Tutorial

A practical guide for analysis of nanoindentation data

Michelle L. Oyen\textsuperscript{a,*}, Robert F. Cook\textsuperscript{b}

\textsuperscript{a} Cambridge University Engineering Department, Cambridge, CB2 1PZ, UK
\textsuperscript{b} Ceramics Division, Materials Science and Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA

\textbf{A R T I C L E I N F O}

Article history:
Received 3 June 2008
Received in revised form
25 September 2008
Accepted 3 October 2008
Published online 15 October 2008

\textbf{A B S T R A C T}

Mechanical properties of biological materials are increasingly explored via nanoindentation testing. This paper reviews the modes of deformation found during indentation: elastic, plastic, viscous and fracture. A scheme is provided for ascertaining which deformation modes are active during a particular indentation test based on the load–displacement trace. Two behavior maps for indentation are presented, one in the viscous–elastic–plastic space, concerning homogeneous deformation, and one in the plastic versus brittle space, concerning the transition to fracture behavior when the threshold for cracking is exceeded. Best-practice methods for characterizing materials are presented based on which deformation modes are active; the discussion includes both nanoindentation experimental test options and appropriate methods for analyzing the resulting data.

\textcopyright 2008 Elsevier Ltd. All rights reserved.

1. Introduction

There is a wide range of biological materials with fundamentally structural or mechanical functions, including wood (Low et al., 2006), nacre (Bruet et al., 2005) and bone (Oyen, 2008a). The study of the mechanical behavior of natural materials – either in a medical or a basic science context – is expanding with the application of modern materials characterization techniques to the study of biological systems. Such techniques include local contact methods for determining mechanical properties, including nanoindentation and atomic force microscopy. In addition to trying to gain fundamental physical understanding, there are direct applications for these methods in biomechanical characterization in fields such as medicine, which use mechanical assays for diagnosis of disease. The recent interest in studying mechanics of natural materials is combined with a growing interest in biomimetic materials synthesis for the creation of robust, and in many cases mechanically-functional, materials based on the “design principles” active in nature (Oyen, 2008a).

Three factors motivate the use of nanoindentation – depth-sensing indentation testing at very small length scales, realistically ranging from nanometer to micrometer length-scales – for the study of the mechanical properties of natural materials. First, in such tests the load and displacement of a small probe, the indenter tip, are continuously monitored as the probe is loaded onto the surface of interest, rendering the method ideal for probing local gradients and heterogeneities in natural materials and for examining their hierarchical and multiscale organization. Second, a major factor in the recent rapid growth in nanoindentation is that no extensive sample preparation is required prior to mechanical testing, in sharp contrast to techniques such as tensile testing for which “dog-bone” samples must be cut or machined. Third, most nanoindentation instruments provide experimental control that allows for purposeful exploration of a variety of different...
deformation modes by changing experimental time scale, indenter tip geometry, and loading conditions.

With nanoindentation, one can measure either structural characteristics, such as indenting a cell membrane (Zhao et al., 2006), or the material properties for a single material, typically in the approximation of a bulk half-space (or composite geometry such as a thin layer or inhomogeneous composite). In this work we focus on the latter, on the use of nanoindentation for material properties characterization. In this regard, nanoindents are simply universal test machines that use local contact, and have the added advantage of automation and ease of use.

It is at this point that we encounter the primary problem with the wide-scale adoption of nanoindentation for material properties characterization in biological materials. Ease of use does not translate to easy interpretation of resulting “material properties” data if the test and method of data analysis are not appropriate for the system being considered. In particular, the commercialization of instrumented indentation instruments has led to optimization of test methods for elastic–plastic materials, and data analyses to extract the elastic and plastic properties, the (elastic) modulus and hardness. It is unlikely that these material characteristics would be the most relevant for a range of biological materials: in particular biological materials are likely to exhibit significant viscous deformation (or viscoelastic deformation) with their associated time-dependencies—in many cases plastic deformation is likely irrelevant compared with elastic and time-dependent deformation modes. As a consequence, blindly placing a natural biological sample into a commercial nanoindenter with a Berkovich indenter tip, pressing “go” and accepting the values of modulus and hardness generically provided as output by the instrument software is unlikely to provide an accurate assessment of the material properties of a biological material. This paper aims to provide guidance in matching the test to the natural material properties of interest so as to make the data interpretation as unambiguous as possible.

First we will discuss and describe typical indentation “fingerprints” for combinations of the four modes of mechanical deformation found in materials: elastic, plastic, viscous, fracture. These descriptions will be found to be most conveniently discussed in terms of (i) a map displaying the relative degree of elastic and plastic deformation and the time scale of the experiment relative to the timescale required for viscous deformation and (ii) a parameter quantifying the accommodation of the indentation by the relative amounts of plasticity and fracture. These phenomenological considerations will be followed by a discussion of the best experimental and analytical options for running indentation tests and best-practice guidelines for analyzing the results when elastic, elastic–plastic, viscoelastic, poroelastic, viscous–elastic–plastic, and brittle fracture properties are of interest.

2. Indentation deformation

In this section we consider the modes of deformation, and the methods by which one can identify these modes of deformation based on viewing an indentation load–displacement trace. We first consider four modes of mechanical deformation that are commonly observed during indentation testing: elastic, plastic, viscous, and fracture.
2.1. Modes of deformation

Elastic deformation is time- and rate-independent, instantaneously reversible, and there is no threshold (at very small deformations elastic deformation is present). There are adiabatic and isothermal limits to the elastic modulus, although for engineering materials such as metals and ceramics the difference between these two limits is difficult to detect in typical mechanical characterization experiments. The elastic modulus, \( E \), of a linearly elastic material is the proportionality constant between stress \( \sigma \) and strain \( \epsilon \) \( (\epsilon = E\sigma) \); in a nonlinearly elastic material the instantaneous value of the proportionality constant \( E \) is not single-valued but is strain-dependent. In anisotropic materials, the indentation modulus, \( M \), is used frequently in place of \( E \), where \( M \) is a weighted average of the anisotropic elastic constants relative to the indentation direction (Swadener et al., 2001).

Viscous deformation or flow is strongly time-dependent, irreversible and does not have a threshold for onset. Viscous flow changes the apparent value of the elastic modulus as a function of experimental time or frequency scales and as such it is often considered collectively with the elastic deformation—different mechanisms for viscous dissipation (molecular rearrangement, resistance to flow) are combined with elastic deformation to produce a response that is viscoelastic or poroelastic. The time constant for deformation can be defined in an empirical viscoelastic sense, where the time constant is the ratio of a function of viscosities \( (\eta) \) to moduli \( (E) \) such that \( \tau = \frac{f_1(\eta)}{f_2(E)} \) or in the simplest case of a Maxwell model, \( \tau = \eta/E \). Alternately, the time constant can be defined in the context of poroelastic mechanics, where there is a characteristic length-scale, \( l \), and a hydraulic permeability coefficient, \( \kappa \), from Darcy’s law (Wang, 2000). In this case, the time constant is \( \tau = l^2/\kappa \).

While elastic deformation and viscous flow do not have detectable thresholds, plastic and fracture deformation modes are characterized by threshold values. Both can lead to discontinuities in the indentation load–displacement trace, although the plastic discontinuities are observed usually only at extremely small loads and displacements while the fracture discontinuities are seen only at very large loads and displacements. In other words, the threshold for indentation-induced plastic deformation is usually smaller than the threshold for indentation fracture. (However, this may not be the case if the probe is very obtuse or very acute, and is in clear distinction to macro-scale systems – containing pre-existing cracks – in which fracture often occurs on loading prior to the onset of plastic deformation.) Both fracture and plastic deformation modes are irreversible, time- and rate-independent or weakly time- or rate-dependent. The clear difference between plastic and fracture deformation modes is that plastic deformation does not change the apparent value of the elastic modulus, while fracture does. Fracture is associated with the creation of new surfaces, and the macroscopic defect reduces the apparent sample stiffness. Plastic deformation involves extremely small (atomic-level) defects and does not usually reduce the apparent sample stiffness.

In characterizing plastic deformation, it is traditional to consider the yield stress \( \sigma_Y \), the point on the stress–strain curve at which the response first irreversibly deviates from linearity. The typical measure of indentation plasticity is the hardness, more appropriately termed the “contact Hardness” \( \left( H_c \right) \) and referring to the mean supported contact stress—the peak load \( (P_{\text{max}}) \) divided by the contact area under load \( (A_c) \):

\[
H_c = \frac{P_{\text{max}}}{A_c}.
\]

There is often a linear relationship between the hardness and the yield stress in metals (Tabor, 1951), particularly when the indentation is dominantly plastic (i.e., the elastic deformation is small compared with the plastic deformation). There is a more specific measure for characterizing the resistance to plastic deformation directly, \( H \), confusingly at times also termed the “true” hardness. There is a clear relationship between \( H \) and \( H_c \) in both elastic–plastic (Sakai, 1999; Oyen, 2006a) and viscous–elastic–plastic materials (Oyen and Ko, 2007). In characterizing fracture the fundamental parameter is the fracture resistance, \( R \), the free energy per area required to create new fracture surface. The commonly-used fracture metric of the mode-I critical stress intensity factor, \( K_{IC} \), or toughness, \( T \), is related to \( R \) by \( R = K_{IC}^2 / E \) or \( R = T^2 / E \) (Lawn, 1993).

Now that we have defined the four principal deformation modes and the material properties used to characterize these deformation modes, we will move on to next examine how these different deformation modes appear in an indentation load–displacement plot, and how we can compare materials exhibiting different types of deformation.

2.2. Maps and fingerprints

It is reasonably straightforward to identify the deformation modes active during indentation testing simply by examining the shape of the indentation load–displacement \( (P-h) \) trace. Fig. 1 presents a flow-chart for identification of the deformation modes present in an indentation test in the context of a series of simple yes-or-no questions. Representative load–displacement data are shown for the five “end-points” of the flow-chart: elastic-only deformation (spherical indentation of fused silica glass), brittle fracture (spherical indentation of silicon at large loads, (Oliver et al., 2008)), elastic–plastic deformation (Berkovich indentation of fused silica glass), viscous–elastic deformation (spherical indentation of a glassy polymer) and viscous–elastic–plastic deformation (Berkovich indentation of a glassy polymer). We use the term brittle here to refer to systems in which the indentation hysteresis is caused predominantly by fracture, as opposed to that for elastic–plastic systems in which hysteresis is caused predominantly by plastic deformation mechanisms. Brittleness will be considered in greater detail later in this section.

There is a continuum of responses with different relative amounts of viscous, elastic and plastic deformation, and these possibilities can be illustrated compactly on a deformation mechanisms map (Fig. 2). The abscissa of Fig. 2 represents the ratio \( d \sim 1/t_{\text{fr}} \) of a viscous–elastic time constant \( t \) relative to the experimental time frame \( (t_{\text{fr}}) \); note that the time constant contains the elastic modulus in the denominator regardless of whether viscoelastic or poroelastic deformation is considered. Large values of \( d \)
Fig. 1 – A decision tree structure for identifying dominant modes of deformation in instrumented indentation load–displacement data. Characteristic “fingerprints” for each of five response types are shown as output from the decision tree: E, elastic; B, brittle; P, plastic; V, viscous.

lead to small amounts of viscous deformation, and hence $d$ characterizes the indentation viscous flow resistance (the physical meaning of large $d$ is that the contact cycle is rapid – relative to the viscous deformation time scale – so that viscous flow during indentation is truncated). The ordinate of Fig. 2 represents the ratio $y \sim (H/E)^{1/2}$ (Cook and Oyen, 2007). Large values of $y$ suggest small amounts of plastic deformation, and hence $y$ characterizes the indentation plastic yield resistance (the physical meaning of large $y$ is that the yield stress of the material is large – relative to the modulus – so that plastic deformation by indentation is difficult). The map in two dimensions is the full range of viscous–elastic–plastic (VEP) responses with fully viscous, elastic and plastic deformation in three of the corners. The bold contour lines on Fig. 2 indicate regions where the indentation deformation at peak load is 2/3 of the type labeled. Regions inside these contours away from the center indicate indentation behavior dominated by one deformation mode. The fine contour lines indicate regions where the indentation deformation at peak load is 1/2 of the type labeled. Regions close to two fine contour lines away from the center indicate indentation behavior dominated by two deformation modes. The central region enclosed by the three fine contour lines indicates indentation behavior including all three deformation modes.

The indentation load–displacement traces from Fig. 1 for E, EP, VE and VEP responses are placed on the map (Fig. 2) in Fig. 3. The two sets of responses (for glass and polymers) illustrate the effect of increasing indenter acuity: for the sphere the deformation is totally E or VE, while for the sharper probe the plastic deformation threshold is exceeded and the response is EP or VEP. Similarly, the difference between the glass and polymeric materials is apparent in that the timescales for flow (as given by the time constant of the material)
Fig. 2 – A behavior map detailing the viscous–elastic–plastic indentation space: the indentation viscous flow resistance $d$ is proportional to the ratio of material time constant to experimental time frame; the indentation plastic yield resistance $y$ is proportional to the square-root of the hardness to modulus ratio (Cook and Oyen, 2007).

Fig. 3 – Indentation behavior map from Fig. 2 with indentation data “fingerprints” overlaid for (clockwise from upper right-hand corner) elastic (E), elastic–plastic (EP), viscous–elastic–plastic (VEP) and viscous–elastic (VE) material responses. The four representative experimental plots are for two indenter probe geometries with different acuity (spherical and Berkovich) and two materials with different time constants (silicate glass and glassy polymer).

are far more similar to the experimental indentation time frame for polymers, which exhibit viscous flow in both sets of responses.

It is important to recognize that the coordinates used here to characterize indentation behavior include both material properties and indentation contact variables, and thus both can influence the balance between the deformation modes during a contact event. Materials with large hardness/modulus ratios ($y$) or large time constants ($d$) for flow will obviously tend to VE or EP responses as plastic or viscous deformation, respectively, are suppressed. However, indentation contacts with probes of large acuity ($\psi$, where $2\psi$ is the included angle of the equivalent conical indenter tip) or that extend over a large period of time ($t_{\text{R}}$) counter these effects by imposing large amounts of shear (and thus driving plastic deformation) or long contact periods (and thus enabling viscous deformation; the driving force for viscous deformation may also be increased by large acuity probes). Hence large-acuity pyramidal probes (Vickers, Berkovich, or, in the extreme, cube-corner) loaded for relatively short contact times are used to study elastoplastic indentation in ceramics and dielectrics and small acuity spherical probes loaded for relatively long contact times are used to study viscoelastic indentation in polymers.

In the context of this map (Figs. 2 and 3) we have not considered fracture. We next consider the trade-off between plastic and fracture behavior during indentation, examining in a little more detail the contact stress fields developed in an indented material. On initial contact between an indenter and a material, the material deforms elastically and the work performed by the indenter, $w$ (given by the integral $\int Pdh$) results in a change in elastic energy, $\Delta U_E$, associated with elastic deformation of the material. This change in elastic energy is in turn given by an integral throughout the material volume ($V$) of the change in the elastic energy density, $\bar{u}_E$, $\int \bar{u}_EdV$. The change in elastic energy density is inhomogeneously distributed and of course $\bar{u}_E$ is greatest near the indentation contact site. If the indentation load, and hence indentation size, are made large enough, $\bar{u}_E$ reaches a critical value at some point in the contact field beyond which the material can no longer continue to deform by elastic deformation. In many materials, indentation to greater loads and displacements leads to irreversible plastic deformation in the contact field and the distribution of plastic deformation usually occurs to minimize $\Delta U_E$. Such minimization of course requires less work to be performed by the indenter and hence the onset of plasticity leads to a decrease in the slope of the (now irreversible) indentation $P$–$h$ curve. Similar ideas apply to the onset of viscous deformation. Unloading from indentation loads above such thresholds of course leads to hysteretic indentation responses as observed in Fig. 1.

Increasing indenter acuity or indentation to large enough loads leads to greater $u_E$ values and hence plastic deformation occurs in nearly all materials during non-spherical indentation, also observed in Fig. 2. An extreme case of this mode of indentation deformation is shown in Fig. 4a, which shows the Berkovich indentation response for Al. Such a material–indenter combination exhibits almost complete indentation irreversibility; a $P$ indentation fingerprint that would be sited in the lower right of the maps of Figs. 2 and 3. Fig. 4b shows a Vickers indentation in steel characteristic of indentation in a P material. Metals are the typical small-$y$P materials and tend to have distributed plastic deformation zones that extend considerable distances from the indentation contact area (although this may not be obvious
on surface observation as in Fig. 4b). Some materials, however, do not have such easy plastic deformation mechanisms and demonstrate considerable elastic recovery during indentation. An example is shown in the map of Fig. 2 for soda-lime silicate glass; an EP indentation fingerprint. Minerals, ceramics, semiconductors, and glasses are the typical large-$y$ EP materials and distinct from the P materials tend to have compact plastic deformation zones that are restricted to the material beneath the contact area. The localized plastic deformation zone imbedded in a much larger elastically deformed matrix in such EP materials leads to often-negligible decrease in the slope of indentation $P$–$h$ curve from a purely $E$ response.

Material within the localized plastic deformation zones of EP materials that exhibit limited plasticity has an almost invariant large $u_E$ value of $H^2/E$. Material in the elastic matrix just exterior to the zone also has a large $u_E$ value just slightly less than this value, as such material is in a state of incipient plastic deformation. Further away from the contact, $u_E$ decreases significantly (Lawn et al., 1980; Cook and Pharr, 1990). As a consequence of this large energy density inhomogeneities in the material, either naturally occurring or generated by the indentation contact itself, can lead to fracture in the material. Fracture occurs for the same reason that plasticity occurs: the material can no longer sustain an increase in $u_E$. Considerations of the nature of the stress field and material properties determine whether a material will fracture or plastically deform in order accommodate an imposed load or displacement but common experience suggests that materials plastically deform under shear and fracture under tension. For most materials, the stress field under a conical indentation contact is composed of hydrostatic pressure and shear and exterior to the contact of radial compression and circumferential tension (Cook and Pharr, 1990). As a consequence large-$y$ EP materials exhibit plasticity beneath conical contacts and predominantly radial fracture exterior to the contact. Spherical indentations also exhibit hydrostatic pressure and shear beneath the contact, but radial tension, especially on the surface, exterior to the contact (Johnson, 1985). As a consequence, at small indentation loads large-$y$ EP materials also exhibit plasticity beneath spherical contacts and circumferential surface fracture exterior to the contact. At large indentation loads with spherical indenters, the distinction between the conical and spherical stress fields is lost and radial fracture occurs.

Just as with plasticity, indentation fracture usually occurs to minimize $\Delta U_F$ and thus requires less work to be performed by the indenter. The onset of fracture thus leads to a decrease in the slope of the indentation $P$–$h$ curve. In many cases, however, the fracture events are not large enough or oriented appropriately such that no appreciable change in slope of the $P$–$h$ response is observed on fracture (Lawn et al., 1980). This is particularly so under displacement-controlled indentation. However, under load-controlled indentation, especially with acute probes or to very large indentation loads, such that the scale of the fracture events is comparable with the indentation displacements, the effects can be significant. Unloading from indentation loads above the fracture threshold in these cases also leads to hysteretic indentation responses. An extreme case of this mode of indentation is shown in Fig. 4a, which shows a load-controlled, large-load spherical indentation response for Si (Oliver et al., 2008). Such a material-indenter combination exhibits almost load-free displacement excursions associated with discrete fracture events at the indentation. The development of this response is as follows: small radial cracks oriented normal to the surface are initiated at the
contact periphery during indentation loading and extend on continued loading. The small size and orientation of these cracks does not greatly distort the indentation stress field and hence their formation leads to only small changes in the \( P-h \) response (and very careful experiments can detect such changes on acute indentation (Morris et al., 2004)). The orientation of these radial cracks in the indentation loading stress field is not stable, however, and at great enough loads the cracks turn to propagate nearly parallel with the surface, forming shallow lateral cracks (Cook and Pharr, 1990). The orientation of these cracks and their much greater size greatly distorts the indentation stress field and hence their formation leads to large changes in the \( P-h \) response. If the indentation load is great enough the lateral cracks propagate to the surface, removing a chip of material and hence support of the indenter. Such events lead to the dramatic displacement excursions observed in Fig. 4a. (Thin film systems, in which there is a weak film-substrate interface, provide an intrinsic easy path for lateral crack propagation and very often exhibit such behavior (Toivola et al., 2002).)

The tendency for a material to plastically deform or fracture under a given set of indentation conditions (indentor acuity, indentation load or displacement) can be quantified by the brittleness. Brittleness is a measure of component or system performance that takes into account the plastic deformation and fracture properties of the material involved as well as the externally imposed conditions (these latter appear as geometry terms, in common with many measures of performance (Ashby, 1999)). The material component of brittleness is the so-called brittleness index, an intrinsic system length scale. Systems larger than this length scale exhibit fracture, in this case radial cracks of the same orientation as the indenter. Such events lead to the dramatic displacement excursions observed in Fig. 4a. (Thin film systems, in which there is a weak film-substrate interface, provide an intrinsic easy path for lateral crack propagation and very often exhibit such behavior (Toivola et al., 2002).)

A characteristic length scale for plastic deformation at an indentation, the plastic displacement, \( h_p \), is related to the resistance to plastic deformation, \( H \), by (Cook and Oyen, 2007)

\[
P = \frac{h_p}{h^2} = \frac{a_1}{a_2} H
\]

where \( P \) is the indentation load and \( a_1 \) is a dimensionless geometrical constant. A characteristic length scale for fracture at an indentation, the radial crack length, \( c \), as shown in Fig. 4c, is related to the fracture resistance, \( R \), by

\[
P = \frac{c^{3/2}}{H} = \frac{a_3}{a_3} \sqrt{R E}
\]

where \( a_3 \) is another dimensionless constant characterizing both geometry and the \( E/H \) ratio of the material (Lawn et al., 1980). (Eq. (3) is derived from considerations of the decrease in \( u^* \) exterior to the localized plastic deformation in large-\( \gamma \) EP materials and as such both elastic and plastic deformation parameters, \( E \) and \( H \), are involved.) The brittleness index is the length scale at which the scales of plastic deformation and fracture are equal. From the above two equations this dimension is:

\[
h_p^* = c^* = \left( \frac{a_3}{a_1} \right)^2 \frac{R E}{H^2}
\]

Indentations with plastic displacements greater than \( h_p^* \) exhibit fracture, in this case radial cracks of the same length, \( c^* = h_p^* \) (there thus exists a conjugate threshold load, \( P^* \), for fracture). The materials properties parameter, \( H^2/RE \), is a measure of the tendency to brittle indentation behavior as there is a greater likelihood of fracture under given indentation conditions (load, geometry) for materials for which this parameter is greater. The intuitive expectation that materials with small values of \( R \) are more likely to crack is consistent with this parameter. The geometry parameter \( (a_1/a_3)^2 \) is similarly a measure of the tendency to brittle behavior, as there is a greater likelihood of indentation fracture of a given material if this parameter is greater. The intuitive expectation that probes of greater acuity are more likely to lead to indentation fracture is borne out by the strong dependence of \( a_1/a_3 \) on the effective included angle of the probe: for a given indentation contact pressure (determined by \( H \) and \( E \) ), \( a_1/a_3 = \cot(\sqrt[3]{3} \psi \) (Lawn et al., 1980; Cook and Oyen, 2007) such that acute probes can exhibit extremely small threshold fracture loads.

A dimensionless ratio, the brittleness, \( b \), can be formed using the indentation displacement, \( h \), and the above dimension by

\[
b = \frac{h}{h_p} = \frac{h_p^2}{\left( \frac{a_3}{a_3} \right)} \left( \frac{R E}{H^2} \right)
\]

This parameter has much in common with \( d \) used earlier to describe the viscous flow resistance; it is a combination of an experimentally-controllable parameter (the first term in parentheses, setting the indentation size and probe acuity) and a material parameter (the second term in parentheses, setting the threshold indentation dimension for fracture). The parameter \( b \) can thus be used to rank (a one-dimensional map) indentation systems in increasing brittleness, with large values of \( b \) indicating a brittle (B) indentation and small values indicating a sub-threshold plastic (P) indentation. This scheme is illustrated in Fig. 4a for the indentation of Al and Si.

3. Indentation testing and data analysis

Having defined the modes of deformation, and corresponding material properties, we move on to the practical details of nanoindentation testing. This is in two parts: first we examine the experimental choices available for controlling the deformation modes present during an indentation test, and then we move on to examine the analysis of indentation data.
3.1. Experimental choices for indentation testing

There is a number of experimental selections to be made in indentation testing. These have already been mentioned in the discussion above concerning modes of deformation: indenter acuity, time-frames of the indentation test, and the overall type of load profile, including specification of peak load. These will be discussed briefly here prior to considering data analysis. The critical question in this context is “What is the purpose of the test?” Obviously one is undertaking an indentation test to measure something, some aspect of the material properties. The experimental selections are optimized for the properties that are of interest for the material or system of interest—it simply cannot be the case that the elastic modulus and hardness are the properties of interest in every single case, and thus taking that as the default position could be a mistake.

Probe acuity

The most important experimental selection is that of probe geometry. While most nanoindenters are equipped by default with a Berkovich (relatively obtuse) pyramid, there are choices both more obtuse (large sphere) and more acute (cube corner probe) as shown in Fig. 5. As discussed, there are several ways that changing to a more acute probe can affect the type of deformation resulting from an indentation test. With a more acute probe, a transition can take place from E or VE to EP or VEP (Fig. 3). One can also shift the irreversible deformation from P to B (Figs. 4a and 4b). Conversely, reducing the probe acuity, by moving over to spherical indenters, can increase the Poisson’s ratio as well.

Test type

The indentation test is extremely flexible in terms of the displacement-load-time sequence. Most nanoindenters are intrinsically load-controlled or nearly load-controlled, although there are feedback controls to allow for displacement control in some systems. The applied testing sequence is thus a load-time sequence \( P(t) \). An indentation cycle is typically at least a loading and unloading sequence, potentially with a hold (Fig. 6a) or dynamic (frequency-sweep) segment (Fig. 6b) at fixed load. Key experimental controllables are thus the peak indentation load \( (P_{\text{max}}) \), the ramp time \( (t_{\text{r}}) \), and the hold time or other frequency profile at maximum load. There are other types of loading profiles such as the constant indentation strain rate (actually exponential loading (Oyen and Cook, 2003)) and multiple loading and unloading cycles. Many commercial indentation instruments have software similar to that employed in universal test frames, such that almost any load-time profile of interest could be programmed into the instrument. The data analysis must then be matched to the test that was run. This is particularly important for materials with a viscous response in the experimental time frame, as the behavior of a viscoelastic material is loading history dependent.

3.2. Data analysis

The maps and fingerprints then guide the experimental choices and these choices, along with the deformation modes present, then uniquely determine which analysis method is appropriate for extracting material properties from the indentation data. Although the Oliver and Pharr (1992) method for extracting elastic–plastic properties is built into most commercial indentation systems, other analyses are more appropriate for certain combinations of test and material deformation. We now explore the methods of data analysis for the five different deformation combinations considered here (E, EP, VE, VEP, BEP). Note that we restrict this discussion to materials that are bulk half-spaces (i.e., not thin films) and for which the microstructure does not have a length-scale comparable with a characteristic indentation contact dimension. In either of these two cases, length-scale effects arise, quantified by a ratio of the contact geometry to film thickness or contact geometry to microstructural length-scale or inhomogeneous material characteristic feature size. The additional length-scale introduces features of the indentation response that are beyond the scope of the current work, which are well-covered elsewhere (Rar et al., 2002; Oyen et al., 2004; Constantinides et al., 2006; Oyen and Ko, 2008).

Elastic

The simplest and most straightforward data analysis occurs in the case of purely elastic deformation, where a direct fit of the load–displacement \((P-h)\) data to the elastic indentation solution can be performed. For a spherical indenter the elastic relationship is (Johnson, 1985):

\[
P = \frac{4\sqrt{R}}{3} \frac{E}{(1-\nu^2)} h^{3/2}
\]

(6)

where \( \nu \) is the Poisson’s ratio of the material and \( R \) is the tip radius. For a conical or pyramidal indenter the elastic relationship is Sneddon (1965):

\[
P = \frac{\pi \tan \psi}{2\nu^2} \frac{E}{(1-\nu^2)} \frac{h^2}{R^2} \]

(7)

where \( \gamma = \pi/2 \), the ratio of the total depth to contact depth in the theoretical expression (Sneddon, 1965); experimental values for have been found to deviate substantially from this value, especially in compliant materials (Thurn et al., 2002; Oyen and Bushby, 2007).

Elastic–plastic

A number of authors have considered the deconvolution of indentation load–displacement \((P-h)\) data to obtain the elastic modulus \((E)\) or plane strain modulus \((E/(1-\nu^2))\) and
the contact hardness ($H_C$) (Doerner and Nix, 1986; Oliver and Pharr, 1992; Field and Swain, 1993). Here we only consider the Oliver and Pharr’s analysis; a less common approach for elastic–plastic indentation data analysis, specific to spherical indentation is the approach of Field and Swain and the reader is referred to other sources (Field and Swain, 1993; Bushby, 2001) for details on this approach. In the Oliver–Pharr approach, the load–displacement data from unloading is fit to a simple power-law expression:

$$P = \alpha (h - h_f)^n$$

with fitting parameters $\alpha$, $h_f$, $m$. From these parameters, the stiffness at peak load ($S = dP/dh|_{P_{\text{max}}}$) is calculated. The contact depth is computed from the stiffness and peak point ($P_{\text{max}}$, $h_{\text{max}}$):

$$h_c = h_{\text{max}} - \beta \frac{P_{\text{max}}}{5}$$

where $\beta$ ($\leq 1$) is a dimensionless geometry parameter related to the shape of the probe (typically taken as $\beta = 0.75$ for a paraboloid of revolution). A calibration function $A_C(h_C)$ is used to obtain the contact area. The stiffness and contact area are used to calculate the reduced modulus ($E_R$, the series sum of the plane strain moduli of the material and the indenter) via

$$E_R = \frac{S\sqrt{\pi}}{2\sqrt{A_C}}$$

and the contact hardness by Eq. (1). In practice, the user of a commercial indentation instrument sees very little of the curve-fitting or computation. The typical operator feeds the area function calibration for a particular indenter tip $A_C(h_C)$ into the instrument software, and the parameters ($E$ or $E_R$ and $H_C$) are automatically output by the instrument software. The accuracy of the obtained modulus and hardness are completely dependent on the user-defined calibration function, and no real understanding of the contact mechanics is actually required. It is for this reason that the approach has become so common, and that many attempts have been made to adapt the approach to different deformation modes, such as viscoelastic or viscoelastic–plastic (Feng and Ngan, 2002). However, it is unclear as to why (outside of convenience due to the “built-in” nature of the Oliver–Pharr approach) this analysis would be used in all cases of material deformation, when more appropriate analysis techniques for visco deformation, in particular, are available and remarkably simple to implement.

Viscous–elastic

We have considered two cases in the context of VE material behavior, the linearly viscoelastic responses based on empirical exponential functions and as typically considered for polymeric materials including elastomers (Findley et al., 1989), and the poroelastic responses considering fluid flow through a linearly elastic porous medium (Wang, 2000). The latter case of poroelastic behavior does not have a closed-form analytical solution for indentation, and the reader is referred to computational studies (Selvadurai, 2004) and approximations to numerical solutions (Agbezuge and Dersiewicz, 1974; Oyen, 2008b). For considerations of analytical viscoelasticity, the viscoelastic indentation response is straightforward (Lee and Radok, 1960; Johnson, 1985) and data analysis is extremely simple: a single curve-fit just as in the case of the elastic expressions (Eqs. (6) and (7)). The key difference is that the elastic expressions are a fit to load–displacement ($P$–$h$) data, while the viscoelastic case involves a fit to displacement–time ($h$–$t$) data for a particular loading history $P(t)$. In the case where a displacement could be instantaneously applied, the displacement–time creep response would be of the form:

$$h^\gamma(t) = k_G P_{\text{max}}J(t)$$

where for a sphere $n = 3/2$ and the geometrical constant $k_G = 3/(8R^{1/2})$; for a cone or pyramid $n = 2$ and $k_G = \pi \tan \psi/(2\gamma^2)$. In the case where the creep function $J(t)$ is a straightforward Prony series,

$$J(t) = C_0 - \sum C_i \exp (-t/\tau_i)$$

the displacement–time ($h$–$t$) response for step-loading creep is of the form:

$$h^\gamma(t) = B_0 - \sum B_i \exp (-t/\tau_i).$$
A fit to Eq. (13) for the B coefficients and the time constants ($\tau_j$) leads directly to the relaxation function (Eq. (12)), where the time constants are the same in both expressions and the C values are obtained by algebraic rearrangement of Eqs. (11)–(13). This simple exercise was explored above in Eqs. (11)–(13) for a step-loading condition, which is experimentally impossible, but the experimentally tractable condition of creep following a finite constant loading-rate ramp, $P(t)=kt$, with a total ramp time $t_0=P_{\text{max}}/k$, also gives (Oyen, 2005, 2006b, 2007) a simple expression of the form of Eq. (13):

$$h^n(t) = k_0P_{\text{max}} \left[ C_0 - \sum C_i \exp \left(-t/\tau_i\right) RCF_i \right]$$

which differs from the step-hold creep response (Eq. (11)) by a ramp correction factor ($RCF_i$) (Oyen, 2007) given by

$$RCF_i = \frac{\tau_i}{\tau_R} \left[ \exp \left(t_0/\tau_i\right) - 1 \right].$$

Again, by direct comparison of Eqs. (14) and (13), the coefficients of the relaxation function (Eq. (12)) can be obtained by straightforward arithmetic calculations. These viscous–elastic expressions (here for a viscoelastic solid with a finite equilibrium modulus) hold for cases in which the contact impression recovers in a time-dependent manner following unloading, in contrast to the permanent indentation impressions resulting in a purely viscous material or one with a plastic deformation component.

In the case of dynamic sinusoidal loading for frequency-based measurements, the input is an oscillation at force amplitude $P_0$, where this value $P_0$ is a small perturbation on the primary peak force $P_{\text{max}}$ such that $P_0 \ll P_{\text{max}}$ (Fig. 6) and $P(t) = P_{\text{max}} + P_0 \sin(\omega t)$ for frequency $\omega$. The response is a displacement of amplitude $h_0$ and a phase shift $\phi$ such that the response to the load perturbation has a $g_0 \sin(\omega t - \phi)$ dependence. The real ($E'$) and imaginary ($E''$) parts of the elastic modulus are then calculated where the complex modulus is $E'' = E' + iE''$ and

$$E' = \frac{1 - \nu^2}{2\sqrt{\nu}} \frac{S\sqrt{\pi}}{2\sqrt{AC}}$$
$$E'' = \frac{1 - \nu^2}{2\sqrt{\nu}} \frac{\omega D\sqrt{\pi}}{2\sqrt{AC}}$$

analogous to Eq. (10) and where $D$ is the damping coefficient.

The real and imaginary moduli can be calculated directly from:

$$E' = (1-\nu^2) \frac{\sqrt{\pi}}{2\sqrt{\nu}} \frac{P_0}{h_0} \cos \phi$$
$$E'' = (1-\nu^2) \frac{\sqrt{\pi}}{2\sqrt{\nu}} \frac{P_0}{h_0} \sin \phi.$$

The derivations of these expressions and caveats concerning their use can be found in other sources (Huang et al., 2004; Cheng et al., 2006; Herbert et al., 2008). In practice, a typical experiment would include a number of different measurements for different frequencies (“frequency sweep”) and the evaluation of the real and imaginary parts of the complex modulus as a function of frequency.

Viscous–elastic–plastic

The addition of plastic deformation to viscous and elastic deformation – such as when a sharp indenter is used on a glassy polymer – requires an addition to the viscoelastic analysis in order to account for the plastic deformation. There are several mechanisms for doing this within the framework of the viscoelastic analysis described above. One clever way to do this is to perform a multi-stage test where the indenter is penetrated into the surface, unloaded almost completely but not enough to lose contact, and then after a delay to allow for viscoelastic recovery (and thus effectively reset the viscoelastic time-clock to $t = 0$) reloaded in the same location (Zhang et al., 2005). The idea is predicated on the idea that the plastic deformation will take place during the first loading cycle, and that after unloading the analysis can be purely viscoelastic. This method in fact relies on the same concept as the Oliver–Pharr analysis, that the indentation of a “hole” caused by plastic deformation is not fundamentally different in terms of the elastic (or here, viscoelastic) deformation.

There is a number of empirical models for direct analysis of VEP materials. For a relatively simple series model of viscous, elastic and plastic deformation elements (Oyen and Cook, 2003), two different techniques have been developed for doing a single curve fit to unloading data from a triangle-wave load–unload cycle. The first involves direct fits to the displacement-time data (Oyen and Ko, 2007) and the second uses analysis of the data in normalized coordinates (Cook and Oyen, 2007).

Brittle–elastic–plastic

The relationship between indentation load and crack length and knowledge of the indentation geometry and mode (E or EP) enables the fracture resistance or toughness (the combination $(RE)^{1/2}$ is recognized as $T$) to be determined from indentation tests. Simple scaling analyses show that the equation above relating $P/c^{3/2}$ to $T$ describes many crack geometries at indentation contacts, relying as it does on the concept of a center loaded crack (Lawn, 1993; Cook and Pharr, 1990). Hence the length of cracks oriented perpendicular to the surface or nearly so, such as cone cracks at spherical E contacts and half-penny and radial cracks at conical, pyramidal, or spherical EP contacts are all described by Eq. (3) with different values of the geometry term $c_3$. [See Cook and Pharr (1990) for exact definitions of theses crack types and their evolution during the indentation contact cycle—the complicated and transient indentation stress fields generated in large-$c$ EP materials with compact plastic deformation zones lead to a variety of indentation fracture sequences, including crack propagation during unloading.] Perhaps the easiest technique to implement is that for Berkovich or Vickers EP indentation, for which the value of $c_3$ has been extensively studied and calibrated. Inversion of Eq. (3) gives

$$T = \frac{P/c^{3/2}}{a_3}$$

such that measurement of the crack length, $c$, at a given indentation load, $P$, or preferably the load-invariant average value of $(P/c^{3/2})$, enables $T$ to be estimated using (Anstis et al., 1981)

$$a_3 = 62.5(H/E)^{1/2}.$$
the indentation loads required to generate cracks (and thus exceed the brittleness index) are large enough such that direct observation is possible (Fig. 4c; the cracks may be difficult to observe clearly in translucent or opaque materials). For many applications, such observation may not be possible or desirable, in which case the load–displacement trace may be used to detect the threshold load or displacement for cracking. Discontinuities in the gradient of the P–h trace (either \( \frac{dP}{dh} \) (Toivola et al., 2002) or, more sensitively, \( \frac{dP}{dh^2} \) (Morris et al., 2004) can be correlated with observed indentation cracking and thence used as a means to determine the cracking threshold: \( P^* \) can be taken as the load at which discontinuities appear, although of course this will depend on the indenter chosen for the study.

As noted, \( P^* \) can be decreased significantly through the use of acute probes, suggesting that toughness could be estimated at extremely small length scales using the approach outlined above. Superficially, radial cracks generated by small-scale nanoindentation using acute probes such as the cube corner appear very much like such cracks generated at much larger contacts using obtuse probes such as the Berkovich or Vickers. Examination of the data for such cracks however, shows that the conventional indentation fracture analysis based on an EP indentation is incorrect, and that the predominant driving force for acute indentation fracture is an elastic wedging field not present during obtuse indentation. In fact, for most materials, the plastic deformation zone contribution to fracture during acute indentation is negligible (Morris and Cook, 2005; Morris et al., 2005), and the \( \alpha_3 \) formulation given above is not appropriate for acute fracture. Determination of cracking thresholds by discontinuities in load–displacement traces is still applicable however, for qualitative comparisons of indentation fracture behavior.

4. Conclusion

In this paper we have reviewed the different types of deformation present during an indentation test, and the experimental means for isolating different deformation modes during indentation. We have described different material properties that can be obtained from indentation studies exploring these modes of deformation, including the elastic modulus, the contact hardness, resistance to plastic deformation, time constants for time-dependent deformation and the fracture resistance. We note that by simply restricting nanoindentation testing to the most common test techniques – using Berkovich probes and Oliver–Pharr data analysis, to obtain only two of these parameters (\( E \) and \( H_c \)) – one can miss the richness of data that comes from exploring a larger range of deformation modes. We have shown how simple curve fits and arithmetic calculations can be used to identify a range of material properties, to characterize fully the deformation modes found by varying the indenter probe geometry and experimental test conditions in a range of material types.

The techniques and analyses presented here are particularly suited for exploring the mechanical behavior of biological materials by nanoindentation testing. While the description elastic–plastic, and thus the Oliver–Pharr method for data analysis, may apply in a few limited cases for tissues in the body, such as dehydrated bone or dentin, most biological tissues do not exhibit elastic–plastic indentation responses. Tooth enamel has been shown to be brittle–elastic, and the toughness has been measured successfully by both microindentation (Xu et al., 1998) and nanoindentation (Marshall et al., 2001). Soft tissues are viscous–elastic and have been characterized with models based on viscoelasticity (Hayes and Mockros, 1971; Mattice et al., 2006), using a poroelastic framework (Mak and Mow, 1987) or incorporating both poroelasticity and flow-independent viscoelastic responses (Huang et al., 2001). In the case of hydrated bone, depending on indenter acuity, the observed response can vary. Bone behavior can be considered as viscous–elastic during blunt indentation, using viscoelastic (Bembey et al., 2006) or poroelastic (Oyen, 2008b) analyses. For sharp indenters the response can be considered as viscous–elastic–plastic (Oyen and Ko, 2007). In all of these specific cases of biological materials, a richness of information is found in the indentation response when an analysis method is selected appropriate for the modes of deformation actually present in the tissue under contact-probe loading.

REFERENCES
