δ_P then becomes

$$\left.\frac{\partial \delta_P}{\partial x}\right|_t = \left.\frac{\partial \delta_P}{\partial \ln \dot{\epsilon}}\right|_\sigma \cdot \left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t + \left.\frac{\partial \delta_P}{\partial \ln \sigma}\right|_{\dot{\epsilon}} \cdot \left.\frac{\partial \ln \sigma}{\partial x}\right|_t. \tag{93}$$

If we define the variations of δ_P (with a sign that makes them normally positive) by

$$b \equiv \left.\frac{\partial \delta_P}{\partial \ln \dot{\epsilon}}\right|_\sigma \tag{94}$$

and

$$c \equiv -\left.\frac{\partial \delta_P}{\partial \ln \sigma}\right|_{\dot{\epsilon}} \tag{95}$$

the general differential equation for the development of nonuniform deformation becomes (eqs. (33), (21) and (84))

$$\boxed{\frac{\partial}{\partial \epsilon}\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t + (\delta + b)\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t = -\frac{1}{M}\frac{\mathrm{d}\ln A}{\mathrm{d}\epsilon} - c\left.\frac{\partial \ln A}{\partial x}\right|_t.} \tag{96}$$

Here, we have kept area gradients distinct from strain gradients, in order to be independent of the *tensile* test.

For tension (eqs. (85), (86) and (22)), we get

$$\frac{\partial^2 Y}{\partial \epsilon^2} + (\delta_P + b)\frac{\partial Y}{\partial \epsilon} - cY = 0. \tag{97}$$

This equation was first introduced[13] in a slightly different form: then path independence in a $(\sigma,\dot{\epsilon})$-description had been assumed (the glide resistance being the only state parameter) and, furthermore, M had been assumed constant. In that case,[13]

$$M\frac{\partial^2 Y}{\partial \epsilon^2} + (H - 1 + B)\frac{\partial Y}{\partial \epsilon} - CY = 0 \tag{98}$$

where B and C are partial derivatives of H with respect to $\ln \dot{\epsilon}$ and $(-\ln \sigma)$, respectively. The connection to the new (and preferable) coefficients is

$$b = \frac{B}{M} - \delta\left.\frac{\partial \ln M}{\partial \ln \dot{\epsilon}}\right|_\sigma \tag{99}$$

and

$$c = \frac{C}{M} + \delta\left.\frac{\partial \ln M}{\partial \ln \sigma}\right|_{\dot{\epsilon}} \tag{100}$$

The most important feature of eqs (97) and (98) is that the last term is negative. In the case of eq. (98), this was previously shown[13] to be a consequence of the fact that stress strain curves are usually convex upward ($C > 0$). For the more general case of eq. (97), it is demonstrated in Fig. 27: at higher stresses, deformation (is faster and) decelerates less. This is especially evident at the minimum creep rate.

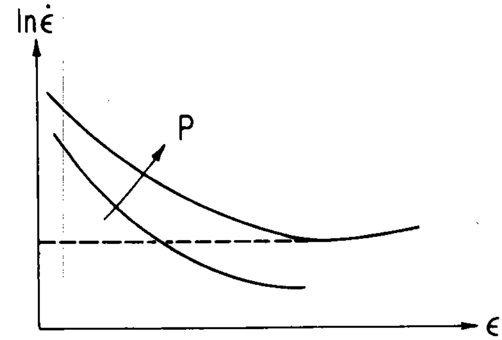

FIG. 27. Typical creep curves for two different constant loads. The fact that the deceleration parameter δ_P (the magnitude of the slope in this diagram) decreases with increasing stress (at constant strain rate) causes eventual unbounded nonuniformity of deformation to be a necessity.

The consequence of the negative sign in eq. (97) is analogous to having a negative restoring force in an oscillator: the "displacement" Y must eventually increase without limit. Thus, *all* tensile deformation (for which the parameters in eq. (97) are positive) must eventually become nonuniform, and unboundedly so!

The *two branches* of neck development can now be specified in a somewhat more quantitative way by inspection of eq (97): it involves two "strain constants"; namely, the ratio of the coefficients of $(\partial^2 Y/\partial\epsilon^2)$ and $(\partial Y/\partial\epsilon)$ for short-term behavior; and the ratio of the coefficients of $(\partial Y/\partial\epsilon)$ and Y for long-term behavior.[13] The first is almost the same as we had seen before: $1/(\delta_P + b)$ instead of $1/\delta_P$. The second is of special interest when $\delta_P = 0$; then,

$$\frac{b}{c} = \left.\frac{\partial \ln \sigma}{\partial \ln \dot{\epsilon}}\right|_{\delta_p} \equiv N. \tag{101}$$

This is an unusual rate sensitivity: *viz.*, at constant deceleration.* For the case $\delta_P = 0$, $1/N$ is in fact the stress exponent in a power-law expression for the minimum creep rate!

The long-term behavior is thus governed by the rate sensitivity of (more or less) steady-state creep. If N were assumed constant, integration of eq. (97) under neglect of the first term would give

$$Y = Y_0 \exp(\epsilon/N). \tag{102}$$

Note that N is not a *decay* constant, but governs the rate of *increase*. The meaning of Y_0 in eq. (102) will normally be that of a strain gradient (eq. (86)) that resulted from earlier neck evolution.

Equation (102) was derived by Hart[9] for the rate of neck growth in the steady-state limit. (Unfortunately, he used the same symbol m for the steady-state rate sensitivity as for the instantaneous rate sensitivity, here called M.) In fact, if every cross section of the specimen were in steady state, and steady state obeys a power law

$$\dot{\epsilon} \propto \sigma^{1/N} \tag{103}$$

then the gradients of stress and strain-rate must be linked by

$$\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t = \frac{1}{N} \left.\frac{\partial \ln \sigma}{\partial x}\right|_t$$

*Previously,[13] it had been defined at constant hardening rate, and M had been assumed constant.

or, with the definitions (35) and (85),

$$\left.\frac{\partial \epsilon'}{\partial \epsilon}\right|_x = -\frac{Y}{N}. \tag{104}$$

This is the equation that led to eq. (102), and follows from eq. (97) for $(\partial^2 Y/\partial\epsilon^2) \simeq 0$ and $\delta_P = 0$.

In conclusion, we emphasize two points. First, that when all material elements have reached the steady-state limit, the *specimen* is *not* in steady state; it necks at a rate given by the steady-state stress exponent. Second, that just before steady state is reached, near the condition $\delta_P = 0$, the material parameters change markedly during the development of the nonuniformities; the behavior can then only be described by the second-order equation (97).

We have argued before[13] that in this same regime, it may be an adequate approximation to set the parameters b and c constant, while $\delta_P \simeq 0$. In that approximation, eq. (97) can be integrated in closed form. It is especially simple when $b \gg 1$, which is realistic for some high-temperature materials[13] and, in addition, makes any *variations* of δ_P near zero truly insignificant. The strain gradient is then given by

$$Y = Y_1 \exp(-b\epsilon) + Y_2 \exp(\epsilon/N). \tag{105}$$

The constants Y_1 and Y_2 depend on the two initial conditions. Note that the role of δ_P as the decay constant for the short-term branch has, in this limit, been assumed by b, the *curvature* of the creep plot ($\ln \dot{\epsilon}$ vs. ϵ) near the minimum, instead of the *slope*. Taking account of rate sensitive strain hardening, one can therefore broaden the condition (91) for *high localization resistance* (i.e. a potential for *superplasticity*) to read

$$\boxed{\delta + b \gg 1 \qquad (106)}$$

where again the difference between constant-load and constant-stress conditions can usually be neglected.

Numerical solutions for arbitrary but constant values of the coefficients in eq. (98) have been obtained.[13] They show the overwhelming importance of the strain-rate sensitivity of strain hardening, in the low-hardening regime.

There are many important questions of detailed geometry, and its effect on stress distribution, which enter into a proper treatment of necking.[2] We have left these out of our discussion entirely, since we were concerned with the *kinetics* only.

4.4. *Compression, Torsion, Sheet Forming*

Up to now, we have often assumed the macroscopic test piece to be in *tension*. This was not always a necessary restriction, but more in the nature of a convenient simplification. The overriding concern was for a sample with *alternative* straining elements. The easiest exemplification of such a piece of material is a long, slender specimen, as it is used in a tensile test.

Compression specimens are usually squat. Since plastic deformation occurs by shear, the straining elements tend to be inclined toward a uniaxial stress axis, rather than being perpendicular to it. Thus, they interact with the compression platens: they are not *alternative* in the sense that they can act independently of each other. However, compression loading is also sometimes imposed on samples of a length-to-diameter ratio greater than one. In that case, the difference to a tension test becomes merely a matter of *sign*.

The stress, the strain rate, and the strain increment were used in unsigned convention in this article: they are the *equivalent* quantities in the sense used in plasticity theory; measures of the scale rather than of the (tensorial) "direction". The only relations in which the sign was important are of a *geometrical* nature: the connection between strain increment and area change (eq. (2)), and those that followed from it: eqs. (20), (22), (85) through (88), (97), and (98)).* In particular, the difference between the material property δ (the deceleration parameter in a constant-stress creep test) and the critical test parameter δ_P (at constant load) depends on the kind of loading. In many of the applications, we have found this difference to be of no consequence; particularly, when strains were small. Therefore, the relations derived for Lüders front propagation and for jerky flow hold with negligible modifications for compression tests on relatively tall specimens.

Torsion loading is a case intermediate between tension and compression: here, there is no connection at all between area and strain (assuming the length is held constant). Again, when the bar is long enough for the straining elements to be regarded as alternative (which may here not be very long), localizations may develop. An investigation of Lüders front propagation and jerky flow under these circumstances would be interesting.

The development of a *neck* is, of course, a phenomenon restricted to tensile deformation. Nevertheless, physically similar processes can occur in compression and torsion. In compression, an accentuation of a strain-rate gradient that may have originally arisen by fluctuation, or from an area irregularity or a material heterogeneity, would *increase* the cross-sectional area and thus lead to a process called *barrelling*. While the severe barrelling frequently observed is strongly influenced by friction effects at the compression platens, the phenomenon can, in principle, occur by the very same mechanisms that led to localization of flow in tension. This was discussed in detail by Jonas and coworkers[50] for the case of substantial flow softening.[57] Even under mild hardening, though, where the decrease in stress with an increase in area tends to redistribute flow more uniformly, non-negligible strain gradients may arise in the transient.

For a quantitative treatment, we refer to the differential equation for the development of strain-rate gradients in its general form: eq. (96). The area changes on the right-hand side are zero for torsion, so that

$$\boxed{\frac{\partial}{\partial \epsilon}\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t + (\delta + b)\left.\frac{\partial \ln \dot{\epsilon}}{\partial x}\right|_t = 0 \qquad \text{(torsion)}. \qquad (107)}$$

The solutions are therefore identical to those in which we had neglected changes in geometry and in material constants: eqs. (34) and (36), and Fig. 10; only, δ_P there must here be replaced by $\delta + b$. The influence of b may be important at high temperatures, where the rate sensitivity of strain-hardening (and of M itself) is often large due to dynamic recovery. Then the growth or decay of any fluctuation would be slow, and large strain differences can be accumulated in the process. In steady state, deformation is uniform—so long as $b > 0$. Again, the rate sensitivity of strain hardening is crucial.

In compression, the right-hand side of eq. (96) has the opposite signs of those in tension. With

$$\delta_p = \frac{H + 1}{M} \qquad \text{(compression)} \qquad (108)$$

*Jonas and coworkers[50] have used precisely the opposite technique to treat compression: they changed the signs of stress, strain, and strain rate, and left the relation between area and strain intact. This has two disadvantages: the sign of the inequalities changes; and torsion cannot be treated at all.

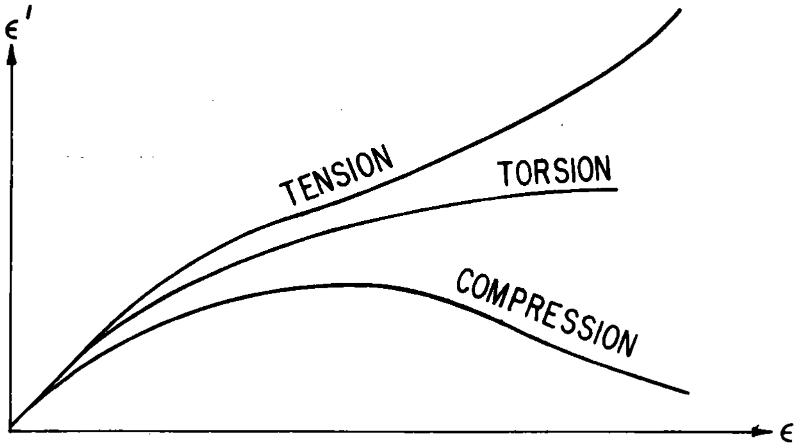

FIG. 28 Qualitative development of an initial strain-rate gradient near steady state.

we get

$$\boxed{\frac{\partial^2 Y}{\partial \epsilon^2} + (\delta_p + b)\frac{\partial Y}{\partial \epsilon} + cY = 0 \text{ (compression)}. \qquad (109)}$$

The important (and not unexpected) result is that the *restoring force* in the last term is now normally positive: area gradients must eventually vanish. (This is not so when c is negative, as may happen under the influence of severe simultaneous recovery or recrystallization at high temperatures, see Fig. 5.) Furthermore, the "damping term", the coefficient of Y', is here always positive. Thus, growth of nonuniformities can never be unbounded—but both growth and decay can be quite slow. Figures 24 and 25 demonstrated this effect under the assumption of a constant δ_P; the limiting value is 1 for steady-state compression.

Incorporating the effects of changes in material properties with strain, Fig. 28 shows the development of the strain gradient after an initial strain-rate gradient for tension (eq. (97)), torsion (eq. (107)), and compression (eq. (109)), in a qualitative way. While growth in tension is unbounded, it is slow but bounded in torsion, and in compression it reverses. The strain gradient at the maximum should be of the same order as the initial relative gradient in area, stress, or glide resistance, if these were what gave rise to the initial strain-rate gradient.

Sheet forming qualifies, under some circumstances, as deformation of alternative straining elements. In many practical applications, however, as well as in biaxial test configurations, the constraints exerted by the surrounding material cause important interactions. A discussion of *uniform* deformation in homogeneous (and isotropic) sheets is nevertheless useful as a baseline.

The case corresponding precisely to uniaxial tension is *plane-strain* deformation in a sheet, Fig. 29. When a very wide strip is subjected to tractions in one direction, x, a mode of localized deformation is possible that consists of thickness reduction only:[82] neighboring elements that do not deform restrain contraction in the lateral direction, y, (and cause a lateral tensile stress $\sigma_{yy} = \sigma_{xx}/2$). The thickness reduction is then the same as the total reduction in cross-sectional area. Its gradient may thus be immediately identified with the quantity Y defined in eq. (84), and the whole rest of the tension formalism can be applied. The only problem remaining is a proper identification of the equivalent stress and equivalent strain rate under these circumstances for any particular material; the von Mises yield

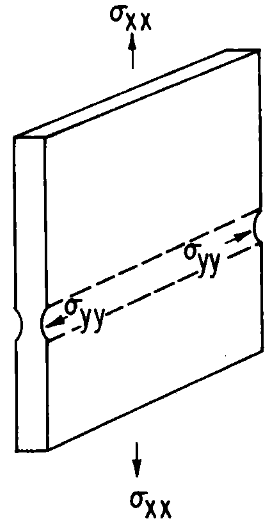

FIG. 29. The thickness reduction in a sheet under plane strain is equivalent to the area reduction in a tensile bar.

condition and the "flow rule" are usually adequate. The equivalence of tensile necking and plane-strain localization in sheets is experimentally confirmed: the critical strain is similar to that given by the Considère criterion.

A *narrow* strip under uniaxial tension behaves somewhat differently. Lateral strains are now not completely impeded (although they may be somewhat different from the thickness strains, due to the geometric anisotropy). In the limit that both are the same, the case is again identical to tension. It is here called *diffuse necking*—the reason being that the extent of the neck tends to be governed by the larger of the lateral dimensions, and is therefore much longer here than in the localized mode discussed for plane-strain deformation. It occurs at about the same strain.

A *localized* mode is also observed in the tension of narrow strips, but it occurs at a strain approximately double that for diffuse necking.[83] It is in fact a plane-strain mode, and for this reason occurs in a thin band that is oriented in the direction of nonextensional deformation (where the plastic contraction is counterbalanced by the length change due to rotation). Inasmuch as this mode invariably leads to failure even if diffuse necking does not, it is the localization of prime concern. Finally, a pure shear mode, which can be most sharply localized, exists at zero,[84] or sometimes small but finite,[12] hardening rate.

A more difficult situation arises when a sheet is deformed under *balanced biaxial tension*, or in fact in any direction in stress space in which the lateral principal stress exceeds half the longitudinal one; or equivalently, in which the lateral strain is *positive*. The observation here is that a localized mode occurs that looks very much like the one in plane strain: it is in a band perpendicular to the major stress axis. Explanations for this mode are still under serious dispute. Inasmuch as there is no nonextensional direction, the geometric effect that gave rise to "localized necking" in strip tension cannot be responsible. One set of explanations is critically based on the preexistence of defects, especially area defects.[1, 2] Since such defects are not necessary to explain the development of nonuniformities in tension, this assumption would seem to make an undue distinction between the two types of stress state. A treatment in terms of strain-*rate* defects has not yet been undertaken. Another set of proposed explanations relies on the development of vertices in the yield surface during the previous deformation.[3] Its physical basis is not well established. Its chief merit is that it does attempt to address the question of material behavior during changes of path in stress space, as they must occur in the biaxial regime for any localization to be possible.[85]

The latter problem may be more appropriately discussed in terms of a dependence of the strain-*hardening* rate on straining direction, or its sensitivity to *changes* in direction.[24] It has been shown in simple cases that changes in slip distribution tend to cause structural instabilities in the material. The resulting negative hardening rate could well be the cause of localization. Experiments that involve deliberate changes in straining path are, however, very difficult. A solution to this problem does therefore not seem to be in sight, even though it is of great technological importance.

5. Conclusions

5.1. *Summary*

All pieces of material that have alternative straining elements are prone to developing nonuniformities in deformation. These may be initiated by some defect in the specimen (either geometric or physical), or by some fluctuation in the strain rate, or as a consequence of an inherent material property. In particular, negative strain hardening and a negative rate sensitivity of the flow stress cause initial instabilities.

Whether deformation tends to localize or to spread after an initial disturbance, and how fast it does so, depends on the deceleration parameter δ: the (negative) slope in a plot of the logarithm of the strain rate vs. the strain, i.e. the quantity measured in primary creep. It is proportional to the work hardening rate, and inversely proportional to an appropriate rate sensitivity.

The material parameters, strain-hardening rate, deceleration rate, and rate sensitivity, are *local* properties. In the very cases where nonuniform deformation is possible, macroscopic measurements are not directly indicative of these *material* properties, but rather of *specimen* properties. For example, negative intrinsic hardening often leads to a plateau in the stress strain curve; conversely, a negative macroscopic stress strain characteristic may be indicative of homogeneous (and even locally decelerating) deformation. A negative rate sensitivity leads to macroscopically jerky flow and can be observed only indirectly. For these reasons, any analysis that attempts to describe material behavior in nonuniform deformation must be based on a careful distinction between local and macroscopic properties.

Two of the common modes of nonuniform deformation, jerky flow and the spreading of a Lüders front, are in fact *de*localizations after an initial instability caused by an inherent material property: negative rate sensitivity in the case of jerky flow, negative strain-hardening in the case of front propagation. Their very different behavior is due to one fundamental and one quantitative difference. In the Lüders case, the material is left, after the front has passed, in a state akin to having been deformed uniformly, and is thus ready to proceed in a homogeneous manner. In jerky flow, this is not the case. Furthermore, the absolute magnitude of the rate sensitivity in the first case is large, in the second small; as a result, the strain-rate discontinuity in jerky flow is too large to be accommodated by steady-state front propagation.

Necking, by contrast, is characterized by *increasing* localization of initially uniform deformation. It is unavoidable; its eventual rate of development is determined by the steady-state rate sensitivity. At an earlier stage, criteria of (instantaneous) instability provide only very limited guidance to material behavior. Rather, it is determined by the details of the initial conditions and by the magnitude of the deceleration parameter δ_P (at constant load). Even when it is positive, but small, differences in strain rate (which may have originated as a fluctuation) persist for substantial strains—and all during this time, the *strain* gradient grows. For this reason, the Considère criterion has been successful as an approximate demarcation line between negligible and severe localization. For this reason also, however, significant localization is possible in torsion and compression, even when the Considère condition is not exceeded.

The strain *rate* emerges as a crucial variable: its change with strain in each material element, its gradient along the specimen, and its influence on the flow stress. It alone makes both a qualitative and a quantitative treatment of the *development* of nonuniform deformation possible. Without rate dependent plasticity, only yes/no criteria can be derived, and they have been found severely wanting.

5.2. *New Results*

A number of suggestions have been made here for the first time. They will now be briefly summarized.

(1) The material parameters and constitutive laws have been formulated *ab initio* in an evolutionary way: by the rate of change of stress and of strain rate with strain. This has the advantage that path dependence is inherently allowed; only when it is unimportant can other relations be uniquely defined. (Section 2.3.)

(2) The differential equation governing neck development has been newly derived based on these evolutionary material parameters. This provides a generalization of the recently suggested[13] formulation. The differential equation is of second order in strain (or time) at every cross section, and thus allows for the necessary two branches of neck development, one bounded and one unbounded. It also describes the various initial defects (if any) by a single formalism. Finally, it shows that the geometric part of the problem, which had been the entire basis of previous treatments, is relatively unimportant when hardening is low, and especially when the rate sensitivity of hardening is high, as it frequently is. Then, even in compression tests where δ_P remains positive (unless there is severe recovery or recrystallization[56]), transient nonuniformities in deformation can be expected. As a general condition for low resistance to localization, independent of boundary conditions, the criterion $\delta \ll 10$ is proposed. (Sections 3.1, 4.3 and 4.4.)

(3) A distinction between two kinds of negative hardening has been suggested: a strain-triggered structural instability, which leads to transient localization at constant stress; and *homogeneous recovery* that occurs even where straining does not, and allows macroscopic negative hardening without instability. Classical "work softening" appears to be of the second type. (Section 3.3.)

(4) A more realistic approximation of material kinetics has been proposed for the case of Lüders front propagation: constant deceleration (logarithmic creep) for all strain rates above the prescribed average, and zero strain rate below. This gives a considerable improvement over the classical exhaustion type of kinetics, both in the prediction of the Lüders strain, and in the description of the front width. Creep at lower strains has been used to explain peculiarities at the end of the Lüders plateau. (Section 4.1)

(5) For the regime of jerky flow, the detailed shape of the material characteristic in the regime of negative rate sensitivity is used to explain the type and severity of serrations. A broad "valley" gives rise to a large strain-rate discontinuity and thus presumably to a high

velocity of front propagation and severe serrations. The depth of the valley determines the maximum value of the load drops and thus, through strain hardening, the strain per jerk. Machine interaction only controls the width of "hopping bands". (Section 4.2.)

(6) The stress relaxation test has been shown to be unresponsive to negative rate sensitivity, except in an extremely soft machine. However, the continuity of strain rate from immediately before to immediately after arrest of the cross-head should be broken when the rate sensitivity is negative. (Section 3.3.)

5.3. *Assessment*

We see the outstanding problems in the four major areas of localized flow as follows.

Lüders front propagation. The root cause of this phenomenon has long been known: a high generation stress. However, this does not control the level of the lower yield stress, which must be decreased if nonuniform deformation is to be minimized. While its relation to various other variables and material properties is now well described, the crucial question of the mechanistic cause has not been solved. Detailed measurements that include a deliberate variation of the rate sensitivity of the material would be of value.

Jerky flow. It appears to be well established that a negative intrinsic rate sensitivity is the cause of jerky flow, or at least a necessary condition. The reason for the strain dependence of the negative contribution to the rate sensitivity is yet under dispute; it was not a subject of this article. The mechanism that controls front propagation, and its connection to the *magnitude* of the negative rate sensitivity, is the most pressing unknown. It could explain the shape of serrations. The stress at which the rate sensitivity becomes zero is of prime importance; it may be the same as that at the bottom of the load drops in type-C serrations.

Necking. While its development is well understood in the regime of superplasticity, the conditions for the initiation of severe necking near the UTS are not sufficiently appreciated. A detailed investigation of specimen shape development, coupled with measurement of the material properties in a creep test under constant load, and with computer integration of the governing differential equation, should provide the necessary insight. A comparative study of barrelling in compression in this light might prove interesting.

A proper treatment of *sheet deformation* in the biaxial stress regime awaits a detailed investigation of constitutive behavior under deliberate path changes. The behavior of the strain-hardening rate after such path changes should be of particular interest in this context, not only the change in size and shape of the yield surface and plastic potential.

In general, we find that the influence of *rate sensitivity* in the material description has been unduly neglected in the past. Inasmuch as it is, to some degree, controllable by metallurgical variables, it could provide a key to performance improvement.

REFERENCES

1. Z. MARCINIAK, *Mechanics of Sheet Metal Forming* (eds.) D. P. KOISTINEN and N–M. WANG, p. 215. Plenum Press (1978).
2. J. W. HUTCHINSON and K. W. NEAL, ibid. p. 127; *Acta Met.* **25**, 839 (1977).
3. A. NEEDLEMAN and J. R. RICE, *Mechanics of Sheet Metal Forming* (eds.) D. P. KOISTINEN and N-M. WANG, p. 237. Plenum Press (1978).
4. S. S. HECKER, *Formability: Analysis, Modeling, and Experimentation* (eds.) HECKER, GHOSH and GEGEL, p. 150. Met. Soc. AIME (1978).
5. I. L. DILLAMORE, *Textures of Materials* (eds.) G. GOTTSTEIN and K. LÜCKE, p. 67. Springer (1978).
6. A. S. ARGON, *The Inhomogeneity of Plastic Deformation* (ed.) R. E. REED-HILL, p. 161. Am. Soc. Metals (1973).
7. A. CONSIDÈRE, *Ann. Ponts Chaussées* **9/6**, 574 (1885).

8. J. D. CAMPBELL, *J. Mech. Phys. Sol.* **15,** 359 (1967).
9. E. W. HART, *Acta Metall.* **15,** 351 (1967).
10. J. KLEPACZKO, *Int. J. Mech. Sci.* **10,** 296 (1968).
11. E. DUNCOMBE, *Int. J. Solids Struct.* **10,** 1445 (1974).
12. R. J. ASARO, *Acta Metall.* **27,** 445 (1979).
13. U. F. KOCKS, J. J. JONAS and H. MECKING, *Acta Metall.* **27,** 419 (1979).
14. F. A. NICHOLS, *Acta Metall.* **28,** 663 (1980).
15. A. H. COTTRELL and B. A. BILBY, *Proc. Phys. Soc. London* **A62,** 49 (1949).
16. W. SYLVESTROWICZ and E. O. HALL, *Proc. Phys. Soc. London,* **B64,** 495 (1951).
17. B. J. BRINDEY and P. J. WORTHINGTON, *Metall. Rev.* **15,** 101 (1970).
18. P. G. MCCORMICK, *Acta Metall.* **20,** 351 (1972).
19. A. VAN DEN BEUKEL, *Phys. Status Solidi* **a30,** 197 (1975).
20. A. KORBEL, J. ZASADZINSKI and Z. SIEKLUCKA, *Acta Metall.* **24,** 919 (1976).
21. E. W. HART, *Acta Metall.* **3,** 146 (1955).
22. A. W. SLEESWYK, *Acta Metall.* **6,** 598 (1958).
23. P. PENNING, *Acta Metall.* **20,** 1169 (1972).
24. J. L. DUNCAN and W. JOHNSON, *Int. J. Mech. Sci.* **10,** 143 (1968).
25. J. C. FISHER and H. C. ROGERS, *Acta Metall.* **4,** 180 (1956).
26. J. M. KRAFFT, *Acta Metall.* **10,** 85 (1962).
27. D. J. VEREL and A. W. SLEESWYK, *Acta Metall.* **21,** 1087 (1973).
28. N. BALASUBRAMANIAN and J. C. M. LI, *Mater. Sci. Eng.* **14,** 37 (1974).
29. F. W. NOBLE and D. HULL, *Acta Metall.* **12,** 1089 (1964).
30. E. PINK, *Mater. Sci. Eng.* **10,** 279 (1972).
31. J. D. PENROSE and A. F. BROWN, *J. Iron Steel Inst.* **209,** 862 (1971)
32. S. H. VAN DEN BRINK, A. VAN DEN BEUKEL, and P. G. MCCORMICK, *Phys. Status. Solidi* **a30,** 469 (1975).
33. M. MAYER, O. VÖHRINGER and E. MACHERAUCH, *Phys. Status Solidi* **a49,** 473 (1978).
34. R. A. MULFORD and U. F. KOCKS, *Acta Metall.* **27,** 1125 (1979).
35. R. A. MULFORD, *Metall. Trans.* **10A,** 1527 (1979).
36. L. J. CUDDY and W. C. LESLIE, *Acta Metall.* **20,** 1157 (1972).
37. D. A. WOODFORD, *Trans. Am. Soc. Met.* **62,** 291 (1969).
38. J. HEDWORTH and M. J. STOWELL, *J. Mater. Sci.* **6,** 1061 (1971).
39. G. L. DUNLOP and D. M. R. TAPLIN, *J. Mater. Sci.* **7,** 84 (1972).
40. G. RAI and N. GRANT, *Metall. Trans.* **6A,** 385 (1975).
41. A. K. GHOSH and R. A. AYRES, *Metall. Trans.* **7A,** 1589 (1976).
42. C. G'SELL and J. J. JONAS, *J. Mater. Sci.* **14,** 583 (1979).
43. C. S. HARTLEY and D. A. JENKINS, *J. Metals* **32,** 23 (1980).
44. U. F. KOCKS, *Constitutive Equations in Plasticity* (ed.) A. S. ARGON, p. 81. MIT Press (1975).
45. A. P. L. TURNER, *Metall. Trans.* **10A,** 225 (1979).
46. J. C. SWEARENGEN, R. W. ROHDE and D. L. HICKS, *Acta Metall.* **24,** 969 (1976).
47. U. F. KOCKS, *Strength of Metals and Alloys* (eds.) P. HAASEN, V. GEROLD and G. KOSTORZ, p. 1661. Pergamon Press (1979).
48. E. W. HART, *Acta Metall.* **18,** 599 (1970).
49. H. MECKING and U. F. KOCKS, *Acta Metall.*, in press (1981).
50. J. J. JONAS, R. A. HOLT and C. E. COLEMAN, *Acta Metall.* **24,** 911 (1976).
51. Y. TOZAWA, *Mechanics of Sheet Metal Forming* (eds.) D. P. KOISTINEN and N-M. WANG, p. 81. Plenum Press (1978).
52. U. F. KOCKS and T. J. BROWN, *Acta Metall.* **14,** 87 (1966).
53. J. WASHBURN and G. MURTY, *Can. J. Phys.* **45,** 523 (1967).
54. Z. S. BASINSKI and P. J. JACKSON, *Phys. Status Solidi* **9,** 805 (1965).
55. Z. S. BASINSKI and P. J. JACKSON, *Phys. Status Solidi* **10,** 45 (1965).
56. U. F. KOCKS, A. S. ARGON and M. F. ASHBY, *Prog. Mater. Sci.* **19,** 1 (1975).
57. J. J. JONAS and M. J. LUTON, *Advances in Deformation Processing* (eds.) J. J. BURKE and V. WEISS, p. 215. Plenum (1978).
58. U. F. KOCKS, *J. Eng. Mater. Tech. (ASME H)* **98,** 76 (1976).
59. H. MECKING, *Work Hardening in Tension and Fatigue* (ed.) A. W. TOMPSON, p. 67. Met. Soc. AIME (1977).
60. A. KELLY and R. B. NICHOLSON, *Prog. Mater. Sci.* **10,** 149 (1963).
61. M. S. WECHSLER, *The Inhomogeneity of Plastic Deformation* (ed.) R. E. REED-HILL, p. 19. Am. Soc. Metals (1973).
62. F. A. MCCLINTOCK, Mass. Inst. Tech, personal communication.
63. R. B. SCHWARZ, *Dislocation Modelling of Physical Processes* (eds.) C. S. HARTLEY *et al.*, in press. Pergamon (1981).
64. J. FRIEDEL, *Les Dislocations,* Gauthier-Villars (1956).
65. U. F. KOCKS and H. MECKING, *Dislocation Modelling of Physical Processes* (eds.) C. S. HARTLEY *et al.*, in press. Pergamon (1981).

66. P. Wycliffe, U. F. Kocks, and J. D. Embury, *Scripta Metall.* **14,** (Dec. 1980).
67. P. Feltham, *Philos. Mag.* **6,** 847 (1961).
68. E. W. Hart and H. D. Solomon, *Acta Metall.* **21,** 295 (1973).
69. H. Fischer, Doctoral thesis, Technische Hochschule Aachen (1975).
70. A. W. Sleeswyk, Groningen, personal communication.
71. C. Y. Cheng and U. F. Kocks, *Sec'd. Int'l. Conf. Strength of Metals and Alloys*, p. 199. Am. Soc. Metals (1970).
72. R. B. Schwarz and J. W. Mitchell, *Phys. Rev.* **B9,** 3292 (1974).
73. W. G. Johnston, *J. Appl. Phys.* **33,** 2716 (1962).
74. J. D. Embury, McMaster University, personal communication.
75. A. H. Cottrell, *Dislocations and Plastic Flow*, Clarendon Press (1953).
76. E. N. da C. Andrade, *Proc. R. Soc.* **A84,** 1 (1910).
77. G. R. Piercy, R. W. Chan and A. H. Cottrell, *Acta Metall.* **3,** 331 (1955).
78. G. T. Hahn, *Acta Metall.* **10,** 727 (1962).
79. P. Haasen, *Physical Metallurgy* (ed.) R. W. Cahn, p. 821. North Holland (1965).
80. A. Korbel and A. Pawelek, *Dislocation Modelling of Physical Processes* (eds.) C. S. Hartley *et al.*, in press, Pergamon (1981).
81. J. M. Jalinier, N. Christodoulou, B. Baudelet and J. J. Jonas, *Formability: Analysis, Modeling, and Experimentation* (eds.) S. S. Hecker, A. K. Ghosh and H. L. Gegel, p. 29. Met. Soc. AIME (1978).
82. A. N. Needleman, *Formability: Analysis, Modeling, and Experimentation* (eds.) Hecker, Ghosh and Gegel, p. 54. Met. Soc. AIME (1978).
83. P. B. Mellor and A. Parmar, *Mechanics of Sheet Metal Forming* (eds.) D. P. Koistinen and N–M. Wang) p. 53. Plenum Press (1978).
84. A. K. Chakrabarti and J. W. Spretnak, *Metall. Trans.* **6A** 733, 737 (1975).
85. J. L. Duncan, McMaster University, personal communication.