ADVANCED ENGINEERING MATERIALS
www.aem-journal.com

Profilometry-Based Inverse Finite Element Method Indentation Plastometry

Trevor William Clyne, Jimmy Edward Campbell, Max Burley, and James Dean*

This is a review of the current state-of-the-art regarding a particular approach to extraction of the (quasistatic) stress–strain relationship of a metallic sample from an indentation experiment. It is based on the application of a relatively high load (kN range) to the sample via a large spherical indenter (≈1 mm radius), followed by measurement of the indent profile. This profile is then used as the target outcome for inverse finite element method (FEM) modeling of the test, aimed at converging on the best fit set of parameter values in a constitutive plasticity law (true stress–true strain relationship). This can then be used to simulate any specified loading configuration, including a conventional tensile test. Commercial products are now available in which the indentation, profilometry, and convergence operations are all automated and completed within a few minutes. The review covers the various conceptual and practical issues involved in implementation and optimization of these procedures, including both those related to the measurement system (experimental and FEM simulation) and those associated with the sample (such as anisotropy, inhomogeneities, and residual stresses). An attempt is made to convey an impression of the expected levels of reliability and also the scope for obtaining insights that are not readily obtainable using other types of test.

1. General Introduction

1.1. Historical Background

The plastic deformation of metals and attempts to study it experimentally have been important for millennia. Systematic investigation started in the 19th century, although it was not until 1900 that the first well-defined hardness test was proposed (by Brinell). This involved pushing a hard (steel) ball into a sample with a specified force and measuring the diameter (in projection) of the resultant indent. The hardness number was defined as the force over the contact area (obtained from the measured diameter). This does give an indication of the resistance that the material offers to plastic deformation, but it is not a well-defined material property—being dependent on the size and shape of the indenter, and the applied load. Furthermore, it is not uniquely related to the fundamental plasticity characteristics, which may be regarded as the yield stress and the way that the post-yielding stress level changes with strain (the "work hardening" behavior): materials with widely varying combinations of these can record the same hardness value.

Several other hardness tests were developed during the 20th century, as was the uniaxial tensile test. The simple stress and strain fields (within the gauge length) during a tensile test allow stress–strain curves to be readily obtained, albeit at a cost in terms of the practical difficulties involved in producing suitably shaped samples, gripping them, and measuring the strain in the gauge length. This stress–strain curve (when expressed as true values of both) does constitute a fundamental representation of how the material behaves during plastic deformation. It is central to prediction of how components behave in practice, with widespread use of the finite element method (FEM) greatly facilitating this. A longstanding aim has thus been to obtain this information using the much simpler and more convenient operation of pushing a hard indenter into a free surface.

The introduction of indentation systems in which both load and penetration depth are monitored, creating richer data sets, led to hopes that they could be used to extract stress–strain relationships. Most such developments also involved refinement of the scale of the test. From the start, instrumented indentation was often conducted with loads in the mN range and depths in the nm range—such that they were described as nanoindenters. This had various attractions, but it has complicated the development of a capability for extracting reliable stress–strain curves. This is mainly because of "size effects"—a term used to describe the observation that small regions do not, in general, exhibit the same deformation characteristics as the bulk. It often relates to grain structure—and the need to deform a suitably large assembly of grains—although size effects can also arise from other causes. The topic is addressed in Section 1.3.

T. W. Clyne, J. E. Campbell, M. Burley, J. Dean
Plastometrex Ltd
204 Science Park, Milton Road, Cambridge CB4 0GZ, UK
E-mail: twc10@cam.ac.UK

T. W. Clyne
Department of Materials Science
27 Charles Babbage Road, Cambridge CB3 0FS, UK

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adem.202100437.

DOI: 10.1002/adem.202100437

Leaving aside for the moment this issue of scale, two main conceptual approaches have been adopted for extraction of stress–strain curves from indentation data. The first, commonly termed the "Instrumented Indentation Technique" (IIT), involves converting load–displacement data directly to stress–strain curves, using analytical relationships. This approach has been very popular,[1–14] and is attractive in terms of quickly and easily obtaining the final outcome, but it involves gross simplifications concerning the actual stress and strain fields under an indenter. A slight variant of the concept involves the use of neural network procedures[15–17] to relate load–displacement data to corresponding stress–strain curves—i.e., to "train" the analytical relationship, although in practice this is subject to similar limitations. In fact, it has become clear[18] that the reliability of the IIT approach is in general very poor. This largely arises from the complexity of the evolving stress and strain fields during indentation, which cannot be analytically linked to the load–displacement data in an accurate, universal way. This applies equally to neural network approaches, which also seek an empirical functional relationship of some sort.

The second is a more rigorous approach, although inevitably more cumbersome. It involves[19–25] repeated FEM simulation of the test, altering the values of the parameters in a constitutive plasticity law until optimum agreement is reached between a measured and a modeled outcome—either the load–displacement plot or the residual indent profile. The current review is focused on the latter (leading to the term "Profilometry-based Inverse FEM Indentation Plastometry"—PIP), which offers certain advantages[26–29] over use of the load–displacement plot. In summary, these include removal of the need for any measurements during loading, elimination of uncertainties associated with machine compliance, scope for detection and analysis of (in-plane) anisotropy, and an improved sensitivity of the measured outcome to the shape of the stress–strain curve. More detail is provided in Section 1.4.

The main conceptual distinction, however, is between direct conversion of load–displacement data to a stress–strain curve (IIT) and iterative FEM to converge on optimal values of parameters in a constitutive law (whether the target outcome for this operation is the load–displacement curve or the residual indent shape). Iterative FEM simulation clearly has the potential to fully capture the nature of the evolving stress and strain fields during the test, whereas this is not possible with the IIT approach.

1.2. Constitutive Laws and FEM Model Formulation

For any approach involving repeated FEM modeling of a plastic deformation process, the true stress–strain relationship (material plasticity response) must be characterized via a (small) set of parameter values. Several analytical expressions have been proposed to characterize the work hardening of metals, but only two are in frequent use. The first is the Ludwik–Hollomon (L—H) equation[30]

$$\sigma = \sigma_Y + K\varepsilon^n \tag{1}$$

where σ is the (von Mises) applied stress, σ_Y is its value at yield, ε is the plastic (von Mises) strain, K is the work hardening coefficient, and n is the work hardening exponent. The second is the Voce equation[31]

$$\sigma = \sigma_S - (\sigma_S - \sigma_Y)\exp\left(\frac{-\varepsilon}{\varepsilon_0}\right) \tag{2}$$

The stress σ_s is a "saturation" level, while ε_0 is a characteristic strain for the exponential approach of the stress toward this level. When implemented in an FEM model, these stresses and strains are von Mises (deviatoric) values, with the hydrostatic components of the stress and strain tensors playing no role in the plasticity response. The range of shapes obtainable by changing the values of the parameters is illustrated by the plots in **Figure 1**. It can be seen that a wide variety of curve shapes can be obtained.

Of course, the true stress–strain curve of a metal is not in fact expected to conform exactly to any analytical relationship. In choosing between formulations such as these two, the key question is which of them is likely to capture most effectively the actual response of a cross section of metals. In practice, there is often a tendency for a plateau to be approached, corresponding mechanistically to the rate of "hardening" due to creation of dislocation entanglements etc. becoming balanced by the rate of "softening" due to processes such as cross-slip and climb of dislocations, which causes them to get annihilated and become less entangled. (This type of balance being approached should not be confused with plateau formation in a nominal stress–strain curve, which can occur even when the true stress–strain curve continues to exhibit significant work hardening.) Of the two relationships shown in Figure 1, Voce is better suited to capturing cases in which the (true) work hardening rate (WHR) does approach zero. Partly for this reason, it is starting to come into wider use than L—H for PIP testing. Some detail about this is provided in Section 2.1, where mention is also made of the relatively small number of cases that cannot be well captured by any expression for which the work hardening rate (gradient of the true stress–true strain curve) monotonically decreases.

Other issues relating to the FEM simulation include those concerning the extent of the domain and the details of the mesh. There are excellent review articles available,[32,33] which cover

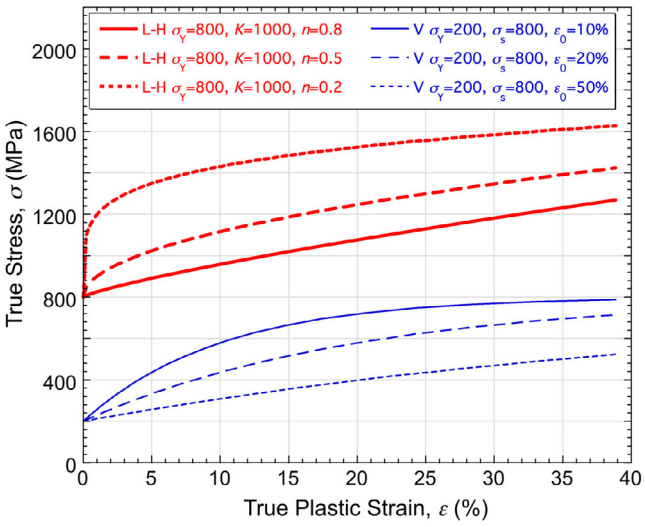

Figure 1. True stress–strain curves in the form of L—H and Voce plots, obtained using Equation (1) and (2), with the parameter values shown in the legend.

them in the context of materials modeling. As with any FEM model, a mesh (with particular types of volume element, a domain, and a set of boundary conditions) needs to be specified. Sensitivity analyses should be run to confirm that the mesh is sufficiently fine to achieve convergence, numerical stability, and mesh-independent results. This usually requires it to be quite fine, particularly in areas where plastic strains will be high. Information specific to indentation modeling is available in the literature.[23,29,34] An example is shown in **Figure 2** of a mesh suitable for spherical indentation. There are about 5000 volume elements in this case, all second-order quadrilateral and/or triangular. Of course, provided both sample and indenter are transversely isotropic, this is a 2D (radially symmetric) mesh.

The (lateral and through-thickness) extent of the domain requires attention. In fact, the sample thickness becomes unimportant beyond a thickness of about 2 indenter diameters, although this does depend on the penetration ratio, δ/R—i.e., the final depth below the original surface level divided by the radius of the indenter. Beyond that thickness, the axial stress—and hence the elastic strain—becomes negligible. The lateral extent, beyond about 2 indent diameters, is also unimportant. Effectively, a semi-infinite system is being modeled. Of course, it is important that the experimental arrangement should also be effectively semi-infinite (if a universal, rather than customized, modeling capability is being used). This is relevant, for example, to the testing of thin sheet. If a 1 mm diameter ball is used, and $\delta/R \approx 20\%$, then the minimum sheet thickness is of the order of 1—1.5 mm. Similar arguments apply to the lateral extent—i.e., to the presence of a free edge near the indent (for both mounted and unmounted samples). Again, for a 1 mm diameter ball, the indent axis should be no closer to such an edge than about 2 mm.

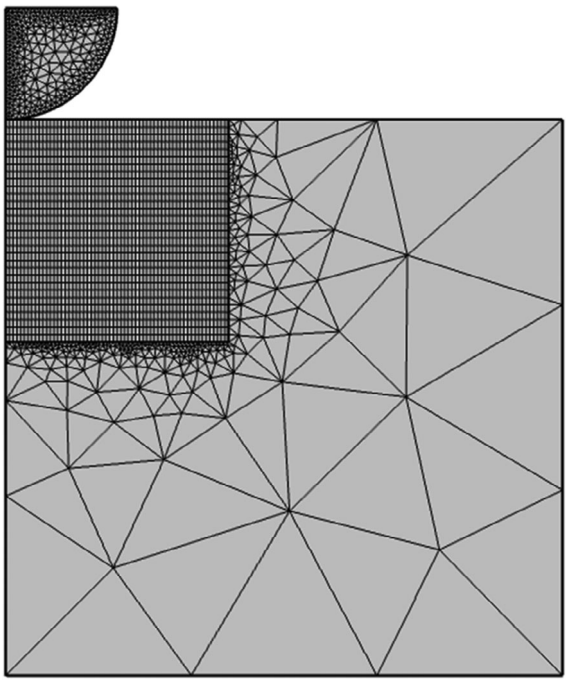

Figure 2. A typical (radially symmetric) FEM mesh for simulation of (spherical) indentation.

A further point concerning boundary conditions relates to interfacial friction.[35–37] This is usually characterized via a fixed value of a coefficient of friction, μ. There is information available in the literature about likely values of μ under different conditions.[38–42] Concerning indentation with a large sphere specifically, the surfaces of both indenter and sample are normally quite smooth and the value of μ is expected to be fairly low—probably somewhere in the approximate range of 0.1–0.2. In practice, the exact value does not have a very strong effect on outcomes, although neglect of friction ($\mu = 0$) will lead to interfacial sliding occurring throughout the test, which is rather unrealistic.

1.3. Indentation Geometry, Representative Volumes, and Plastic Strain Levels

The choice is between a sphere and a more complex ("sharp") shape. An attraction of sharp indenters is that they tend to create indents with large diameters—i.e., the load requirements for creating indents large enough to measure easily are reduced. In practice, however, there are strong incentives to use spherical indenters. These have been highlighted in several publications.[21–23,43,44] For example, a sphere is much less prone to becoming damaged than are shapes having edges or points. It is also easier to specify and manufacture. Furthermore, there is reduced risk with spheres of encountering the computational problems that can arise with simulation of behavior in regions of high local curvature (edges or points). Finally, at least with (approximately) isotropic materials, a spherical indenter allows the FEM modeling to be radially symmetric (2D), which is not possible with most shaped indenters.

While modeling of the indentation process is effectively scale-independent, there is nevertheless an important issue concerning the size of the plastically deformed volume. For any type of test, it must be large enough for its response to reflect that of the bulk (if, indeed, it is bulk properties that are being sought). For a typical (polycrystalline) metal, this usually translates into a requirement for a "many-grained" volume to be deformed because the plastic response of the bulk is influenced by characteristics such as grain size and shape, crystallographic texture, and grain boundary structure (influencing the ease of intergranular sliding and grain rotation). Indenting within a single grain, or even deforming a small assembly of grains, is unlikely to allow the stress–strain relationship of the bulk to be accurately obtained. Grain size is commonly in the range of a few microns up to a few hundred microns, so creation of indents that are submicron in depth, with diameters of no more than a few microns, does not in general involve deformation of a representative volume.

For a (typical) grain size of around 100 μm, an indenter diameter of 2 mm, and a penetration depth of 200 μm ($\delta/R = 20\%$), part of the free surface around an indent (with a diameter of around 1 mm) is shown in **Figure 3**. It can immediately be seen that the deformed volume contains "many" grains—actually something like 10^4 grains in the region that has been deformed to plastic strains beyond a few %. In practice, any number of grains beyond a few tens is probably acceptable, so that the maximum (equiaxed) grain size is around 500 μm. Grain sizes larger than this are sometimes found, particularly in castings, but these

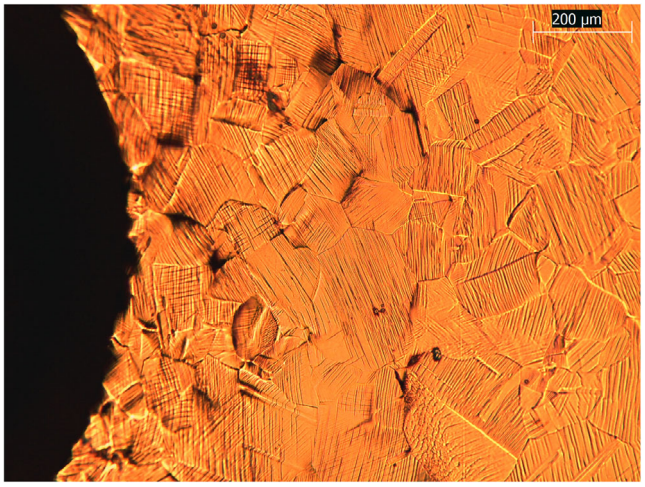

Figure 3. Optical micrograph of the free surface around an indent created in a copper sample by a spherical indenter with a radius of 1 mm.

constitute only a small fraction of commonly encountered metallic components. Of course, if grains are highly elongated in particular directions, then this could cause certain complications—often tied in with the possibility of texture (Section 3.2).

Several other features are often apparent in micrographs such as that of Figure 3. Some of these convey information about the mechanisms involved in the plastic deformation. For example, the sets of parallel lines within the grains are persistent slip bands—surface steps with heights of a few microns, created by many dislocations in a particular slip system gliding to the free surface. (An example is given in Section 3.2 of the appearance of these in a single crystal, where they can be related to the specific crystallography.) It can be seen that multiple slip systems were operative in most of the grains, particularly those closer to the indent. Careful inspection of Figure 3 also reveals the presence of both prior annealing twins (with parallel sides) and deformation twins that have formed within grains. There is also evidence of grain rotations that have occurred during the deformation, both around the "rim" of the "crater" and in the form of intergranular depressions on the free surface.

In addition to the requirement to deform a volume that is large enough to be representative of the bulk, it is important that the plastic strains being induced in the deformed region should be in an appropriate range. In most cases, interest centers on a range up to several tens of % (because these are levels that commonly occur prior to fracture of most metallic materials). A test in which peak strains are well below this range cannot be used to obtain reliable information about the plasticity characteristics of interest. This is a different requirement from that of the volume (relative to the grain size). For a spherical indenter, the distribution of plastic strain depends on the δ/R ratio, with the stress and strain fields during indentation being independent of the absolute scale. The distribution of strain also depends on the plasticity characteristics of the sample, but in general it is found that a δ/R value of about 20% produces a suitable strain distribution. This is illustrated by **Figure 4**, which shows fields of equivalent (von Mises) plastic strain for this value, for two materials (having the nominal stress–strain curves shown). In both cases, strains of up to several tens of % are created (although these are higher, and less diffuse, in the as-received material, which exhibits less work hardening).

This issue can be taken a little further because it is not simply a question of the peak strain level. If this level had been created only in a very small volume, then the influence on the overall outcome of that part of the stress–strain curve would be very limited. The real requirement here is an insight into how much of the deformation that affects the indentation response takes place in different ranges of strain. Descriptions of how such distributions of plastic strain can be obtained from an FEM simulation are available in the literature.[23,45] Outcomes of analysis of this type are shown in **Figure 5**, with Figure 5a showing fractions of the plastic work done, as a function of the strain at which it took place, in the (annealed Cu) material of Figure 4, for three different δ/R ratios. Also shown (Figure 5b) is a plot of the average strain at which the plastic work was done, for several metals, as a function of δ/R. These average strains tend to be lower, at a given δ/R, for metals that exhibit greater work hardening. In general, however, a δ/R value of about 20% ($\delta = 200\,\mu m$ in Figure 5b) corresponds to a weighted average plastic strain of around 10—15%, which is considered suitable.

1.4. Target Outcomes and Convergence in Parameter Space

As mentioned earlier (Section 1.1), either the load–displacement plot or the residual indent profile can be used as the target outcome, but the latter offers several advantages over the former and the present review is largely focused on its use. These

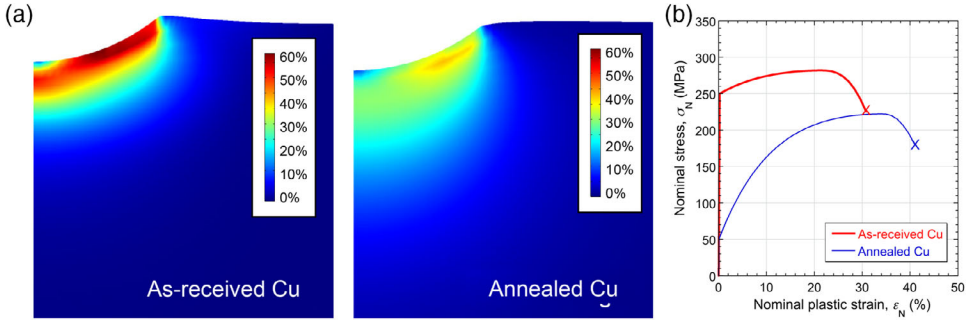

Figure 4. a) Fields of equivalent plastic strain after penetration to a depth ratio, δ/R, of 20%, for as-received and annealed copper samples, having the (nominal) tensile stress–strain curves shown in (b), which correspond to Voce parameter values of 250 MPa, 460 MPa, and 40% (as-received) and 50 MPa, 350 MPa, and 17% (annealed).

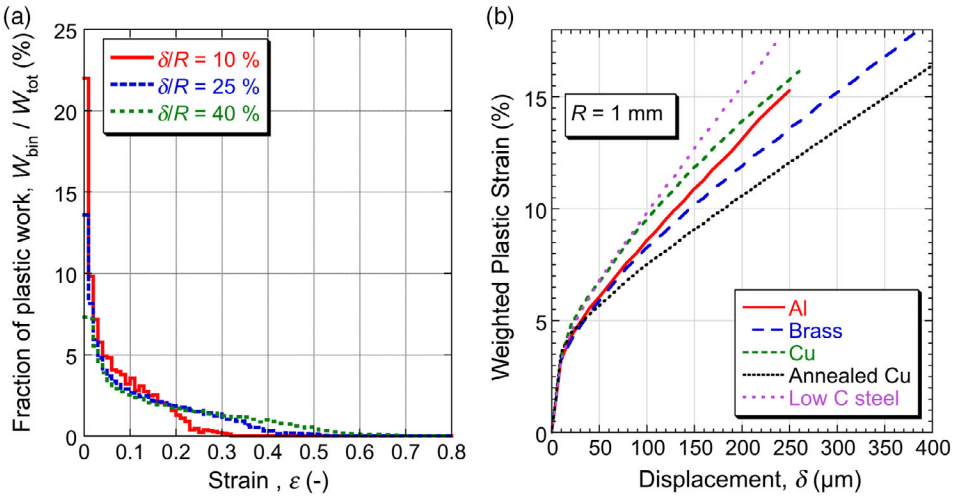

Figure 5. a) Fractions of the total work done in different ranges of strain during penetration of annealed Cu (Figure 1.4b) to three different depths and b) average plastic strains, for several materials, as a function of penetration.[23]

advantages include the fact that no measurements need to be made during the test (apart from noting the final load). If, on the contrary, the load–displacement plot is being used as the target outcome, then (accurate) displacement measurement must be made during the test: this might well require some kind of compliance calibration. Furthermore, use of the (3D) residual indent profile offers potential for the detection and characterization of plastic anisotropy in the sample (Section 3.2), whereas this is not possible if the focus is on load–displacement. Finally, there is an option to measure the profile at more than one depth, which provides additional experimental data and could, for example, be useful in detecting (sharp) variations in material properties with depth. Of course, using the indent profile as target outcome does not preclude the acquisition of load–displacement data, which might be useful for process control—for example, in ensuring that the test is conducted to a particular penetration depth or applied load.

There is then the issue of how the profile is measured. There are two main approaches, one involving (very low load) contact of a stylus that is dragged across the sample and the other based on (noncontact) scanning of an optical beam of some sort. Both exhibit various characteristics that are relevant to this application, with a wider range of types being available for optical measurement. Reviews are available for both stylus[46] and optical[47] systems. It is important to note that, as these measurements are being made on a relatively coarse scale (with typical indent depths of the order of 100—200 μm), very high resolution is not required. An accuracy of around ±1 μm is generally sufficient. It is therefore not appropriate to use Atomic Force Microscopy (AFM) systems or highly accurate optical techniques—partly because these tend to have limited range. Focus on this scale also means that surface roughness is not a critical issue and sample preparation can be relatively crude. Furthermore, surface contamination, oxide films, and so on are unlikely to cause difficulties. It is also worth noting that having no requirement for highly accurate measurement tends to mean that the cost of the equipment can be kept relatively low. This is an additional benefit of working at relatively coarse scale to that of the need to deform a representative volume (Section 1.3).

The postindentation convergence operation is likely to be based on minimization of a parameter characterizing the mismatch between measured and modeled outcomes (indent profiles). If radial symmetry can be assumed—i.e., if the sample is transversely isotropic—then there is just a single profile to consider. There are various options when defining such a misfit parameter for the case of two profiles, but a common approach,[23] is based on the sum of the squares of the residuals, S

$$S = \sum_{i=1}^{N} (\delta_{i,\mathrm{M}} - \delta_{i,\mathrm{E}})^2 \tag{3}$$

where $\delta_{i,\mathrm{M}}$ is the ith value of the modeled displacement (predicted by FEM) and $\delta_{i,\mathrm{E}}$ is the corresponding experimental (target) value. A typical value employed for N might be around 50, covering a range of radial locations up to about twice the indenter radius—i.e., the range over which the height has detectably changed from the baseline. Perfect fit will lead to a value of zero for S. However, with this definition, S is dimensional and has units, so its magnitude cannot give a universal indication of the quality of the fit. For this purpose, the quantity S_{red}, a "reduced sum of squares" is used, defined by

$$S_{\mathrm{red}} = \frac{\sum_{i=1}^{N} (\delta_{i,\mathrm{M}} - \delta_{i,\mathrm{E}})^2}{N\delta_{\mathrm{av,E}}^2} \tag{4}$$

where $\delta_{\mathrm{av,E}}$ is the average of the experimentally measured displacements (across the range of radial locations being scanned). The parameter S_{red} is a positive dimensionless number, with a value that ranges upward from 0 (corresponding to perfect fit). As a generalization, modeling that captures the material plasticity response reasonably well should lead to a solution (set of parameter values) for which S_{red} is less than, say, 10^{-3}. An optimized solution with a poorer fit than this suggests experimental deficiencies and/or an inability to capture the behavior well with

the constitutive law being used. In practice, S_{red} values around 10^{-4} or below are commonly obtained.

In practical terms, it is important that an efficient algorithm be used for convergence on the optimal combination of parameter values in the constitutive law concerned. Of course, such scenarios are commonly encountered in various branches of science and engineering and several mathematical procedures have been proposed. An example of this is the Nelder–Mead algorithm, with the details of its application to this specific problem being fully described in the literature.[23,45] A typical example of its operation is shown in **Figure 6**, which refers to a particular metal (using the L—H constitutive law and a load–displacement plot as the target outcome). As mentioned earlier, the current state-of-the-art is likely to be based on the Voce law and on using a residual indent profile as target outcome, but this convergence operation is nevertheless illustrative. It can be seen that about 50 runs of the FEM model were required for convergence, with the S_{red} value falling to below 10^{-4} by that point. Of course, a related issue is the time required to reach this point. In practice, this depends not only on the selected mesh and the computational power available, but also on the potential use of various mathematical devices designed to speed up the procedure. It is certainly desirable, for a commercial setup (Section 1.5), to ensure that the total time required for indentation, measurement, and data processing is no more than a few minutes.

1.5. Integrated Facilities

An integrated facility is clearly required, in which experimental measurements and FEM simulations (both repeated running of the indentation model and subsequent simulation of a tensile test) are under some kind of automated software control. In fact, packages of this type are now starting to become commercially available and one such has been used to acquire some of the results presented here. This is the Plastometrex (PLX) Indentation Plastometer, which is shown in **Figure 7**. It has an integrated stylus profilometer, which is actuated automatically after completion of the indentation procedure. The indenters are WC—Co cemented carbide spheres of either 1 or 2 mm diameter (interchangeable). Further technical details are available on the PLX website (www.plastometrex.com). Similar systems are also available from other suppliers.

2. Operation of an Indentation Plastometer

2.1. Basic Operation

It should first be noted that the elastic constants of the material (Young's modulus and Poisson ratio) are required for the FEM simulation. While the focus of the whole operation is on plasticity, elastic deformation inevitably takes place during the test and does affect the outcome (load–displacement data and residual

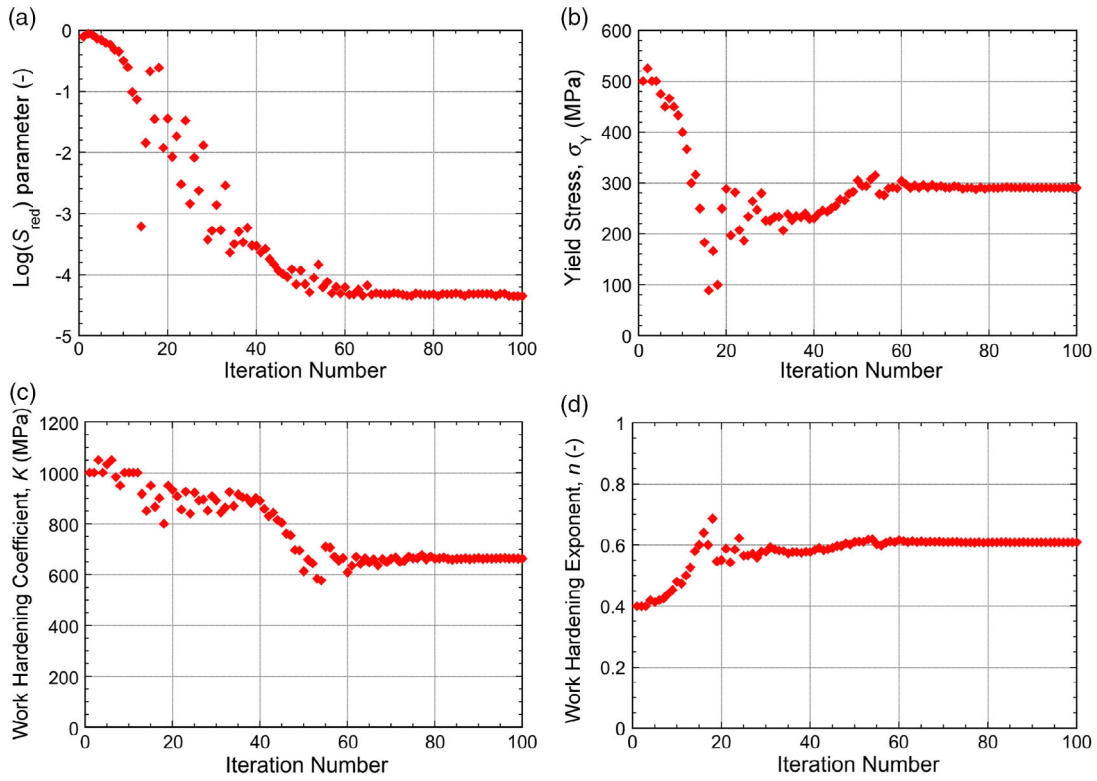

Figure 6. Nelder–Mead convergence[23] on an optimal (L—H) parameter set, targeting a load–displacement plot from indentation of a brass sample, showing the evolution with iteration number of: a) the misfit parameter, S_{red}, b) yield stress, c) work hardening coefficient, and d) work hardening exponent.

Figure 7. Image of the PLX Plastometer, with the key components labeled.

indent shape). However, the sensitivity of the outcome to these elastic constants is relatively small. It is therefore not necessary to specify them accurately—something like ±10–15% should be adequate. Details are provided below (Figure 14 in Section 2.3). Furthermore, elastic constants exhibit very little sensitivity to microstructure (unlike plasticity). They are dictated by interatomic forces and therefore depend essentially just on the types of atom present. For example, all ferrous alloys have very similar elastic constants because Fe is the predominant atom present in all cases, despite the fact that their plasticity characteristics cover a huge range. Selection of the pair of elastic constants from a small set of possibilities is usually appropriate.

While it is common to specify just two elastic constants, which is all that is required for an isotropic material, some metallic samples are anisotropic—i.e., they respond differently when loaded in different directions. Single crystals tend to be quite strongly anisotropic, but polycrystals may also be at least slightly anisotropic if they are textured (have a nonrandom distribution of grain orientations). Texture is usually associated with the way in which the sample was manufactured. Anisotropy can also arise from other sources, such as the presence of aligned hard particles or fibers. It may well affect both plastic and elastic deformation. For both of these, this introduces a need for a larger number of parameter values and in general the level of complexity can rise considerably. On the contrary, it may be so weak that its effects can be neglected. In fact, the PIP procedure allows the presence and approximate nature of any (plastic) anisotropy to be picked up quickly and easily. The issue of anisotropy is covered in more detail in Section 3.2.

Once the elastic constants have been specified, indentation has been conducted, and the profile has been measured, the iterative FEM algorithm is implemented and the solution (set of three plasticity parameter values) is obtained. With the PLX Plastometer (Figure 7), this complete operation takes just a couple of minutes or so. Associated with this solution is a value of S_{red} (Equation (4)), representing the goodness-of-fit between measured and modeled profiles. More pictorially, the two can be superimposed, as shown in **Figure 8** for as-received and annealed copper (with two different penetration depths in each case).

The modeled profiles in Figure 8 are those obtained using the Voce parameter sets in **Table 1**, which also shows corresponding L–H sets and the S_{red} values. It should be noted that, although the parameter sets are different for the two depths (in all cases), they actually correspond to similar stress–strain curves. This can be seen from the (true) stress–strain curves shown in **Figure 9**. However, it should also be noted that they are not the same. This could be attributed to small differences in the material response at different depths, although a more likely explanation in this case is that it is due to the response being sampled over different ranges of plastic strain— see Figure 5.

This is tied in with the fact that it is not actually expected that the true stress–strain curve can be captured perfectly over the complete range of strain by any analytical relationship. It can also be seen in Figure 9 that the Voce and L–H formulations are leading to slightly different curve shapes, particularly for the annealed material at high plastic strains. It can be argued that the two sets of L–H curves are rather inconsistent in this case because it is unlikely that the flow curve for annealed material would in fact rise above that for the as-received material at high strains: they are essentially the same material (with and without prior work hardening) and so might be expected to be capped at similar stress levels at high strains—as indeed happens with the Voce law. As discussed in Section 1.2, there is now a general perception that the Voce law is likely to be slightly more effective than L–H in capturing the behavior over the complete strain

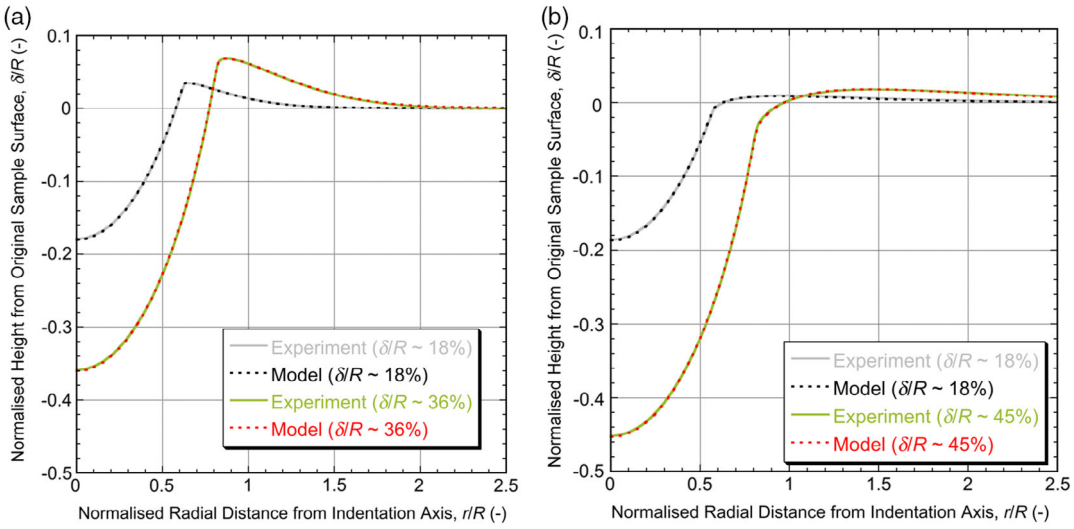

Figure 8. Measured and (best fit) modeled indent profiles, obtained using Voce parameter sets, with two penetration depths, for a) AR–Cu and b) Ann–Cu. Reproduced with permission.[29] Copyright 2019, Elsevier.

Table 1. Best fit (L–H and Voce) plasticity parameter values for the two coppers, for two penetration depths.

Parameter	AR–Cu		Ann–Cu	
	$\delta/R \approx 20\%$	$\delta/R \approx 0\%$	$\delta/R \approx 0\%$	$\delta/R \approx 40\%$
L–H yield stress, σ_Y [MPa]	258.1	237.7	43.5	41.7
L–H WH coefficient, K [MPa]	200.6	221.2	569.9	543.6
L–H WH exponent, n [–]	0.674	0.536	0.635	0.633
L–H misfit parameter, S_{red} [–]	$10^{-4.16}$	$10^{-4.56}$	$10^{-4.05}$	$10^{-4.80}$
Voce yield stress, σ_Y [MPa]	266.8	249.7	48.9	46.2
Voce saturation stress, σ_s [MPa]	421.4	468.0	354.9	358.7
Voce characteristic strain, ε_0 [%]	53.2	41.1	16.6	16.5
Voce misfit parameter, S_{red} [–]	$10^{-4.17}$	$10^{-4.52}$	$10^{-4.31}$	$10^{-4.69}$

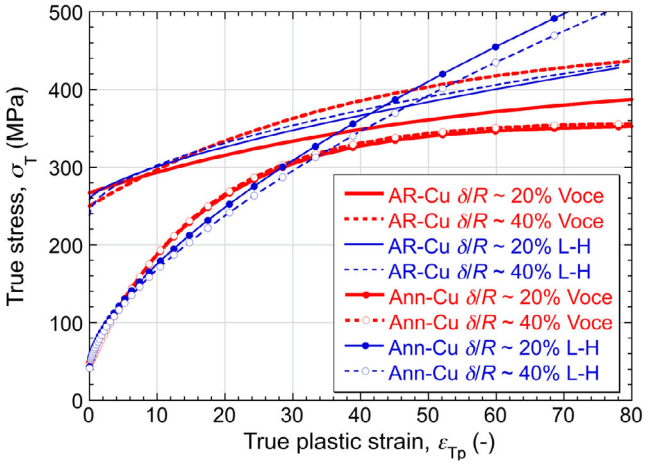

Figure 9. The true stress–true strain curves represented by the plasticity parameter sets in Table 1.

range of interest, particularly for metals that exhibit pronounced work hardening.

Mention should also be made of the fact that some metals exhibit (true) stress–strain curves that cannot be well captured using either L—H or Voce formulations, or indeed using any well-established functional form. One example is provided by single crystals, which often show an initial "easy glide" regime with approximately constant stress level (in which only one slip system is operative), followed by work hardening as further slip systems are activated and dislocations start to interact with each other and become less mobile. Another is materials that show an initial load drop as dislocations escape from their solute "atmospheres," followed by a pulse of straining at constant load while they move, driven by released strain energy. However, in practice single crystals constitute a rather special class of materials, with their behavior exhibiting a strong dependence on their orientation with respect to the loading axis, while few metals of industrial significance actually exhibit a pronounced load drop. The

need for constitutive laws to describe the plasticity of metals in which the work hardening rate does not decrease monotonically has therefore been limited.

However, there is a class of metals, of increasing industrial significance, in which the work hardening rate does increase sharply with increasing strain, followed by a decrease. The resulting stress–strain curves are often described as sigmoidal. Among the most prominent of such materials is the class usually referred to as metastable austenitic stainless steels (MASS). The rise in work hardening rate comes in this case from the mechanical stimulation of martensite formation, which often starts to take place only after appreciable plastic deformation of the austenite has occurred (with little work hardening). The exact behavior depends on several factors, including the temperature (because the thermodynamic driving force for martensite formation is greater at lower temperature) and the stacking fault energy, with a low value facilitating the cross-slip that can be involved in the phase transformation.[48–51] The composition is important, including the carbon and nitrogen levels. High levels of these tend to retard martensite formation, although very low levels of carbon lead to a reduction in the hardness of the martensite.

There is in fact a relatively small "window" of composition and conditions for which the effect is pronounced. However, there is strong interest in promoting such behavior because it is very favorable in terms of energy absorption and crashworthiness. In fact, usage of steel sheet of this type is increasing markedly in the automotive industry. To apply the PIP procedure to metals of this type, an analytical expression is needed that can capture such stress–strain curves, preferably with a relatively small number of adjustable parameters. This may prove to be viable within a fairly short period.

2.2. Cross-Checking with Other Experimental Outcomes

There are several ways in which the reliability of a particular (true) stress–strain relationship (obtained via PIP testing) can be validated. An obvious one is via comparison with outcomes from uniaxial testing. This can be done in tension and/or compression, although the latter does tend to have an element of uncertainty associated with the effects of interfacial friction[38–42,52–54] and comparisons with tensile test outcomes are more common. The easiest comparison to make is between the nominal stress–strain curves, with that for the tensile test being obtained directly and that from PIP testing via FEM simulation of the tensile test (using the inferred true stress–strain relationship and the sample dimensions used in the test—although these only affect the postnecking part of the plot).

A comparison of this type is shown in **Figure 10** for the as-received and annealed copper, with the modeled predictions based on using the Voce parameter values in Table 1 (for 20% penetration—i.e., on two of the true stress–strain curves in Figure 9). It can be seen that the agreement is good and extends to the postnecking parts of the curve. For the AR–Cu material, which necks more than the Ann–Cu one, the shape of the tensile sample at the point of rupture and the plastic strain field within it are shown in **Figure 11**. It can be seen that this field becomes highly inhomogeneous as the neck develops.

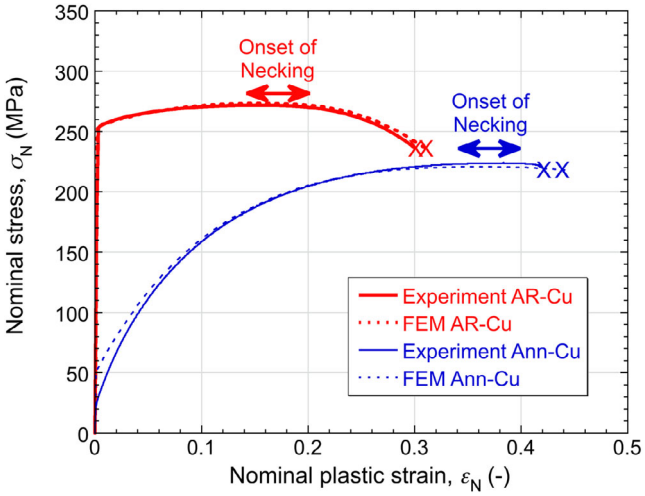

Figure 10. Nominal stress–strain curves from tensile testing of two copper alloys, comparing directly measured plots with those obtained via FEM simulation of the test, using PIP-derived true stress–strain curves.

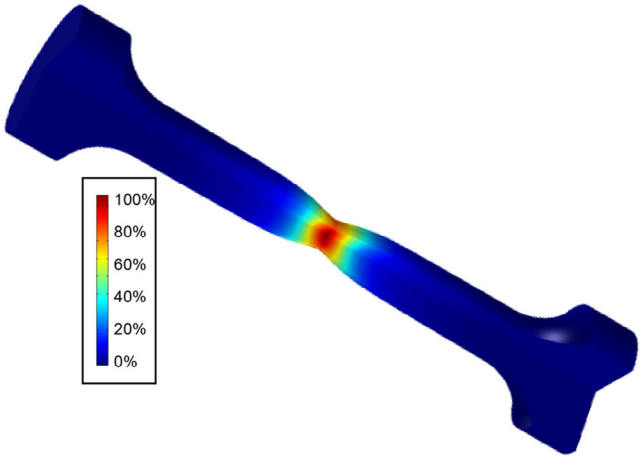

Figure 11. Outcome from FEM simulation of the tensile test on the AR–Cu, showing the equivalent plastic strain field at the point of rupture.

For both materials, the point of rupture has been predicted on the basis of the true strain within the neck reaching a critical value. This is in fact a commonly used criterion for fracture. The value concerned certainly cannot be defined universally, but in general it is found that something in the approximate range of 50–100% is appropriate, at least for reasonably tough metals. The relationship between the nominal stress–strain curve and the true one is shown in **Figure 12**, for the as-received copper. Up to the onset of necking (while the stress and strain fields are uniform), they are related by the well-known analytical relations (obtained by conserving volume). The true stress is given by

$$\sigma_T = \frac{F}{A} = \frac{FL}{A_0 L_0} = \frac{F}{A_0}(1+\varepsilon_N) = \sigma_N(1+\varepsilon_N) \tag{5}$$

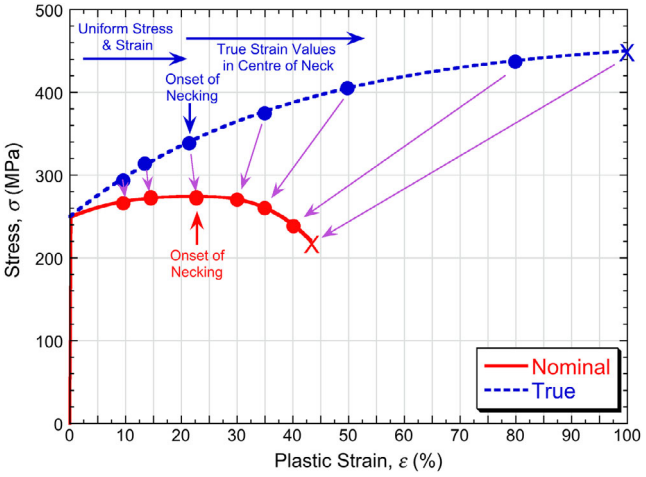

Figure 12. True and nominal stress–strain curves for a tensile test on the AR–Cu sample referred to in Table 1.

where σ_N is the nominal stress and ε_N is the nominal strain. Similarly, the true strain can be written

$$\varepsilon_T = \int_{L_0}^{L} \left(\frac{dL}{L}\right) = \ln\left(\frac{L}{L_0}\right) = \ln(1+\varepsilon_N) \quad (6)$$

Once necking has started, however, there is no analytical relationship and the stress and strain fields must be obtained by numerical modeling. The plot in Figure 12 shows how the true stress and strain in the (center of the) neck relate to different points along the nominal stress–strain curve, for the AR–Cu. A point that should be noted here is that the so-called "ductility" (nominal strain at fracture) is not of any inherent significance. It varies substantially with the dimensions of the tensile sample. (The stress–strain data in Figure 10 and Figure 12 refer to a cylindrical sample, with the ratio of reduced section length to diameter being 6 and that of the gauge length to the diameter being 2.5.) The critical true strain at fracture, on the contrary, is expected to be a material property, although FEM modeling is required to obtain it from a tensile test outcome. Once it has been established for a particular material, it can be used for other purposes, including prediction of fracture for components in service (via FEM simulation).

An alternative comparison between the PIP-based FEM outcome and the experimental tensile test relates to the shape of the neck at the point of fracture. Obtaining this experimentally during the test is quite easily done, via a video recording with suitable resolution. This is free from any of the possible uncertainties associated with measurement of strain. Video stills[18] from the end of tensile tests (immediately before fracture) on two materials are shown in **Figure 13**. It can be seen that the agreement regarding the neck shape is good in both cases. Such comparisons also allow estimates of the peak value of the true strain (in the neck) at the point of fracture, as described earlier. For both of these cases, it has reached about 100%.

2.3. Sensitivity Issues

Central to much of the PIP approach is the issue of sensitivity. If a reliable stress–strain curve is to be obtained, then the measured outcome (indent profile) must be suitably sensitive to it. Of course, this relates to the accuracy of the experimental measurements, as well as to the nature of the change in an inferred curve when the profile changes in some way. Tied in with this rather complex issue is the possibility that some extraneous effect, such as residual stresses or material anisotropy, could be influencing the outcome and potentially affecting the sensitivity. Sample-specific issues such as these are treated in Section 3, but an attempt is made here to identify some inherent sensitivity characteristics. This is done by conducting a series of FEM simulations of the indentation operation, with no experimental data involved.

For example, **Figure 14**a shows a set of stress–strain curves, in the form of one default case and several variants of it. These were all obtained using the sets of Voce parameter values shown, plus two elastic constants (a Young's modulus, E, and a Poisson ratio, ν). Figure 14b is the corresponding set of indent profiles, all for an applied load of 2 kN. Several features are apparent here. These include the effect of the value of E. It can be seen that, if it is changed substantially—reduced by a factor of about 3 in this case—then there is a noticeable (although relatively small) effect on the profile. This arises because the value of E influences the way that the plastic strain field develops, with a higher value leading to reduced penetration: there is then a "spring-back" effect during unloading, when the elastic strains are partly relaxed (with little or no further plasticity). The net effect is that a higher value of E leads to a lower final penetration depth, although the effect is a fairly weak one (and depends slightly on the value of ν). The practical consequence is that E can be chosen from a small set of specified values, with no need for any concern about the exact figure. Furthermore, given that the effect of the ν value is small, and that the range within which it falls is much narrower than for E, taking its value to be 0.33 in all cases is acceptable.

The second effect illustrated in the figure is that making the material a little harder across the complete strain range leads, as expected, to a reduced penetration depth and also to slightly lower pileup. The third effect relates to a case in which the material is initially softer, but work hardens more strongly. These two changes more or less cancel out in terms of the penetration depth, but there is a strong effect in the form of a reduced pileup height, which tends to be much lower for materials with a high work hardening rate. (This is due to the fact that the strains tend to be high in the pileup region—see Figure 4—so that the height is very sensitive to the work hardening characteristics.)

As a further illustration of sensitivity effects, **Figure 15** shows a compilation of outcomes from running an FEM model with various combinations of the three Voce parameter values (each combination corresponding to a pair of yield stress (YS) and ultimate tensile stress (UTS) values). It should be borne in mind that the load would normally be adjusted so that the penetration depth (Figure 15a) was in a selected range—typically around 200 μm for this indenter size ($R = 1$ mm). In that sense, only the lower part of Figure 15a is directly relevant. Nevertheless, such plots do provide information of interest, bearing in mind the expected accuracy of

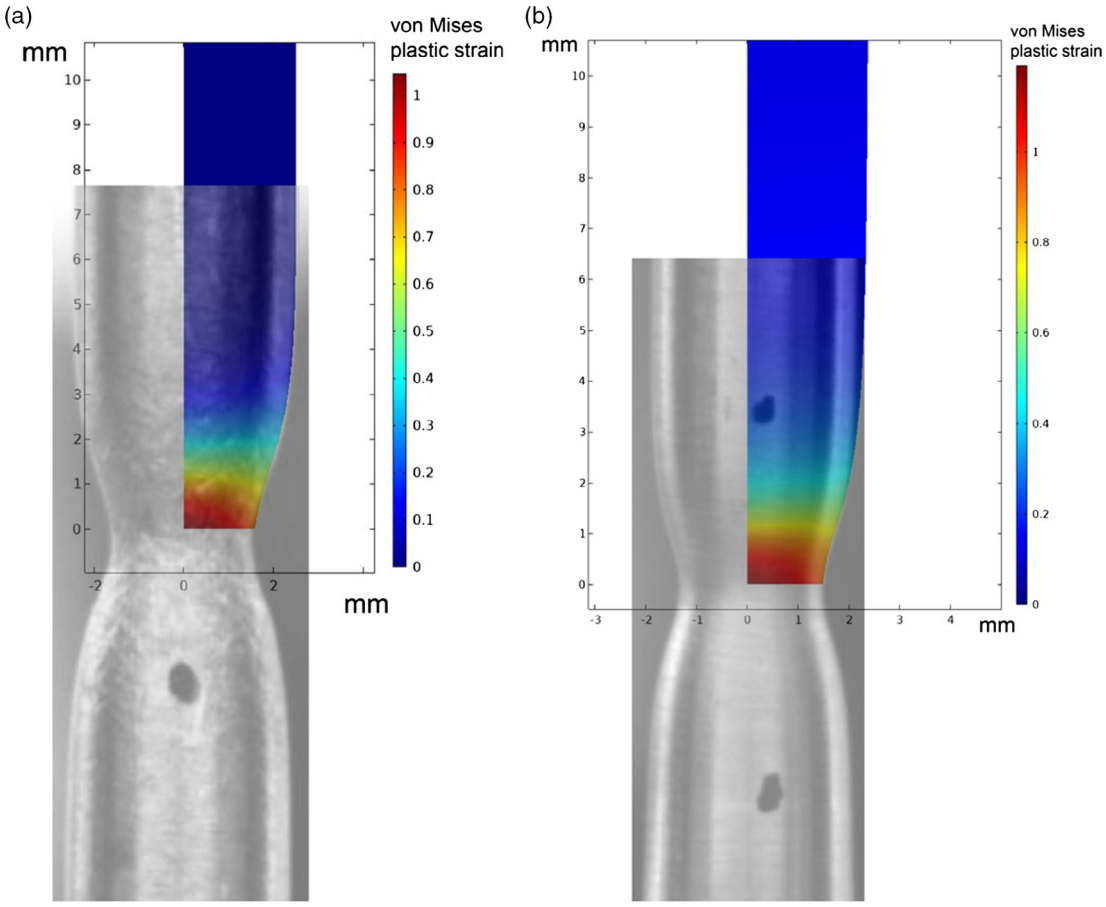

Figure 13. Superimposition of video stills[18] of the neck region immediately before fracture during tensile tests and corresponding FEM model outcomes at that point, showing predicted neck shapes and plastic strain fields, for: a) Cu and b) Al.

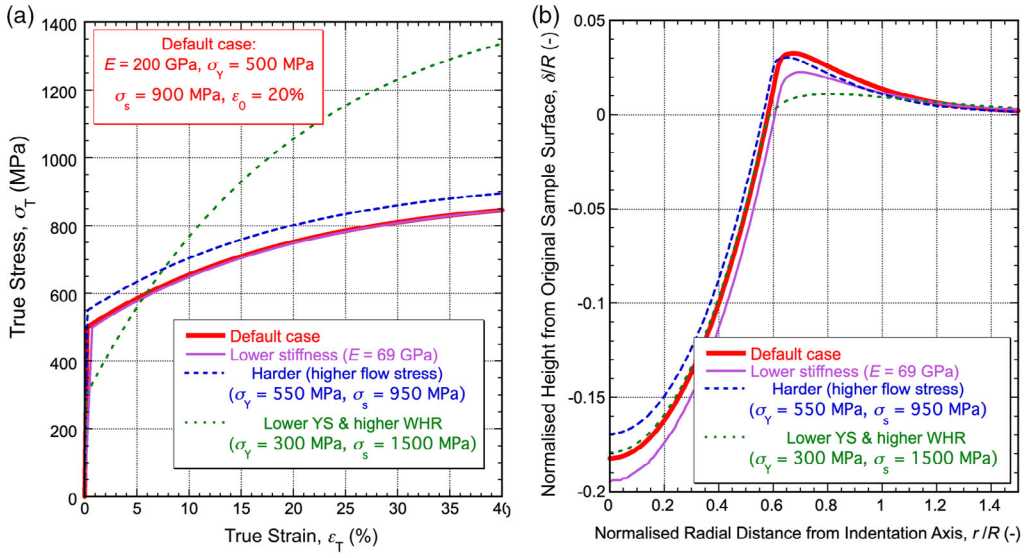

Figure 14. a) A set of (true) stress–strain curves and b) the corresponding set of indent profiles, for an applied load of 2 kN and an indenter radius of 1 mm.

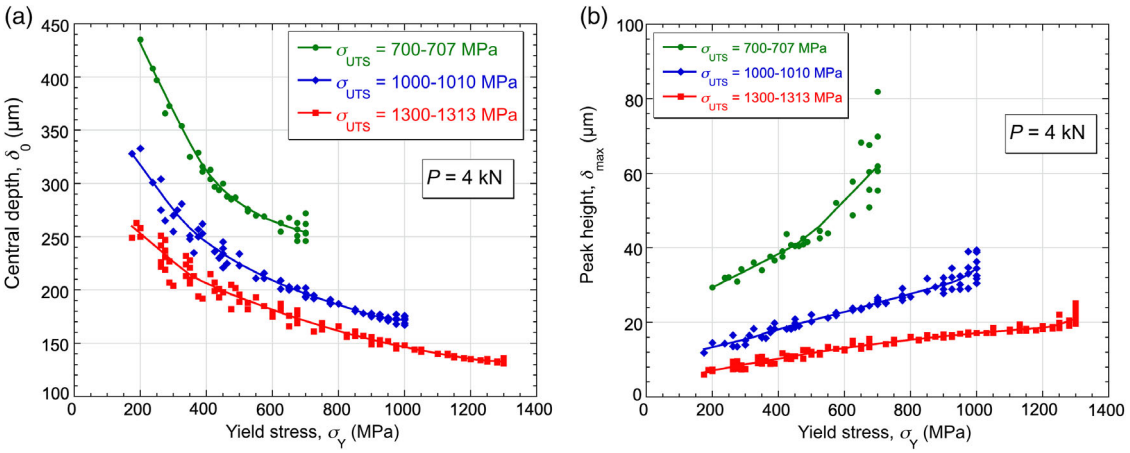

Figure 15. Sensitivity plots from FEM outcomes, showing a) indent depth and b) pileup height, with an applied load of 4 kN and an indenter radius of 1 mm, for various combinations of yield stress and UTS.

the measurement. Typically, the resolution of a profile measurement is about 1—2 μm. With a YS of 400 MPa and a UTS of 1300 MPa (red line in the plot), representing a material with strong work hardening, a change in the measured depth of 2 μm (1%) would alter the inferred YS value by about 10 MPa (2.5%). On the contrary, with a YS of 800 MPa and a UTS of 1000 MPa (blue line), corresponding to little work hardening, this change would only alter the inferred YS by about 5 MPa (0.5%).

It may also be noted again that inferred work hardening characteristics are quite sensitive to the pileup height (Figure 15b). These heights increase rather sharply as the YS approaches the UTS—i.e., as work hardening becomes very weak. For the YS = 400 MPa, UTS = 1300 MPa case, an increase in pileup height from 10 to 12 μm would lead to the inferred YS being raised to about 500 MPa. On the contrary, for the YS = 800 MPa, UTS = 1000 MPa case, an increase in pileup height from 28 to 30 μm would only raise the inferred YS by about 50 MPa. In general, however, the pileup height tends to have a strong effect on the outcome, partly because it reflects the response of material that has been subjected to relatively high plastic strains. It needs to be measured with care. Of course, the actual inferred Voce set will depend on both central depth and peak height (and indeed on the complete shape of the profile), so considering these inferred changes in isolation is potentially misleading. Nevertheless, there are certain points concerning sensitivities that can be deduced from such plots.

3. Microstructural ("Sample-Specific") Effects

3.1. Mechanisms of Plastic Deformation

The mechanisms involved in plastic deformation of metals are, of course, well established. They predominantly involve glide of dislocations, although there are many subtleties involved in how this occurs in different situations[45] and other processes, such as diffusion and deformation twinning, may also be relevant. Of course, most measurements concerned with plasticity are conducted with little or no direct reference to microstructure, with the sample treated as if it were simply a homogeneous continuum. Treatment is also commonly based on the assumption that its properties are isotropic, although this may not be strictly true —see Section 3.2.

However, an awareness of the mechanisms involved can be helpful in interpreting the outcome of tests. For example, if the strain rates are very high, then dislocation glide may be unable to take place quickly enough to accommodate the imposed strain, leading to changes in mechanism, a requirement for higher stresses, and a dependence of the stress–strain relationship on strain rate (which is neglected in the range taken to be a "quasistatic" regime). Information about the strain rate sensitivity of the plasticity can be obtained[55] from PIP-based procedures. Similarly, if loads are maintained for extended periods, then "creep" may occur, during which deformation largely proceeds via diffusional processes, and can occur progressively even under a constant applied load. Temperature is thus relevant and even relatively small changes can strongly affect the significance of diffusion. Again, it is possible to use a variant of indentation plastometry[56] to obtain information about creep characteristics.

Furthermore, even if conditions are such that classical quasistatic plasticity takes place, it is helpful to appreciate that certain microstructural changes, such as the introduction of fine precipitate arrays or refinement of the grain size, can influence the mobility of dislocations and therefore the stress–strain relationship. This also applies to the impairment of mobility as plastic straining proceeds, due to a rise in dislocation density and the formation of entanglements, leading to work hardening.

In addition, microstructural effects can lead to other types of complication (for both uniaxial and PIP testing). These include the presence of anisotropy, arising from some kind of bias in the orientation of individual grains or from other microstructural features (such as alignment of fibers or "stringers" of particles). This is covered in Section 3.2. Another effect, which is particularly relevant to PIP testing, is the possible presence of inhomogeneity in the microstructure. A common example is provided by samples in which the region near a free surface is harder than the interior—perhaps due to a case-hardening (e.g., carburizing)

treatment or to the fact that those regions have undergone higher plastic strains during production (e.g., by rolling of plate). As PIP testing interrogates a relatively shallow layer (≈few hundred microns deep), it is well suited to picking up such variations. On the contrary, if a comparison is made between such a PIP-inferred stress–strain curve and one obtained by tensile testing of a much deeper cross section, then an apparent discrepancy may arise. These issues are covered in Section 3.3. Finally, while this is not necessarily apparent on examining a microstructure, it should be recognized that a particular region (being subjected to PIP testing) may contain residual stresses. This also is a "sample-specific" issue, which could affect the inferred stress–strain curve. This is explored in Section 3.4.

3.2. Sample Anisotropy

As noted in Section 2.1, while most metallic samples are at least approximately isotropic—i.e., they respond identically when loaded in different directions—some are noticeably anisotropic. This effect is a very strong one for single crystals. For both uniaxial loading and PIP testing, the response often varies substantially with the orientation of the crystal relative to the loading axis. The strength of the elastic anisotropy of common metals, as indicated, for example, by the range of values exhibited by the Young's modulus as the loading direction is changed, varies from a factor of about 3 (for copper) to a negligible difference (for tungsten). Data concerning the elastic anisotropy of single crystals are available in Nye.[57] These confirm that even stronger anisotropy than that of Cu is exhibited by some metals with crystal structures of lower symmetry than the cubic system, such as Zn (hexagonal) and Sn (tetragonal).

Single crystals also exhibit marked plastic anisotropy. Furthermore, while some metallic single crystals exhibit little elastic anisotropy—for example, that of Al is quite low—virtually all of them are plastically anisotropic. During uniaxial loading, slip normally starts on a single system (plane and direction), with others being triggered as straining proceeds. The details depend on crystal orientation and operative slip systems. Similarly, indentation also causes slip to occur only on a small number of systems and the resultant indent is not radially symmetric. An example is shown in **Figure 16**, for a Ni-based superalloy that was indented along the [001] axis. In face-centered cubic crystals such as this, slip occurs on (111)-type planes in [110]-type directions. These slip planes intersect (001)-type surfaces (such as the free surface in Figure 16a) along [110]-type directions, which are inclined at 45° to the vertical and horizontal directions ([100] and [010]). The lines running in these directions are persistent slip bands, where sets of dislocations have reached the free surface, creating steps. It can be seen that essentially just two slip systems have been operating in this crystal. Furthermore, not only is the shape of the rim noncircular, but the indent profiles differ sharply in different directions, with much more pileup in the ±45° directions (Figure 16b). Little has happened near the rim in the horizontal and vertical directions, with the material being out of contact with the indenting sphere in those regions. It is not really appropriate to attempt to obtain a stress–strain curve via the PIP methodology for single crystals. It might be possible to extract some information from indentation experiments, although account would need to be taken of the orientation of the crystal and of the operative slip systems in it.

Of course, most metallic samples are polycrystalline. Nevertheless, it is quite common for there to be at least some texture—i.e., for the distribution of orientations of individual grains to be nonrandom. This can arise during solidification, when cubic crystals usually exhibit a preference for growing in <100> directions, and also during processes such as rolling and drawing, when grain rotations occur in a systematic way. It is possible to relate processing conditions to texture, and texture to mechanical properties, although this is a complex area.[58–60] In general, textured materials tend to exhibit anisotropy in a similar way to the corresponding single crystal, although, of course, the strength of the effect tends to be much weaker. However, it may still be significant. An example is provided here of another superalloy sample, which in this case was made by additive

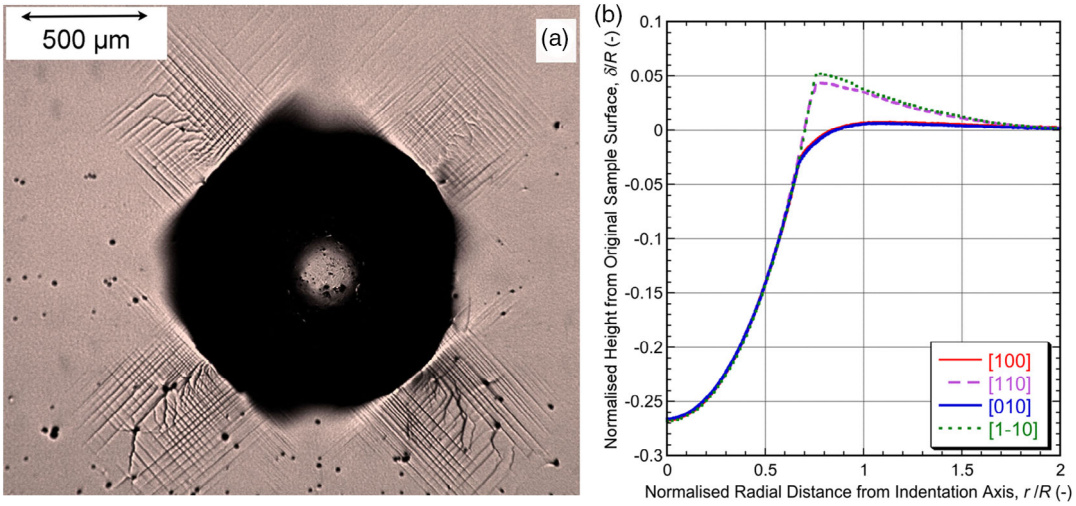

Figure 16. a) Optical micrograph of the free surface around an indent in a CMSX-4 sample, obtained using a 1 mm radius sphere, and b) corresponding indent profiles at four different scan angles. (The crystal was oriented with [001] along the indentation axis and with [100] and [010] in vertical and horizontal directions in the micrograph.).

manufacturing.[61] This leads to a fairly strong texture, with a tendency for <100> directions to be aligned parallel to the growth axis. Outcomes of indenting this sample, along and normal to this axis, are shown in **Figure 17**.

It can be seen in Figure 17 that, while the material is transversely isotropic about the growth axis, so that an indent made with its axis parallel to that direction is radially symmetric, this is not the case when the indention axis is along a transverse direction. The indent diameter is then larger in the growth direction than in the other transverse direction and, even more strikingly, the pileup height is greater for the profile along the growth direction (Figure 17, bottom right). This is a reflection of the fact that the material is "softer" along the growth direction than transversely. This is confirmed by the two tensile test stress–strain curves shown in Figure 17. The standard PIP procedure can only be applied to a single (radially symmetric) profile and the outcome of applying it to the profile in the bottom left of Figure 17 is compared with these tensile test curves in the figure. It can be seen that it is closer to the transverse curve. This is consistent because an indentation test is effectively sampling the response in all directions, so a bias toward that in the transverse directions is expected in this case. The present PIP state-of-the-art does not allow stress–strain curves for different directions to be obtained, although the full 3D modeling (with more plasticity parameter values) that this would require is currently the subject of a development program.

While anisotropy does complicate the interpretation of (nonradially symmetric) indent profiles in terms of obtaining stress–strain curves, it should be noted that they constitute a quick and easy way of detecting its presence. By routinely scanning profiles in multiple directions, not only can any anisotropy be noted, but a clear indication of its nature and approximate strength can be obtained. Moreover, the sensitivity of this detection is much better than could be obtained by conducting uniaxial tests in different directions, even for cases in which this might be a practical proposition. Furthermore, it is often possible to make indents in different sections of a component, allowing a full 3D picture to be built up of the nature of the anisotropy. In particular, this can be done on transverse sections of relatively thick plates or sheets, giving a quick indication of whether properties in the through-thickness direction are significantly different from those in the in-plane directions. In practice, this is not uncommon, although it is rarely studied.

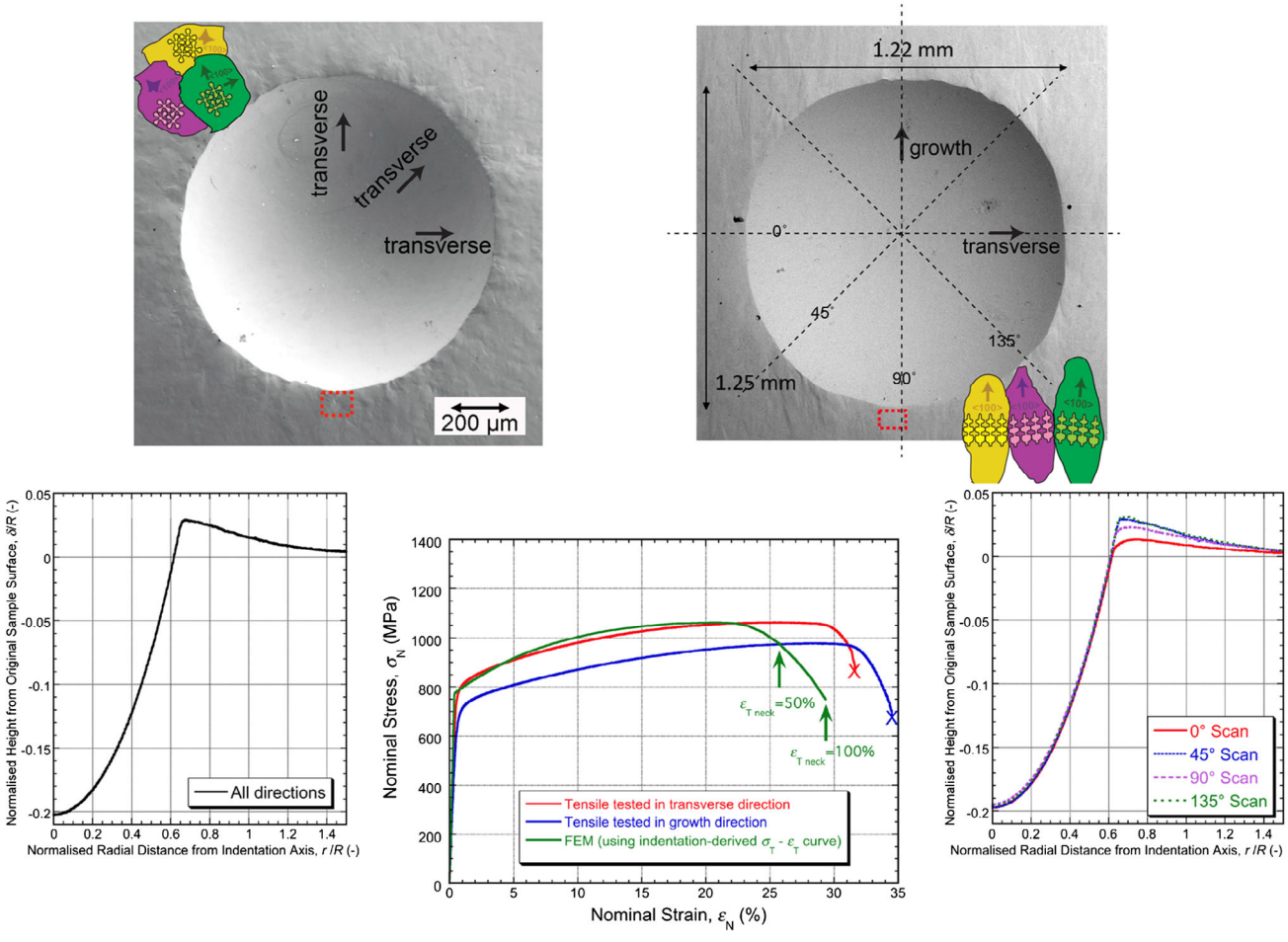

Figure 17. Information from indentation plastometry of a Ni-based superalloy sample made by additive manufacturing showing (top) scanning electron microscopy (SEM) micrographs of indents made with the axis either parallel (left) or normal (right) to the growth direction and (bottom) corresponding indent profiles and a comparison between stress–strain curves obtained by PIP and by tensile testing. Reproduced with permission.[61] Copyright 2021, Elsevier.

3.3. Sample Inhomogeneity and Subsurface Properties

While the indentation involved in PIP is conducted on a relatively large scale (Section 1.3), there is still scope for mapping of properties over a surface and investigating local regions. The "lateral resolution" is of the order of the ball diameter (which is typically 1 or 2 mm), although the spacing between adjacent indents should be at least slightly more than this (to ensure that none of the volume being interrogated has been deformed by the previous indent). This should be adequate for many purposes, such as study of weldment and heat-affected zone (HAZ) properties, but it does mean that, particularly for relatively thin sheet, it may be difficult to map properties across a transverse section. On the contrary, the "depth resolution" of PIP is better than this because the region that is substantially deformed during the test typically extends to a depth of only about 20–30% of the indenter radius (from the original free surface, for a δ/R ratio of 20%—see Figure 4). This corresponds to about 100—150 and 200–300 μm for 1 and 2 mm diameter balls. This means that, if the outer free surface of a component is indented, then it is actually a relatively thin near-surface region that is being tested. This could exhibit properties that are different from those in the interior.

Tied in with this is the possibility that the surface preparation techniques—cutting, grinding, polishing, and so on—may affect the properties of the near-surface region. There has, of course, been work in this area. For example, Safari et al.[62] observed that the near-surface layer of obvious grain deformation in Ti—6—4 alloy increased in thickness from about 10 μm for a cutting speed of 100 m min^{-1} to about 50 μm at 300 m min^{-1}. They also made subsurface microhardness measurements. Unfortunately, this has limitations because the indents themselves are typically a few tens of μm in lateral extent. (Smaller indents could be made using a nanoindenter, but these would present problems of grain-to-grain variation, effects of surface oxide films, and so on.) They did observe greater hardening with higher feed rates and higher cutting speeds, with indications of the hardness being raised down to depths of several hundred microns, but there was a lot of scatter. It may also be noted that the cutting was done dry and with relatively deep cuts. Also, Ti has a relatively low thermal conductivity, with consequences in terms of subsurface heating. In addition, as noted by Telrandhe et al. ,[63] the prior microstructure of the Ti—6—4 alloy can have a strong effect on the machining characteristics

Slightly more systematic subsurface microhardness profiles in Ti—6—4 were obtained by Ramakrishnan et al. ,[64] although this was after machining with an abrasive water jet. For a material with a "bulk" Vickers hardness of about 310 kgf mm^{-2}, they observed an increase down to about 200 μm below the surface, being about 340 kgf mm^{-2} at 50 μm depth (and recognizing that accurate hardness measurement at shallower depths than this can be difficult). Of course, the cutting mechanisms are rather different with this machining technique. As with conventional cutting, however, it is possible to obtain a good surface finish in combination with considerable subsurface "damage."

The most common metallographic preparation technique is that of mechanical abrasion (grinding), often followed by polishing. The difference between the two is not well defined, but grinding is commonly conducted with relatively coarse abrasive particles attached to a surface (usually paper-like), while during polishing they are finer and suspended in a liquid. There has been extensive work on the resultant subsurface damage for ceramics and semiconductors, but less concerning metals (despite its extensive application to them). Of course, the mechanisms of material removal tend to be different in the two cases, with much less danger of cracking in metals. It is nevertheless clear that subsurface damage can arise in metals during this type of treatment, usually due to induced plastic deformation causing changes in dislocation density and possibly in grain shape, texture, grain boundary structures, and so on. The changes are in general expected to be more significant, and to extend deeper, in soft metals and in those with high initial work hardening rates—in fact, these two characteristics often go together. Nevertheless, substantial effects can be observed in much harder metals, such as the superalloy studied by Österle and Lee ,[65] who found that efficient cooling during grinding reduced the degree of damage. Furthermore, Zeng et al.[66] found that mechanically polished CoCrMo alloys exhibited hardened surface layers (containing martensitic phases and deformation twins), although they tended to be only ≈1 μm or so in thickness. Much deeper deformed layers have been found in other cases, with a sensitivity to alloy type being found for various cast irons[67] and steels.[68,69]

There are, of course, other surface preparation techniques, including electrodischarge machining, electrolytic polishing, sand blasting, and so on, all of which have different characteristics in terms of subsurface damage. In general, however, it can probably be concluded that, if the preparation is reasonably careful, the subsurface region affected by it can be kept quite thin (<≈10 μm). Such layers can be neglected for PIP testing purposes (although not for nanoindentation!).

Of course, this is a separate issue from that of a component exhibiting inhomogeneities, such as a harder near-surface region, due to higher prior plastic straining during a rolling process, or indeed due to a case hardening treatment. An example of the kind of effect that can arise is provided here, relating to 16 mm-thick steel plate. Conventional PIP testing of free surfaces generated (nominal) stress–strain curves reflecting a somewhat harder response than that obtained by conventional in-plane tensile testing (of most of the section). Tensile tests were then conducted using 2 mm-thick samples taken from close to the two free surfaces and from the center. PIP testing was also conducted at similar depths. The outcome is presented in **Figure 18**, which shows stress–strain curves from the two techniques. It can be seen that similar trends are observed from both, with the near-surface regions being noticeably harder than the center.

Data from a number of such tests are also shown in **Figure 19**, which presents data in the form of YS and UTS values taken from the two types of test, plotted against the through-thickness location. The main message to take from such comparisons is that care needs to be taken concerning the possibility of components and samples exhibiting spatial variations in microstructure and properties. Of course, PIP testing does offer the potential for exploring and understanding such variations in much more detail than is possible using conventional techniques. (This applies equally to anisotropy and to inhomogeneities—and of course both may be present in individual samples.)

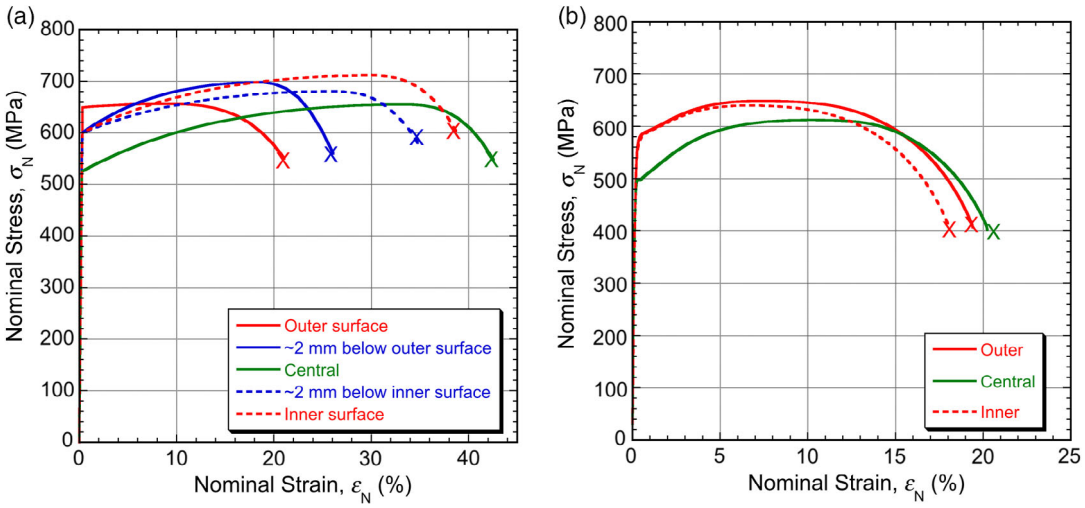

Figure 18. Tensile (nominal) stress–strain curves for different through-thickness locations in a steel plate, obtained a) via PIP processing and b) via conventional tensile testing.

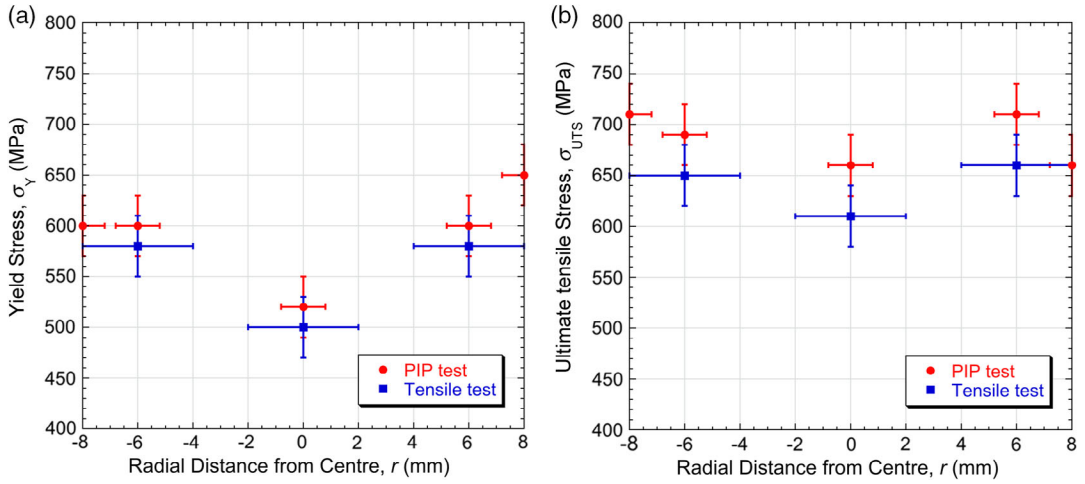

Figure 19. Plots against through-thickness location (distance from the center) of a) yield stress and b) UTS, obtained by PIP processing or conventional tensile testing.

3.4. Residual Stresses

It has long been clear that residual stress in the near-surface region can affect the way that indentation occurs. An attraction of PIP procedures is that it is a simple matter to incorporate any given set of residual stresses into the FEM model. The fact that indentation is usually conducted to a relatively shallow depth, and in a region of limited lateral extent, means that residual stress state within the region being tested can be taken to be uniform. (This still leaves scope for mapping of residual stress over a surface, by making indents in different locations.) Furthermore, there is no stress normal to a free surface, so the residual stress state is fully characterized by the two (in-plane) principal stresses. In some cases, these two principal stresses will be the same, so that only a single value needs to be obtained. In other cases, three parameters are needed, two being values of the two principal stresses and the third specifying their orientation.

Several other general points can be identified. Indenting initially creates mainly a compressive stress in the through-thickness direction. According to both von Mises and Tresca criteria, yielding takes place when the difference between this stress and the initial in-plane residual stress reaches the (uniaxial) yield stress. Yielding will thus be promoted by a tensile residual stress and inhibited by a compressive one, with the effect expected to be stronger with tensile stresses. As penetration takes place, the stress field becomes more complex, but the trends are still expected to be in the same direction and some asymmetry is likely to remain in terms of the magnitude, as well as the sense, of the effects of compressive or tensile residual stress. It may also be noted that, as a broad generalization, residual stresses in near-surface regions are commonly compressive. This will tend to make the material appear slightly harder than it actually is when subjected to PIP testing. Moreover, while residual stresses in a sample could affect the outcome

of uniaxial testing (tensile or compressive), any such effects are likely to be very small (because all residual stresses in a free-standing sample must force-balance to zero when integrated over the whole volume).

Further complexity arises with unequal biaxial residual stresses, but a key question is whether the sensitivities are such that inferring residual stresses in this way is likely to be practicable. In this context, it may be noted that profilometry at least has potential for establishing the directionality of any anisotropy in the residual stress state, whereas this is not possible if only load–displacement data are used.

There have in fact been many studies aimed at using indentation data to measure residual stresses in a sample. Unfortunately, a large proportion of them have been focused on assessment of hardness numbers and/or have involved conducting the indentation on a very fine (nano) scale. The ill-defined nature of hardness and the various problems associated with nanoindentation (Section 1.3) mean that little weight can be given to this body of work. However, some recent studies have involved use of large spherical indenters and inverse FEM analysis, offering more promise for clarification of the prospects of measurement of residual stresses in this way (and also, rather importantly, for assessing whether the unknown presence of such stresses could lead to significant errors in the PIP methodology for plasticity characterization).

The most relevant work in this area[70–75] has involved experimental indentation of samples in which controlled "artificial" residual stresses were created by the external application of (equal or unequal biaxial) forces. Most of these studies have been based on the load–displacement plot being used as the experimental outcome, although in some of them[71,73] the residual indents were also viewed optically, with it being noted that their shapes became (approximately) elliptical when the residual stresses were unequal. However, the recent study of Burley et al.[75] was based on PIP methodology, giving scope for quantification, as well as detection, of anisotropic residual stresses (in a similar way to that described in Section 3.2 for anisotropy in the plasticity of the sample).

The following two figures are taken from this work, which in fact concerns only equal biaxial residual stresses. **Figure 20** shows a set of experimental residual indent profiles, obtained with the superimposed (von Mises) residual stresses shown in the legend, all of which are well below the yield stress (≈200 MPa). The two sets of curves correspond to two different values of the indentation load, P. The value of P is usually chosen so as to give a suitable penetration ratio (δ/R). In general, this should be of the order of 20%, generating plastic strains in approximately the required range (Figure 5). This was the case here for $P = 1.0$ kN (Figure 20b), but penetrations were lower for $P = 0.25$ kN (Figure 20a). General trends are as expected in both cases, with tensile stresses promoting penetration and compressive stresses inhibiting it. However, the relative differences between the curves for different residual stress levels are greater with shallower penetration. As the indenter goes deeper, the stress field becomes more complex, with more lateral movement of material. The stress field from the indentation load thus departs from the initial situation, in which it is largely a compressive stress in the loading direction. With deeper penetration, the effect of the residual stress level is less significant, particularly for compressive ones.

The procedure for obtaining the applied (equal biaxial) residual stress from the profile is to run the model with a series of residual stress values, comparing measured and predicted profiles for each experimental case and evaluating the misfit parameter, S_{red}, each time. The inferred level is the one with the smallest value of this parameter. As there is only one parameter to infer, convergence—i.e., systematic movement in parameter space—can be obtained from a simple linear plot. For each applied residual stress level, this procedure was conducted for several different indentation loads.

Ideally, there would be sharp minima in S_{red} for each curve, at the actual applied residual stress level for the curve concerned. In practice, there were minima, and they did exhibit a broad trend of increasing inferred stress level as the actual level is raised. However, many of the minima were not very sharp (indicating relatively poor sensitivity) and agreement between inferred and

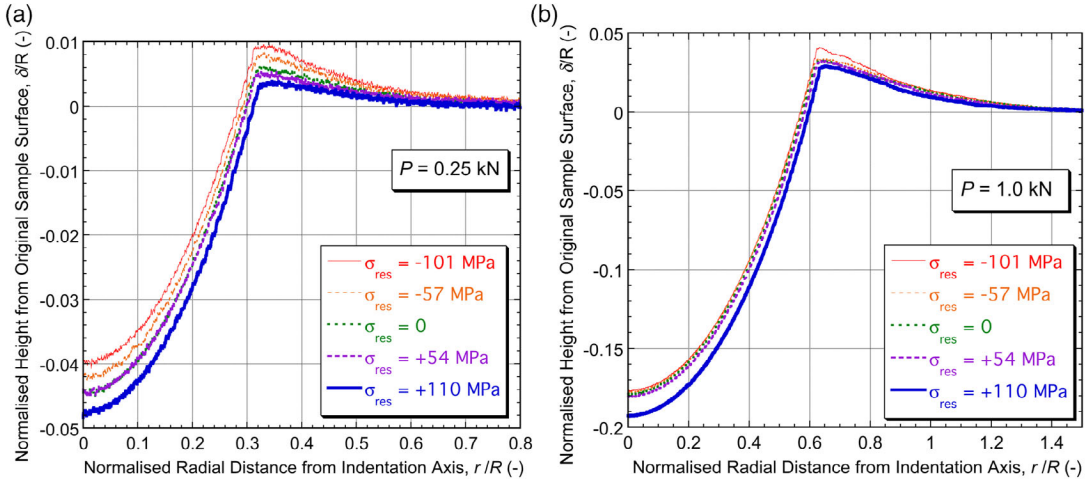

Figure 20. Experimental indent profiles for five levels of applied (equal biaxial) residual stress, for indentation of copper samples with a 1 mm radius ball and loads of a) 0.25 kN and b) 1.0 kN.

actual levels was not close, particularly for higher applied loads. This correlates with the residual indent profiles shown in Figure 20, where the influence of the residual stress level is seen to be lower with higher applied loads.

Figure 21 gives a practical perspective on outcomes, showing the values of σ_Y and σ_{UTS} that would be obtained via this procedure. (The Voce sets were converted to these properties to assist interpretation.) With zero residual stress (green curves), these values were the "correct" ones ($\sigma_Y \approx 200$ MPa and $\sigma_{UTS} \approx 260$ MPa). The red (tensile) and blue (compressive) data show certain features. For example, the sensitivity (i.e., the displacement to "incorrect" inferred values by the residual stress) is greater for shallow penetration (low δ/R). With $P = 0.25$ kN ($\delta/R \approx 4$–5%), σ_Y has been moved to below 100 MPa by the tensile residual stress, and up to about 250 MPa by the compressive one, whereas the corresponding values for $P = 1$ kN ($\delta/R \approx 18$—20%) are about 150 and 220 MPa. For the σ_{UTS} values, the displacements are lower and in the opposite directions, although the rather flat "curves" for the $P = 0.25$ kN case are indicative of poor resolvability. The "resolution" is better for the $P = 1$ kN case and the "displacements" (downward by a few MPa for the compressive stress and upward by about 20 MPa for the tensile one) are small.

As these results relate to a high residual stress (relative to the yield stress, although not in absolute terms), these outcomes suggest that errors in inferred stress–strain curves due to residual stresses are likely to be fairly small, particularly with deep penetration. The sensitivity is greater for tensile stresses than for compressive ones. In practice, as mentioned earlier, near-surface residual stresses are often compressive—for example, due to rolling, shot peening, and so on—which is advantageous in terms of being able to neglect their effects, but a disadvantage if they are to be measured.

Being able to extract plasticity characteristics via PIP testing, even if (unknown) residual stresses are present, is clearly important. As an example, welds often have very different plasticity

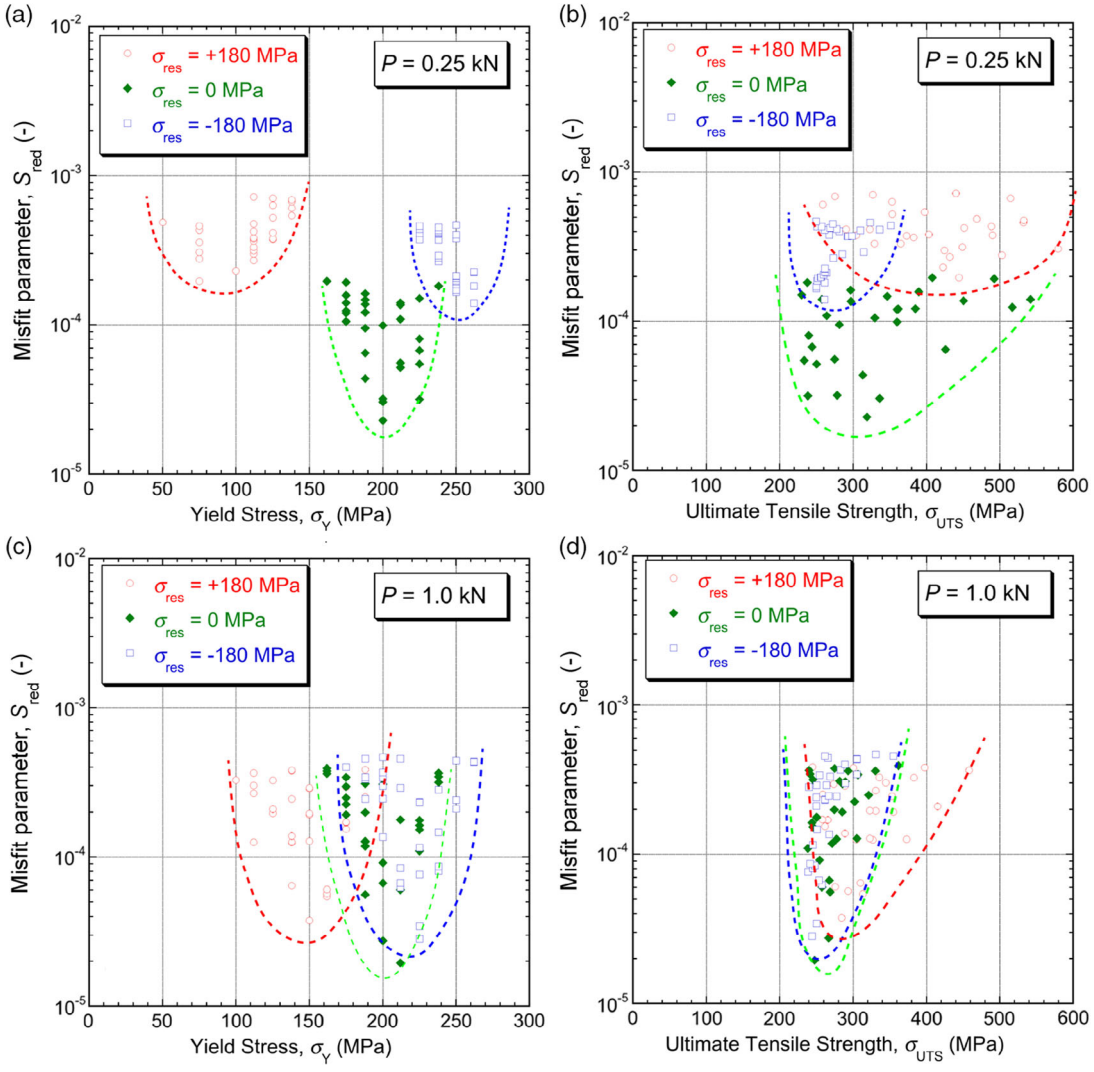

Figure 21. Scatter plots of S_{red} values for various Voce sets, characterizing the misfit between indent profiles obtained with and without residual stress. This is done both for shallow penetration ($P = 0.25$ kN), plotting against corresponding values of a) σ_Y and b) σ_{UTS}, and for deep penetration ($P = 1$ kN), plotting c) σ_Y and d) σ_{UTS}.

characteristics from those of the parent metal and also commonly contain residual stresses. Dimensions of most welds are such that conventional (tensile) testing is difficult, so use of PIP to obtain stress–strain curves for them, and even for different parts of a weld and for nearby HAZ regions, is very attractive. This is an area requiring comprehensive study.

There is also the issue of whether the PIP approach can be used to measure residual stress levels. The results presented here highlight that the sensitivities are not favorable for this. However, using shallow penetration has the potential to improve them. It should be noted that this relates to relative penetration (of a spherical indenter) and not to absolute depths. A figure of, say, 1—2% (i.e., about 10—20 μm depth with a 1 mm radius sphere) could give fairly good sensitivity in terms of inferring the residual stress level (provided the plasticity characteristics are accurately known). However, shallow penetration brings its own problems, in terms of the necessary surface finish, the high resolution required for the depth measurement, and so on. Also, indents could be located in single grains, leading to scatter as a result of the dependence on grain orientation. In general, obtaining accurate residual stress levels in this way presents severe challenges, particularly if they are relatively low or are unequal biaxial. Despite the optimism expressed in a number of publications, the prospects do not look very promising.

4. Summary and Future Perspective

The following points can be noted regarding PIP technology. 1) Integrated facilities are now commercially available, so an understanding of their capabilities and features is starting to become highly significant. It seems likely that usage of the technology will expand considerably over the next few years. 2) If "bulk" properties are required, then a "representative" (many-grained) volume must be tested. Moreover, plastic strains should be created in the approximate range of up to several tens of %. These requirements lead, in turn, to a need for an indenter radius of around 0.5—1 mm, a penetration depth of a few hundred μm, and a load capability in the kN range. The so-called "nanoindenters" are completely unsuitable. 3) Using the profile of the residual indent as the target outcome, rather than the load-displacement plot, offers many advantages. It is easier from an experimental point of view, its sensitivity to the stress–strain curve is greater, and it allows detection and exploration of sample anisotropy. 4) Using inverse FEM modeling to infer the stress–strain relationship is a far superior approach to that of attempting to convert a load–displacement plot to such a curve via some kind of functional relationship (which inevitably requires very crude approximations and assumptions). Inverse FEM fully captures the evolving stress and strain fields during the test and also has potential for the incorporation of residual stresses, material anisotropy, and so on. With suitably accurate experimental data and model formulation, convergence can readily be obtained on a "unique" solution. While "manual" iteration is slow and cumbersome, commercial packages now incorporate a capability for fast, automated operation, with no need for any user input or familiarity with FEM simulation.

Acknowledgement

This article is part of the *Advanced Engineering Materials* Hall of Fame article series, which highlights the work of top scientists in the field of engineering materials. Relevant support for TWC has been received from EPSRC (grant no. EP/I038691/1) and from the Leverhulme Trust, in the form of an International Network (grant no. IN-2016-004) and an Emeritus Fellowship (grant no. EM/2019-038/4).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

hall of fame article, indentation, inverse finite element method, plastometry, profilometry, stress–strain curves

Received: April 13, 2021
Revised: July 14, 2021
Published online:

[1] B. X. Xu, X. Chen, *J. Mater. Res.* **2010**, *25*, 2297.
[2] R. O. Oviasuyi, R. J. Klassen, *J. Nucl. Mater.* **2013**, *432*, 28.
[3] C. Yu, Y. H. Feng, R. Yang, G. J. Peng, Z. K. Lu, T. H. Zhang, *J. Mater. Res.* **2014**, *29*, 1095.
[4] Z. Song, K. Komvopoulos, *Mech. Mater.* **2014**, *76*, 93.
[5] S. Pathak, S. R. Kalidindi, *Mater. Sci. Eng. R-Reports* **2015**, *91*, 1.
[6] J. S. Weaver, S. R. Kalidindi, *Mater. Design* **2016**, *111*, 463.
[7] C. Chang, M. A. Garrido, J. Ruiz-Hervias, Z. Zhang, L. L. Zhang, *Adv. Mater. Sci. Eng.* **2018**, 8316384.
[8] F. Pohl, *Exp. Tech.* **2018**, *42*, 343.
[9] T. R. Zhang, S. Wang, W. Q. Wang, *Results in Phys.* **2018**, *8*, 716.
[10] H. Chen, L. X. Cai, C. Bao, *Strain* **2020**, *56*, 14.
[11] A. R. H. Midawi, N. Huda, C. H. M. Simha, A. P. Gerlich, *Metallogr., Microstruct., Anal.* **2020**, *9*, 884.
[12] K. Goto, I. Watanabe, T. Ohmura, *Mater. Des.* **2020**, *194*, 8.
[13] F. Y. Huang, Y. W. Liu, J. C. Kuo, *Applied Nanoscience* **2021**, 15.
[14] S. Mohan, N. Millan-Espitia, M. Yao, N. V. Steenberge, S. R. Kalidindi, *Exp. Mech.* **2021**, *61*, 641.
[15] A. H. Mahmoudi, S. H. Nourbakhsh, R. Amali, *J. Test. Eval.* **2012**, *40*, 211.
[16] K. Jeong, H. Lee, O. M. Kwon, J. Jung, D. Kwon, H. N. Han, *Mater. Design* **2020**, *196*, 109104.
[17] L. Lu, M. Dao, P. Kumar, U. Ramamurty, G. E. Karniadakis, S. Suresh, in *Proceedings of the National Academy of Sciences of the United States of America* **2020**, Vol. *117*, pp. 7052–62.
[18] J. E. Campbell, H. Zhang, M. Burley, M. Gee, A. T. Fry, J. Dean, T. W. Clyne, *Adv. Eng. Mats.* **2021**, *23*, 2001496.
[19] C. Heinrich, A. M. Waas, A. S. Wineman, *Solids and Structures* **2009**, *46*, 364.
[20] J. Dean, J. M. Wheeler, T. W. Clyne, *Acta Mater.* **2010**, *58*, 3613.
[21] D. K. Patel, S. R. Kalidindi, *Acta Mater.* **2016**, *112*, 295.
[22] J. Dean, T. W. Clyne, *Mech. Mater.* **2017**, *105*, 112.
[23] J. E. Campbell, R. P. Thompson, J. Dean, T. W. Clyne, *Mech. Mater.* **2018**, *124*, 118.
[24] L. Meng, P. Breitkopf, B. Raghavan, G. Mauvoisin, O. Bartier, X. Hernot, *Inter. J. Mater. Forming* **2019**, *12*, 587.
[25] H. Xue, J. X. He, W. N. Jia, J. L. Zhang, S. Wang, S. Zhang, H. L. Yang, Z. Wang, *Structures* **2020**, *28*, 2752.
[26] J. Lee, C. Lee, B. Kim, *Mater. Des.* **2009**, *30*, 3395.

[27] W. Z. Yao, C. E. Krill, B. Albinski, H. C. Schneider, J. H. You, *J. Mater. Sci.* **2014**, *49*, 3705.
[28] M. Z. Wang, J. J. Wu, Y. Hui, Z. K. Zhang, X. P. Zhan, R. C. Guo, *Mater. Sci. Eng. a-Struct. Mater. Properties Microstructure and Processing* **2017**, *679*, 143.
[29] J. E. Campbell, R. P. Thompson, J. Dean, T. W. Clyne, *Acta Mater.* **2019**, *168*, 87.
[30] J. H. Hollomon, *Trans. Am. Inst. Min., Metall. Eng.* **1945**, *162*, 268.
[31] E. Voce, *J Inst. Met.* **1948**, *74*, 537.
[32] T. Belytschko, R. Gracie, G. Ventura, *Modell. Simul. Mater. Sci. Eng.* **2009**, *17*, 043001.
[33] F. Roters, P. Eisenlohr, L. Hantcherli, D. D. Tjahjanto, T. R. Bieler, D. Raabe, *Acta Mater.* **2010**, *58*, 1152.
[34] H. Ghaednia, S. A. Pope, R. L. Jackson, D. B. Marghitu, *Tribol. Int.* **2016**, *93*, 78.
[35] A. E. Giannakopoulos, S. Suresh, *Scr. Mater.* **1999**, *40*, 1191.
[36] B. Taljat, G. M. Pharr, *Int. J. Solids Struct.* **2004**, *41*, 3891.
[37] V. Karthik, P. Visweswaran, A. Bhushan, D. N. Pawaskar, K. V. Kasiviswanathan, T. Jayakumar, B. Raj, *Int. J. Mech. Sci.* **2012**, *54*, 74.
[38] M. Fardi, R. Abraham, P. D. Hodgson, S. Khoddam, *Adv. Eng. Mater.* **2017**, *19*, 1700328.
[39] M. Bol, R. Kruse, A. E. Ehret, *J. Mech. Behav. Biomed. Mater.* **2013**, *27*, 204.
[40] G. Torrente, *Mater. Res.-Ibero-American J. Mater.* **2018**, *21*, e20170905.
[41] X. G. Fan, Y. D. Dong, H. Yang, P. F. Gao, M. Zhan, *J. Mater. Process. Technol.* **2017**, *243*, 282.
[42] D. Duran, C. Karadogan, *Strojniski Vestnik-Journal of Mechanical Engineering* **2016**, *62*, 243.
[43] P. Hausild, A. Materna, J. Nohava, *Mater. Des.* **2012**, *37*, 373.
[44] G. Pintaude, A. R. Hoechele, *Mater. Res.* **2014**, *17*, 56.
[45] T. W. Clyne, J. E. Campbell, *Testing Of The Plastic Deformation Of Metals*, Cambridge University Press, Cambridge, U.K **2021**.
[46] E. C. Teague, F. E. Scire, S. M. Baker, S. W. Jensen, *Wear* **1982**, *83*, 1.
[47] A. G. Marrugo, F. Gao, S. Zhang, *J. Opt. Soc. of America a-Optics Image Sci. Vision* **2020**, *37*, B60.
[48] Y. C. Zhang, M. C. Li, H. Y. Bi, J. Q. Gu, D. X. Chen, E. Chang, W. Zhang, *Mater. Sci. Eng. a-Struct. Mater. Prop. Microstruct. Proc.* **2018**, *724*, 411.
[49] M. H. Zhang, H. Y. Chen, Y. K. Wang, S. J. Wang, R. G. Li, S. L. Li, Y. D. Wang, *J. Mater. Sci. Technol.* **2019**, *35*, 1779.
[50] T. W. Mukarati, R. J. Mostert, C. W. Siyasiya, *Mater. Sci. Eng. a-Struct. Mater. Prop. Microstruct. Proc.* **2020**, *792*, 139741.
[51] T. Masumura, T. Tsuchiyama, *ISIJ Int.* **2021**, *61*, 617.
[52] Z. H. Yao, D. Q. Mei, H. Shen, Z. C. Chen, *Tribology Letters* **2013**, *51*, 525.
[53] J. Zhou, P. He, J. F. Yu, L. J. Lee, L. G. Shen, A. Y. Yi, *J. Vacuum Sci. Technol. B* **2015**, *33*, 031213.
[54] X. Wang, H. Li, K. Chandrashekhara, S. A. Rummel, S. Lekakh, D. C. Van Aken, R. J. O'Malley, *J. Mater. Process. Technol.* **2017**, *243*, 465.
[55] M. Burley, J. E. Campbell, J. Dean, T. W. Clyne, *Int. J. Impact Eng.* **2018**, *112*, 180.
[56] M. Burley, J. E. Campbell, J. Dean, T. W. Clyne, *Int. J. Mech. Sci.* **2020**, *176*, 105577.
[57] J. F. Nye, *Physical Properties Of Crystals - Their Representation By Tensors And Matrices*, Clarendon, Oxford **1985**.
[58] S. R. Kalidindi, *Int. J. Plasticity* **2001**, *17*, 837.
[59] P. R. Dawson, S. R. MacEwen, P. D. Wu, *Int. Mater. Rev.* **2003**, *48*, 86.
[60] H. R. Wenk, P. Van Houtte, *Reports on Progress in Physics* **2004**, *67*, 1367.
[61] Y. T. Tang, J. E. Campbell, M. Burley, J. Dean, R. C. Reed, T. W. Clyne, *Materialia* **2021**, *15*, 101017.
[62] H. Safari, S. Sharif, S. Izman, H. Jafari, *Int. J. Adv. Manuf. Technol.* **2015**, *78*, 651.
[63] S. V. Telrandhe, A. K. Saxena, S. Mishra, *J. Mech. Sci. Technol.* **2017**, *31*, 2177.
[64] S. Ramakrishnan, V. Senthilkumar, D. L. Singaravelu, *Mater. Res. Express* **2019**, *6*, 116583.
[65] W. Osterle, P. X. Li, *Mater. Sci. Eng. a-Struct. Mater. Properties Microstructure and Processing* **1997**, *238*, 357.
[66] P. Zeng, A. Rana, R. Thompson, W. M. Rainforth, *Wear* **2015**, *332*, 650.
[67] R. B. Da Silva, M. L. S. Lima, M. F. Pereira, B. S. Abrao, L. R. R. Da Silva, E. C. Bianchi, A. R. Machado, *Int. J. Adv. Manuf. Technol.* **2018**, *99*, 1839.
[68] B. Weiss, A. Lefebvre, O. Sinot, M. Marquer, A. Tidu, *Surface & Coatings Technol.* **2015**, *272*, 165.
[69] N. Srinivasan, B. S. Kumar, V. Kain, N. Birbilis, S. S. Joshi, P. V. Sivaprasad, G. Chai, A. Durgaprasad, S. Bhattacharya, I. Samajdar, *Metall. Mater. Trans. A* **2018**, *49A*, 2281.
[70] T. H. Pham, S. E. Kim, *Mater. Sci. Eng. a-Struct. Mater. Properties Microstructure and Processing* **2017**, *688*, 352.
[71] G. J. Peng, Z. K. Lu, Y. Ma, Y. H. Feng, Y. Huan, T. H. Zhang, *J. Mater. Res.* **2018**, *33*, 884.
[72] T. H. Zhang, W. Q. Cheng, G. J. Peng, Y. Ma, W. F. Jiang, J. J. Hu, H. Chen, *MRS Communications* **2019**, *9*, 360.
[73] G. J. Peng, F. G. Xu, J. F. Chen, H. D. Wang, J. J. Hu, T. H. Zhang, *Metals* **2020**, *10*, 440.
[74] R. Moharrami, M. Sanayei, *Measurement* **2020**, *158*, 107718.
[75] M. Burley, J. E. Campbell, R. Reiff-Musgrove, J. Dean, T. W. Clyne, *Adv. Eng. Mats.* **2021**, *23*, 2001478.

Bill Clyne is the chief scientific officer at Plastometrex Ltd., a start-up in Cambridge, and Emeritus Professor in Cambridge University, having held a Chair in Mechanics of Materials there for 25 years. His interests cover the thermomechanics of composites, metals and coatings, focusing on novel mechanical testing methods. He has authored 400 papers, 3 textbooks, and 10 educational software packages. He was President of the Federation of European Materials Societies (http://www.fems.org/) and is now an honorary member. He is the director of DoITPoMS (www.doitpoms.ac.uk/), which hosts an Educational Materials Science website. He is an FREng, a Helmholtz International Fellow, and a Leverhulme Emeritus Fellow.

Jimmy Campbell has a Ph.D. in materials science from Cambridge and is now the CTO (chief technical officer) at Plastometrex, with expertise in mechanical testing of metals with a focus on indentation and finite element modelling.

Max Burley has an undergraduate and master's degree in materials science from the University of Cambridge. He went on to do a Ph.D. in materials science with Professor Bill Clyne on the extraction of mechanical properties from indentation data. Currently, he is the chief hardware engineer at Plastometex Ltd.

James Dean has an undergraduate degree in materials science from Imperial College, London, and a Master Degree in gas turbine engineering from Cranfield University. He has a Ph.D. in materials science from Cambridge, and is the former coordinator of the EPSRC-funded Centre for Doctoral Training in Computational Methods for Materials Science at the Cavendish Laboratory (also in Cambridge). He founded Double Precision Consultancy (a mathematical modeling consultancy) in 2012, which was sold to Element Materials Technology in 2021. He is a cofounder of Plastometrex and its current CEO.