Scaling, dimensional analysis, and indentation measurements

Yang-Tse Cheng\textsuperscript{a,},* Che-Min Cheng\textsuperscript{b}
\textsuperscript{a}Materials and Processes Laboratory, General Motors Research and Development Center, Warren, MI 48090, USA
\textsuperscript{b}State Key Laboratory for Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100080, China

Available online 19 July 2004

Abstract

We provide an overview of the basic concepts of scaling and dimensional analysis, followed by a review of some of the recent work on applying these concepts to modeling instrumented indentation measurements. Specifically, we examine conical and pyramidal indentation in elastic–plastic solids with power-law work-hardening, in power-law creep solids, and in linear viscoelastic materials. We show that the scaling approach to indentation modeling provides new insights into several basic questions in instrumented indentation, including, what information is contained in the indentation load–displacement curves? How does hardness depend on the mechanical properties and indenter geometry? What are the factors determining piling-up and sinking-in of surface profiles around indents? Can stress–strain relationships be obtained from indentation load–displacement curves? How to measure time dependent mechanical properties from indentation? How to detect or confirm indentation size effects? The scaling approach also helps organize knowledge and provides a framework for bridging micro- and macroscales. We hope that this review will accomplish two purposes: (1) introducing the basic concepts of scaling and dimensional analysis to materials scientists and engineers, and (2) providing a better understanding of instrumented indentation measurements.

\#2004 Published by Elsevier B.V.

Keywords: Scaling; Analysis; Indentation

1. Introduction

The perception of hardness is perhaps as ancient as human existence. It is the natural consequence of our ability to sense the surrounding world through physical contact. In fact, we develop an intuitive appreciation of “hard” and “soft” as infants before we learn to talk or walk. However, it is only fairly recent (about 100 years or so) that quantitative scales of hardness and measurement methods were developed for materials property characterization. The first modern technique for measuring the hardness of metals was due to Brinell [1], who was concerned about the consistency of steels produced by the company he was employed by. He took steel plates of each batch, placed a hard steel ball between them and squeezed them together in a vice. The size of the dents was chosen to represent the hardness of his steel. Today, a standard Brinell test consists of pressing a hard steel ball normally onto the surface of the metal under investigation. The Brinell hardness is given by the load divided by the surface area of the indentation. Other standard hardness tests have been developed and are routinely used, including the Vickers, Berkovich, Knoop, and Rockwell tests [2–4]. Typically, an indenter of a given geometry and material type, such as a sphere,
a cone, or a pyramid made of steel or diamond, is pushed into the solid. Upon unloading, a permanent depression in the surface of material is examined with a microscope. The hardness value is usually defined as the ratio of the indentation load and either the surface or projected area of residual indents, though the depth of indentation may also be used as a measure of hardness (i.e., in the Rockwell indentation test). These standard hardness tests are vital in nearly all areas of materials science and engineering.

Since the early 1970s, instrumented micro- and nano-indentation techniques have been developed and are now widely available [5–11]. In instrumented indentation experiments, load and indenter displacement are recorded simultaneously during the entire loading and unloading process. These instruments also allow the precise control of either load or displacement during indentation experiments. Instrumented indentation can be accurately made using forces as small as a few micro Newtons over depths in the nano-meter range (Fig. 1). Thus came the term “nanoindentation”. Together with improved understanding of indentation mechanisms and advances in analysis methods, such as that due to Doerner and Nix [12] and Oliver and Pharr [13], instrumented indentation has become an indispensable tool for probing mechanical properties at small length scales.

Since instrumented indentation can also be performed at the load range typical of conventional Vicker’s and Rockwell hardness tests, it is one of the few experimental techniques that can be performed at both micro- and macro-scales, allowing the investigation of materials behavior across length scales from nano- to millimeters. Recently, the field of “multiscale” modeling has attracted significant attention [14–18]. Theoretical models are being developed for describing materials behavior using quantum mechanics for a few atoms, molecular dynamics for hundreds of millions of atoms, dislocation dynamics for interacting dislocations, and continuum mechanics at macroscopic scales. Because it offers an opportunity for comparison with experiments, modeling indentation from small number of atoms to continuum has become popular in recent publications on multiscale modeling [14,18–21].

Indentation measurements have been applied to a wide range of materials, ranging from metals, ceramics, polymers, and composites. The application areas include microelectronics, optoelectronics, coatings for low-emission window glasses, and tribological coatings [22–28]. Presently, there is also a growing interest in probing biological materials [29–35] and food products [36,37]. Instrumented indentation measurements have also been used in the study of the deformation and fracture of rocks for a better understanding of rock mechanics related to geological evolution of the planets [38,39].

A review of every aspects of indentation is beyond the scope of the present paper, though there are several reviews [40–46], bibliographies on indentation modeling [47,48], books [2,3,49,50], special journal volumes [24,27,28], and topical conference proceedings [22–26] dedicated to the broad areas of indentation hardness measurements. Instead, this review will focus on a scaling approach to indentation modeling [51–55]. This approach helps understand several basic questions in instrumented indentation measurements, including:

- What information is contained in the indentation load–displacement curves?
- What is hardness? How does hardness depend on mechanical properties and indenter geometry?
- Can stress–strain relationships be obtained from indentation load–displacement curves?
- How to measure time-dependent mechanical properties from indentation?
- How to detect or confirm indentation size effects?

We show that the scaling approach to indentation modeling provides some new insights into these issues. It also helps organize knowledge and provide a framework for bridging micro- and macro-scales.
We begin with a tutorial on dimensional analysis. Although dimensional analysis is routinely used in certain branches of science and engineering, most materials scientists and engineers are not exposed to this powerful analysis tool. The basic idea of dimensional analysis is that the physical laws do not depend on the arbitrariness in the choice of units of physical quantities. This concept often allows the number of arguments in functions describing physical phenomena to be reduced, thus making them simpler to obtain either from calculations or experiments. This review shows that dimensional analysis is very valuable to the analysis of a range of indentation problems.

Fig. 1. Examples of early experimentally measured load-displacement curves using instrumented micro- and nano-indentation techniques: (a) single crystal Sb [5]; (b) single crystal Ge [5]; (c) electropolished AISI 52100 steel [8]; and (d) electropolished nickel [10]. In [10], it was reported that hardness values measured at depth as small as 20 nm.
Since indentation measurements have been applied to a great variety of materials, ranging from metals, ceramics, polymers, to biomaterials, it is necessary to model indentation using generalized though simplified descriptions for the mechanical properties of solids, including the power-law work-hardening, power-law creep, and linear viscoelastic effects. Dimensional analysis is then used to explore various aspects of indentation in these model systems with the aid of finite element calculations when necessary. The focus is on indentation using self-similar conical and pyramidal indenters with some discussions on spherical indenters.

This review is intended to serve two purposes: (1) introducing the basic concepts of scaling and dimensional analysis to materials scientists and engineers, and (2) providing a better understanding of instrumented indentation measurements.

2. Indenter geometry, geometric similarity, scaling, and dimensional analysis

2.1. Indenter shapes and sizes

Indenters of different geometry such as the conical, pyramidal and spherical of various sizes are available for indentation measurements. Several practical considerations are behind these choices. For example, it is more convenient to make a three-sided than a four-sided sharp pyramidal diamond tip, because of the fact that three intersecting polished planes converge naturally to a single apex, while merging four planes at one point requires much more skills. Furthermore, although technologies exist for making hardened steel ball bearings, making very small size perfectly spherical diamond balls is nontrivial and would require expensive techniques such as ion milling using focused ion beams. Aside from these practical considerations, certain shapes of indenters possess the desirable property of self-similarity in indentation tests.

2.2. Geometric similarity

The concept of geometric similarity was already established in Euclid’s time. We find, in his “Elements” Volume VI, “Similar rectilinear figures are such as have their angles severally equal and the sides about the equal angles proportional” [56]. Thus, for any two geometric objects that can be described by a group of finite number of length and angle parameters, they are similar if their length parameters are respectively proportional to each other with the same proportional constant and all the respective angles identical, i.e., \( \alpha = \alpha' \); \( \beta = \beta' \); \( \ldots \); \( \omega = \omega' \) and \( \frac{a'}{a} = \frac{b'}{b} = \cdots = \frac{n'}{n} = c = \text{constant} \). An example is illustrated in Fig. 2.

The factor \( c \) is called the similarity or transformation coefficient. We can generate a family of geometrically similar objects by varying the proportional constant. More broadly, two geometric objects are called geometrically similar if one can be transformed into the other by a continuous, distortionless, linear transformation. For example, all circular cones and equal-based pyramids of infinite length are geometrically similar provided their apex angles are the same. In this case there is no length parameter and the apex angle is dimensionless, i.e., having a value irrespective of the size of the cone. Ideally sharp conical and pyramidal (i.e., Vickers and Berkovich) indenters possess no characteristic length scale.

All spheres and circles are geometrically similar as well. A sphere is defined by its radius, which is a length parameter. Two spheres of radius \( r \) and \( R \) can be transformed from one to another by scaling with a factor \( R/r \). Conical or pyramidal indenters with spherical caps of different radii are also geometrically similar provided that their apex angles are the same. The scaling factor is the ratio of tip radii.
The concept of geometric similarity can be extended beyond treating purely geometric problems. In this review, the geometric similarity of conical, pyramidal, and spherical indenters will be exploited in the modeling of indentation experiments.

2.3. Similarity and scaling

While Galileo Galilei is perhaps most famous for his celestial discoveries and the falling-weight experiment on the leaning tower of Pisa, it was in the context of examining the strength of materials that he discovered scaling laws [57]. In his book [58], “Dialogues Concerning Two New Sciences”, Galileo developed a theory of strength of materials for beams, rods, and other geometrical forms in a series of eight Propositions. He formulated his Propositions in terms of ratios of physical quantities, such as the length and diameter of beams and cylinders, rather than the actual physical sizes themselves. For example, he proved in Proposition V that: “Prisms and cylinders which differ in both length and thickness offer resistance to fracture [i.e., can support at their ends loads] which are directly proportional to the cubes of the diameters of their bases and inversely proportional to their lengths” (see Fig. 3a). This proposition shows how the strength of geometrically similar objects varies with their relative size. Since the weight of an object is proportional to the cubes of its size \(L^3\), and the strength is proportional to its cross-section area (i.e., proportional to \(L^2/L = L^2\), according to Proposition V), it follows that the strength to weight ratio scales as \(L^{-1}\) for geometrically similar objects. Hence, Galileo concluded that geometrically similar objects are not equally strong under the action of their own weight. The beams become weaker with increase in dimensions and finally they fail when they are too large. Galileo then arrived at some far reaching statements: “the impossibility of increasing the size of structures to vast dimensions either in art or in nature; likewise the impossibility of building ships, palaces, or temples of enormous size, . . . nor can nature produce trees of extraordinary size, . . . Thus, a small dog could probably carry on his back two or three dogs of his own size; . . . a horse could not carry even one of his own size . . . indeed the smaller the body the greater its relative strength”. In order to accomplish similar functions (i.e.,
sustaining their weight), large animals have evolved to have bones much thicker than that of small animals (Fig. 3b). This explains why large (i.e., elephants) and small animals (i.e., rats) are not geometrically similar. These examples help illustrate that, in the examination of mechanical behavior of solids, similarity can play an important role.

2.4. Dimensional analysis

The scaling concepts form the basis for dimensional analysis. The essential idea of dimensional analysis [59] is that physical laws do not depend on arbitrarily chosen basic units of measurement. Consequently, the functions that express physical laws must process certain mathematical property, called generalized homogeneity, i.e., each of the additive terms in the functions will have the same dimensions or units. This property allows the number of arguments in the mathematical expressions to be reduced, thus making them simpler to obtain either from theories or experiments. This basic idea leads to the central theorem in dimensional analysis, the so-called PI-theorem (or II-theorem), which has been attributed to Buckingham [60].

The notion of dimensional analysis was in use long before the explicit formulation of the II-theorem by such luminaries as Galileo, Newton, Fourier, Maxwell, Reynolds, and Rayleigh [61]. Several well-celebrated cases can be found in the monographs [59,62–65], recent papers [66,67], and historical perspectives [68]. An early application of dimensional analysis to hardness tests was cited on page 124 in Ref. [2]. Three recent overviews on some aspects of scaling concepts applied to

Fig. 3. Illustrations from Galileo’s “Two New Sciences” [58]: (a) two cylinders used in proofing Proposition V; and (b) large animals have evolved to have bones much thicker than that of small animals.
contact mechanics problems are also listed [69–71]. Dimensional analysis has also been used to aid the analysis of indentation-induced-fracture [72–75], a subject that is not further discussed in this paper since several recent comprehensive reviews are available [76,77]. This section provides an introduction and a recipe for dimensional analysis. We will then apply it to indentation problems in later sections of this review.

2.4.1. Physical quantities, units, and dimensions

In general, a physical quantity \( x \) is expressed in terms of a number and the number is obtained by measuring the physical quantity. Measurement is the direct or indirect comparison of a certain quantity with an appropriate standard or unit of measurement. For example, the height of a person is measured to be 1.7 m. There are no absolutely preferred units for physical quantities. The height of the person can be 170 cm or \( 1.7 \times 10^9 \) nm. Obviously, it is more convenient to use meter to describe the height of people than nanometer. The units for measuring physical quantities are divided into two categories, \textit{fundamental} units and \textit{derived} units. Once the units for length and time, taken as fundamental units, are specified, the unit for velocity is then derived. A system of units is a set of fundamental units that is both necessary and sufficient for measuring the properties of a class of phenomena. In Newtonian mechanics, for example, there are three basic units and the units of other physical quantities are called derived units. The three basic units are usually taken as that of length, time, and mass, respectively in the length/time/mass system, although there exist other choices (e.g., the length/time/force system). The number is associated with the unit in the following manner: when the unit is shortened by a factor \( X \), the value will be increased by a factor \( X \). We say that the dimension of \( x \), denoted by \([x]\), is equal to \( X \). Note that the dimension of any physical quantity is invariant with respect to the units chosen. In this sense dimension is an objective quantity. In the length/time/mass system, the dimensions for length, time, and mass are denoted by \( L \), \( T \), and \( M \), respectively. The velocity with the unit of length/time has the dimension \( LT^{-1} \).

It can be shown rigorously that the dimension of any physical quantity is always a power-law monomial [59]. Specifically, the dimension \([z]\) of any physical quantity \( z \) in the length/time/mass system is a function of \( L \), \( T \), and \( M \) in the following form

\[
[z] = L^a T^b M^c,
\]

where the exponents are real numbers associated with \( z \). In fact, it is intuitively obvious that \([z]\) cannot be \( \log(L/T) = \log(L) - \log(T) \) or \( Le^M \). Eq. (1) is a consequence of that all systems within a given class of units are equivalent. In the specific example of length/time/mass, the cm-s-g and m-h-kg systems are equivalent. A quantity is dimensionless if all \( a, b, c, \ldots \) are zero in a given system of units.

A group of physical quantities \( x_1, \ldots, x_i \) is said to be dimensionally independent if there does not exist a set of real numbers \( a_1, \ldots, a_i \), not all equal to zero, such that

\[
[x_1]^{a_1} [x_2]^{a_2} \cdots [x_i]^{a_i} = 1.
\]

Consequently, one cannot form a dimensionless number out of dimensionally independent quantities. We note that in the framework of Newtonian mechanics \( i \) cannot exceed 3. The basic units such as length, time, and mass are a priori dimensionally independent. When the basic units are chosen differently, formally different but in actual fact the same relations hold, for example, instead of the mass/length/time system one may choose the force/length/time system of basic units. When macroscopic thermal phenomena are also under consideration, we need to include an additional
physical quantity and the additional basic unit is usually taken as that of temperature. More basic
units are needed when considering mechanical-thermal-electrical systems.

2.4.2. Physical laws and relations
In physical sciences and engineering quantitative laws or relations can be expressed, at least
conceptually, in mathematical form as

\[ z = f(z_1, \ldots, z_n), \]

(3)
to show that the dependent quantity \( z \) is a function of the independent variables and parameters \( z_1 \) to
\( z_n \). Let there be \( i \) dimensionally independent quantities in this relation with \( i \leq n \). Without loss of
generality, we may set them to be \( z_1 \) to \( z_i \). The remaining variables are then dimensionally dependent.
Consequently, their dimensions can be expresses as

\[
[z] = [z_1]^{a_1} \cdots [z_i]^{a_i} \\
[z_{i+j}] = [z_1]^{a_1} \cdots [z_j]^{a_j}, \quad j = 1, \ldots, n - i.
\]

(4)
Hence, the following \( n - i + 1 \) quantities formed out of the original \( n + 1 \) \( z \)'s are dimensionless.

\[
\Pi = \frac{z}{z_1^{a_1} \cdots z_i^{a_i}} \\
\Pi_j = \frac{z_{i+j}}{z_1^{a_1} \cdots z_i^{a_i}}, \quad j = 1, \ldots, n - i.
\]

(5)
The existence of the exponents in the above expressions is guaranteed by the fact that \( z \) as well as \( z_{i+j} \)
are dimensionally dependent on \( z_j \) (\( j = 1, \ldots, i \)). We note that the construction of the \( \Pi_j \)'s ensures
that they are all independent, since each of them contains an element that is not present in any of the
other \( \Pi_j \)'s.

Eq. (3) can then be written in terms of the new variables \( \Pi \) and \( \Pi_j \) as:

\[
\Pi = f(z_1, \ldots, z_i, \Pi_1, \ldots, \Pi_{n-i}).
\]

(6)
Now by passing from one system of basic units to another the values of \( z_1 \) to \( z_i \) can be made
arbitrarily large or small while the values of the \( \Pi \)'s remain unchanged. Consequently the
requirement that physical relations be objective and independent of units forbids the appearance of
the \( z \)'s in the above equation. Hence, we have

\[
\Pi = f(\Pi_1, \ldots, \Pi_{n-i}).
\]

(7)
Eq. (7) differs from Eq. (3) in that the number of variables is reduced by \( i \), the number of
dimensionally independent quantities, and all the variables in Eq. (7) are dimensionless. This is
known as the \( \Pi \)-theorem. The \( \Pi \)-theorem is intuitively obvious since it is clear that physical laws
should not depend on the choice of units. When \( n - i \) is small this advantage can be quite significant
as will be seen when we consider indentation modeling in later sections.

2.4.3. A recipe and caution on the use of dimensional analysis
Dimensional analysis has broad applications. It is also the foundation of modeling. For
experimentalists it plays an essential role from design of experiment to data analysis. To modelers it
helps to establish equations and yield solutions. Unfortunately, dimensional analysis can also be
misused. To facilitate its successful applications, we provide the following recipe for dimensional analysis:

1. Listing independent variables and parameters that the quantity of interest (i.e., dependent variable) depends on. There should be a relationship for each dependent quantity.
2. Identifying independent variables and parameters with independent dimensions.
3. Forming dimensionless quantities and establishing relationships among dimensionless quantities.

The number of relationships is equal to the number of dependent quantities.

The key to the successful application of dimensional analysis is to first recognize which are the dependent variables and which are the independent variables. One needs to be very careful in choosing the independent variables and parameters, keeping the relevant variables and parameters and leaving out those that are irrelevant. The deeper one understands the nature of the problem on hand the better can one make such choices. In forming the dimensionless II’s one is actually measuring the relevant physical quantities by means of measures inherent to the problem rather than externally imposed gauges such as meters, seconds, and grams. When a certain II’s are very small (or large) in comparison with other dimensionless quantities, one can often drop these II’s all together from a dimensionless relation. But, extreme care should be exercised in doing so as has been pointed out by, for example, Barenblatt [59].

3. Simple models for mechanical behavior of solids

The success of continuum mechanics is its predictive models based on a relatively few measurable parameters that provide reasonably good predictions about the bulk mechanical behavior of materials across many length scales, enabling the construction of bridges, highways, skyscrapers, airplanes, and automobiles. It has been recognized that certain mechanical properties of materials are structurally insensitive and size independent and others are structurally sensitive and size dependent [78–81]. Properties such as the elastic modulus are size independent, the measured values from macroscopic experiments are consistent with that predicted from first-principles quantum mechanics calculations. Other properties are size-dependent, for example the yield strength of polycrystalline solids which depends on the grain size and the mechanical/thermal history that the samples have experienced. In this section, we summarize several simple models for elastic, elastic–plastic, creep, and viscoelastic behavior of materials. These models allow detailed analysis of indentation problems, thus providing insights that might be obscured if more complex mechanics models are used.

3.1. Isotropic linear elastic materials

Most materials deform elastically when deformation is small. The stress–strain relationships of isotropic linear elastic materials can be expressed in two equations by separating the deviatoric stress and strain, $s_{ij}$ and $d_{ij}$, from the dilatational stress and strain, $\sigma_{ii}$ and $\varepsilon_{ii}$, as follows [78]:

$$
\begin{align*}
    s_{ij} & = 2Gd_{ij} \\
    \sigma_{ii} & = 3Ke_{ii},
\end{align*}
$$

(8)

where $G$ is the shear modulus and $K$ is the bulk modulus representing the inverse of compressibility. Using the definition of $s_{ij}$ and $\sigma_{ii}$ together with Eq. (8), the stress–strain relationship can be
expressed as
\[ \sigma_{ij} = 2G\epsilon_{ij} + \lambda\delta_{ij}\epsilon_{zz} = \frac{E v}{1 + v} \epsilon_{ij} + \frac{Ev}{(1 + v)(1 - 2v)} \delta_{ij}\epsilon_{zz}, \]  
(9)

where \( \lambda \equiv K - (2/3)G \) and \( G \) are called Lamé’s constants; \( E \) is Young’s modulus and \( v \) is Poisson’s ratio. The dimension of \( \lambda, G, \) and \( E \) is the same as that of stress and \( v \) is dimensionless. The mechanical properties of isotropic linear elastic solids can be prescribed by any two of these parameters of these parameters, for example, \( E \) and \( v \).

3.2. Power-law work-hardening

With increasing degree of deformation, plastic deformation occurs, leading to permanent deformation. Although many mechanisms of plastic deformation exist, it has been found that, at least for metals, a fairly good approximation is power-law work-hardening \([79,80]\) and the stress–strain curves under uniaxial tension are assumed to be given by

\[ \sigma = E\epsilon, \quad \text{for} \quad \epsilon \leq \frac{Y}{E}, \]

\[ \sigma = C\epsilon^n, \quad \text{for} \quad \epsilon \geq \frac{Y}{E}, \]

(10)

where \( Y \), the initial yield stress, and \( C \), the strength coefficient, have the same dimension as stress and \( n \), the strain-hardening exponent, is dimensionless. To ensure continuity, we note \( C = \sqrt{E/Y} \). Consequently, either \( E, Y, \) and \( C \) or \( E, Y, \) and \( n \) are sufficient to describe the stress–strain relationship. We use the latter set of parameters extensively in the following discussions. When \( n \) is zero, Eq. (10) becomes the model for elastic-perfectly plastic solids. For most metals \( n \) has a value between 0.1 and 0.5 \([79]\) (Fig. 4a).

3.3. Power-law creep

Time dependent plasticity \([78–80]\) occurs when materials are deformed at an elevated temperature, often greater than 0.5 \( T_m \), where \( T_m \) is the absolute melting temperature. Creep is time-dependent plasticity under constant stress at elevated temperatures. Three regions are typically delineated (Fig. 4b): Stage I or primary creep is typically characterized by increasing strain with decreasing strain rate. Stage I is followed by Stage II, known also as the secondary or steady-state creep, where the strain increases at a constant strain rate. Eventually, cavitation and/or cracking cause the strain to increase with increasing rate until rupture takes place. This regime is termed State III. In the steady-state creep regime, which usually persists for substantial portion of the material’s life, the strain-rate sensitivity of materials can be modeled by a simple power-law relation,

\[ \sigma = b\epsilon^m, \]

(11)

where \( m \) and \( b \) are materials constants. The dimension of \( b \) is (stress-time\(^m\)) and \( m \) is dimensionless.

3.4. Isotropic linear viscoelastic materials

Many soft materials, such as polymers and biomaterials, deform viscoelastically. Generalizing Eqs. (8) and (9) to isotropic linear viscoelastic solids, the stress–strain relations are usually given
Fig. 4. Simple models of mechanical behavior of solids: (a) power-law work-hardening; (b) power-law creep; and (c) viscoelastic relaxation modulus.
by \[82–84\]

\[
s_{ij}(t) = 2 \int_0^t G(t - \tau) \frac{\partial d_{ij}(\tau)}{\partial \tau} \, d\tau
\]

\[
\sigma_{ii}(t) = 3 \int_0^t K(t - \tau) \frac{\partial e_{ii}(\tau)}{\partial \tau} \, d\tau,
\]

where \(G(t)\) is the stress relaxation modulus in shear and \(K(t)\) is the hydrostatic stress relaxation modulus (Fig. 4c). Equivalently,

\[
\sigma_{ij}(t) = \int_0^t \left[ 2G(t - \tau) \frac{\partial e_{ij}(\tau)}{\partial \tau} + \lambda(t - \tau) \delta_{ij} \frac{\partial \sigma_{kk}(\tau)}{\partial \tau} \right] \, d\tau,
\]

where \(\lambda(t) = K(t) - (2/3)G(t)\). The dimension of \(\lambda(t), K(t),\) and \(G(t)\) is the same as that of stress. The time dependent Poisson’s ratio and Young’s modulus may then be defined as

\[
v(t) = \lambda(t)/[2(\lambda(t) + G(t))]\]

and

\[
E(t) = 2G(t)(1 - v(t)),
\]

respectively.

Alternatively, the stress–strain relations may be written as

\[
d_{ij}(t) = \int_0^t J_1(t - \tau) \frac{\partial s_{ij}(\tau)}{\partial \tau} \, d\tau
\]

\[
e_{ii}(t) = \int_0^t J_2(t - \tau) \frac{\partial \sigma_{ii}(\tau)}{\partial \tau} \, d\tau,
\]

where \(J_1(t)\) and \(J_2(t)\) are the two independent creep functions for isotropic viscoelastic solids. Obviously, Eqs. (12) and (14) are not independent of each other. In fact, the relationships between the relaxation and creep functions, \(G(t)\) and \(J_1(t)\), as well as \(K(t)\) and \(J_2(t)\), have simple forms after transforming them using Laplace transformation \[82–84\].

In the following, we will systematically investigate indentation in “model systems” that are characterized by Eqs. (10)–(12). We will focus on conical indentation in homogeneous materials with emphasis on finding general relationships for load–displacement curves and contact depths from which the hardness of materials can be evaluated (Fig. 5). These results provide a basis for understanding indentation in materials with more complex mechanical behaviors and configurations (i.e., layered structures, composites, and functional gradients).

4. Indentation in elastic–plastic solids with power-law work-hardening

4.1. Shape of indentation loading curves

An indentation loading curve is the relationship between load, \(F\), and displacement, \(h\), that can be continuously measured during an indentation experiment. The aim of this section is to review the approaches used to describe indentation loading curves, including expressions derived from dimensional analysis for conical and pyramidal indentation in homogeneous elastic–plastic solids. This section serves to answer questions such as what is the expected shape of indentation loading curves and what information is contained in these curves?

The contact between a rigid conical indenter normally loaded into a flat specimen is well known for two types of materials, elastic materials and rigid-plastic materials. For elastic materials, an
analytic expression was obtained by Love [85], Harding and Sneddon [86], and Sneddon [87] between the load, \( F \), and the penetration depth, \( h \), in the case of a nonadhesive rigid conical punch, normally loaded, on the surface of a smooth elastic body as

\[
F = \frac{2Eh^2}{(1-v^2)\pi} \tan \theta \equiv K_e h^2.
\]  

(15)

The load is proportional to the square of the indenter displacement during both loading and unloading. The proportionality factor, \( K_e \), depends on both the elastic parameters \( (E/(1-v^2)) \) as a group and indenter geometry, \( \theta \).

For rigid-plastic solids, Tabor [88], applying the principal of geometric similarity, shows that the mean pressure acting on a conical or pyramidal indenter is the same whatever the size of indentation, as long as the material is uniform and homogeneous. He found, for fully work-hardened metals, the semi-empirical relation for the mean pressure

\[
P_m = C_0 Y,
\]  

(16)

where \( C_0 \) is function of indenter geometry and \( Y \) is the yield strength. Since the mean pressure was empirically found to be independent of the depth of indentation, the load is thus proportional to the square of indentation depth,

\[
F = K_p h^2.
\]  

(17)

More generally for a uniform and homogeneous elastic–plastic solid, the equation for the loading curves cannot be obtained analytically and it is often assumed to be given by

\[
F = K_{ep} h^2,
\]  

(18)

where \( K_{ep} \) is a function of materials properties and indenter geometry. Several equations have been proposed that link \( K_{ep} \) to \( K_e \) and \( K_p \) [89]. It has also been proposed that the indentation loading
curves can be more generally described by the so called Meyer’s law,
\[ F = Kh^x, \]
where \( x \) is not necessary an integer. In the following, the origin of the square dependence and its implications are examined. The dependence of \( K \) on the mechanical properties of materials and indenter geometry will also be explored.

We now apply the recipe of dimensional analysis to conical indentation in an isotropic elastic–plastic solid obeying the work-hardening rule Eq. (10). We assume the indenter is rigid and without surface roughness. The friction coefficient between the indenter and material surface is assumed zero, though nonzero friction will be discussed later. First, we must select the dependent variable and identify all independent variables and parameters. If we choose the force on the indenter, \( F \), as the dependent variable, the indenter displacement, \( h \), is clearly an independent variable. The mechanical properties of materials, Young’s modulus \( (E) \), Poisson’s ratio \( (v) \), initial yield strength \( (Y) \), and work-hardening exponent \( (n) \), are independent parameters. The indenter half angle \( (\theta) \), which characterizes the indenter geometry, is also an independent parameter. On the other hand, quantities such as stress or strain under the indenter are not independent variables, but dependent variables. Parameters such as temperature, humidity, and the rate of indentation are not relevant in the present problem, because these parameters are outside the scope of the model defined by Eq. (10). After identifying all the independent variables and parameters, thus accomplishing Step 1 in the recipe of dimensional analysis, we have a general expression, \( f_I \), for the loading curve,
\[ F = f_I(E, v, Y, n, h, \theta). \]

The next step is to identify, among the six governing parameters, \( E, v, Y, n, h, \) and \( \theta \), the ones with independent dimensions. This is accomplished by noting that two of them, namely \( E \) and \( h \), have independent dimensions (i.e., dimensions of stress and length). The dimensions of \( Y, v, n, \) and \( F \) are then given by \( [Y] = [E], [v] = [E]^0[h]^0, [n] = [E]^0[h]^0, [\theta] = [E]^0[h]^0, \) and \([F] = [E][h]^2 \).

We are now ready to take the third step of dimensional analysis. By applying the \( \Pi \)-theorem in dimensional analysis, we obtain:
\[ \Pi_2 = \Pi_2(\Pi_1, v, n, \theta), \text{ or equivalently } F = Eh^2\Pi_2\left(\frac{Y}{E}, v, n, \theta\right), \]

where \( \Pi_2 = F/Eh^2 \), \( \Pi_1 = Y/E^2 \), \( v, n, \) and \( \theta \) are all dimensionless.

Based on the above dimensional analysis, we can make several important observations for a rigid conical indenter with a given half angle, \( \theta \), indenting into an elastic–plastic solid with work-hardening. First, the force on the indenter, \( F \), is proportional to the square of the indenter displacement, \( h \). This square dependence is common to conical indentation in purely elastic (i.e., when \( Y \to \infty \)), rigid plastic (i.e., \( E \to 0 \)), elastic–perfectly plastic (i.e., \( n = 0 \)), and elastic–plastic solid with work-hardening. The coefficient of the square dependence is of course different for these cases. Second, the parameter, \( F/Eh^2 \), is a function of \( Y/E, v, n \), for a given \( \theta \). Third, the original problem of a function of six parameters is reduced by 2, the number of parameters with independent dimensions in Eq. (20). It is now a function of four parameters, thus simplifying the original problem of six parameters and allowing a systematic evaluation of the effects of each parameter. Finally, the numerical values of the function \( \Pi_2(Y/E, v, n, \theta) \) cannot be known from dimensional analysis alone and must be obtained either through experiments or modeling. Using a finite element analysis software, ABAQUS [90], the function \( \Pi_2(Y/E, v, n, \theta) \) has now been calculated over a wide range of parameter space [51–53,91,92].
Fig. 6. Examples of indentation curves obtained from finite element analysis: Loading and unloading curves in a solid with $E = 200$ GPa, $Y = 2.0$ GPa, $\nu = 0.3$, and several values of $n$ [53].

Fig. 6 shows examples of the calculated loading curves for several sets of values of $E$, $Y$, and $n$ with a Poisson’s ratio 0.3. The loading curves were fitted with a power function: $F = ah^n$, where $a$ and $x$ are two fitting parameters. The exponent, $x$, obtained from all simulations, such as that shown in Fig. 6 is between 1.98 and 2.03. Since errors unavoidably exist in all numerical calculation, we do not expect finite element calculations to produce the exact square dependence. They are, nevertheless, consistent with predictions from dimensional analysis (Eq. (21)).

The relationships between $F/Eh^2$ and $Y/E$ are illustrated in Fig. 7 for a given indenter geometry. As expected, work-hardening has a greater effect on the force required to move the indenter for smaller ratio of $Y/E$. For a large ratio of $Y/E$, the value of $F/Eh^2$ approaches that for purely elastic contact with or without work-hardening. Fig. 7 shows that, in general, loading curves from conical indentation alone cannot be used to detect whether material work-hardens because work-hardening does not change the square-dependence. Furthermore, loading curves alone cannot uniquely determine both $Y/E$ and $n$. Thus, caution must be exercised when mechanical properties of solids are extracted by matching the computed indentation loading curves with that obtained from experiments.

Fig. 7. Scaling relationship between $F/Eh^2$ and $Y/E$ [53].
Although the square dependence was derived for conical indentation, the same approach can be applied to indentation using pyramidal indenters, since pyramidal indenters are also geometrically self-similar. The square dependence is also not limited by the assumptions of rigid indenter and zero friction between indenter and solid. For example, we consider indentation in elastic–plastic solids with work-hardening using an elastic, pyramidal indenter. Let $\Theta$ be the set of angles that describe the shape of a pyramidal indenter, $E_i$ the Young’s modulus of the indenter, $v_i$ its Poisson’s ratio, and $\mu$ the friction coefficient for the interface between the indenter and solid, dimensional analysis shows that for loading the force is given by

$$F = Eh^2 \Pi_s \left( \frac{Y}{E}, v_i, n, \frac{E_i}{E}, v_i, \mu, \Theta \right). \quad (22)$$

Obviously, three-dimensional finite element calculations are needed to evaluate the dimensionless functions in Eq. (22). Such calculations, though computationally intensive, are within the current capabilities of commercial finite element software. It is evident, however, that the square depth dependence remains the same as that for conical indentations in elastic–plastic solids.

Several authors have proposed explicit equations for indentation loading curves for conical and pyramidal indenters. The square dependence on the indentation depth is usually either assumed or verified by finite element calculations, although details vary depending on the underlining assumptions. For example, Larsson et al. [91,92] proposed expressions for indentation loading curves for Vickers and Berkovich indenters based on extensive finite element calculations. Their equation for loading curves was inspired by that of the spherical cavity model and was parameterized in terms of several quantities, including a representative stress corresponding to a “representative” strain of about 0.3. Sakai [93] and Rother [94] suggested that the prefactor in the square-depth dependence is a function of work of indentation. Hainsworth et al. [95] derived an equation for the loading curves where the prefactor contains a function of $HIE^s$, the ratio of hardness to reduced modulus. A modification of the loading curve equation first proposed by Hainsworth et al. was recently made by Malzbender et al. [96] and Troyona and Martin [97].

From dimensional analysis, it is evident that the square dependence is a consequence of the absence of a length scale. Therefore, it should hold for more complex constitutive laws which do not contain length scales, as has been shown recently by Li et al. [98] for anisotropic elastic solids, other plasticity models, and materials with finite compressibility. The dimensionless pre-factor depends on a large number of dimensionless parameters, though the square dependence remains valid.

On the other hand, the square dependence is expected to breakdown if there exists a relevant length parameter, $l$. Indeed, we note from dimensional analysis,

$$F = Eh^2 \Pi_s \left( \ldots, \frac{h}{l}, \ldots \right). \quad (23)$$

The loading curve is thus no longer proportional to the square of indenter displacement. For example, we have shown [99] that loading curves can depart from the square dependence if the conical or pyramidal indenter has a finite tip radius $R$ (Fig. 8). In this case, the tip radius, $R$, enters the above equation as a length parameter in place of $l$. In fact, it has been shown that both tip radius, $R$, and load frame compliance can be determined from the loading curve, thus facilitating the calibration of indentation systems [100].

Hence, the square dependence of loading curves should hold for sharp conical and pyramidal indentation in homogeneous solids without a length parameter. A deviation from the square
dependence signals the presence of length scales or “size effects”. The measurement of the shape of loading curves is therefore a potentially powerful method for detecting relevant length parameters and “size effects”.

4.2. Contact depth, “sinking-in” and “piling-up” phenomena

It is well known that the materials immediately outside the contact area of an indentation usually deform and do not remain flat [2,3]. For a heavily work-hardened metal, such as copper or mild steel, the surface of the specimen around the indentation bulges upwards or piles-up, and for a fully annealed metal, the surface of the specimen around the indentation sinks in [101–103]. Both piling-up and sinking-in behavior are schematically illustrated in Fig. 5. The piling-up and sinking-in phenomena are of interest because they affect the accurate determination of contact areas and thus the hardness measurements. Other interesting issues include whether the pile-up area carries load, the optical appearance of indents on surfaces, and the use of the shape of the indents to infer the mechanical properties of the solids.

Dimensional analysis can help identify the parameters affecting the degree of piling-up and sinking-in [104]. In general, the contact depth is a function of all independent parameters

\[ h_c = g(E, v, Y, n, h, \theta). \]

As we have seen in the previous section, the dimensions of \( Y, v, n, h \), and \( h_c \) can be expressed in terms of the dimension of \( E \) and \( h \), the two parameters with independent dimensions. Applying the
\[ P_{\beta} = P_{\beta}(P_1, \nu, n, \theta), \quad \text{or equivalently,} \quad h_c = h \Pi_{\beta} \left( \frac{Y}{E}, \nu, n, \theta \right), \]  

(25)

where \( P_{\beta} = h_c/h \), \( P_1 = Y/E \), \( \nu \), \( n \), and \( \theta \) are all dimensionless. The linear dependence between \( h_c \) and \( h \) in Eq. (25) has been verified by extensive finite element calculations for a wide range of values of \( E \), \( Y \), and \( n \). Fig. 9 displays the relationship between the calculated \( h_c/h \) and \( Y/E \) for several values of \( n \). It is apparent that the ratio \( h_c/h \) is a function of both \( Y/E \) and \( n \), as predicted by dimensional analysis. Furthermore, the value of \( h_c/h \) can be either greater or smaller than one, corresponding to the piling-up and sinking-in of the displaced surface profiles, respectively. For large \( Y/E \), sinking-in occurs for all values of \( n > 0 \). For small \( Y/E \), both sinking-in and piling-up may occur depending on the degree of work-hardening. In the case of severe work-hardening (i.e., \( n = 0.5 \)), sinking-in is expected even for very small values of \( Y/E \), whereas piling-up is expected for elastic-perfectly plastic solids and for solids with a small work-hardening exponent (e.g., \( n = 0.1 \)).

These simulation results are expected from analytical theories of indentation in elastic solids where sinking-in is expected [87] and in rigid-plastic solids where piling-up is predicted [105]. They are consistent with experimental observations of sinking-in and piling-up phenomena reported in the literature. A recent AFM study of piling-up and sinking in Cu is shown in Fig. 10 [106]. These previous experiments clearly demonstrated the influence of work-hardening on piling-up and sinking-in. Previously, it has been suggested that piling-up and sinking-in was determined by the work-hardening exponent [101,107,108]. Others have shown that for highly elastic solids, such as polymers, sinking-in is often observed [88]. Although it was generally understood that the Young’s modulus, initial yield stress, and work-hardening exponent all play a role in controlling piling-up and sinking-in, these experiments alone do not provide a full picture of the individual contributions of various parameters. Dimensional analysis together with finite element calculations have made our present understanding more clear: For conical and pyramidal indentations, the magnitude of piling-up and sinking-in is determined by neither \( n \) alone nor all three parameters \( Y, E \), and \( n \). It is determined by two parameters \( Y/E \) and \( n \) for a given indenter.
4.3. Relationships between hardness and mechanical properties of solids

Several hardness definitions exist in the standards or protocols for indentation measurements. The one commonly used in recent instrumented indentation literature is given by

$$ H = \frac{F}{A_c} $$

(26)
where $A_c$ is the projected contact area under load $F$. The traditional definition of hardness is

$$H_{\text{trad}} = \frac{F}{A_{\text{trad}}},$$

(27)

where $A_{\text{trad}}$ is the residual projected area of the hardness impression after load removal. The two hardness values can differ significantly, when $A_c/A_{\text{trad}} \neq 1$. This is, for example, the case of indentation in rubbers where $A_c$ is finite and $A_{\text{trad}}$ approaches zero. We will use primarily the first definition (Eq. (26)) in this review.

Hardness, once defined, allows ranking of materials by their hardness values. While it is useful to assign materials their hardness, it has been of persistent interest to understand relationships between hardness and fundamental properties of materials and address questions such as “Is hardness a material property?” “Is it determined by the plasticity of solids?” “Does elastic property play a role?” “Can indenter geometry and loading conditions affect hardness measurements?” More generally, “What is hardness?” This section helps provide some answers to these questions by establishing relationships between hardness, elastic modulus, yield stress, work-hardening exponent, and indenter geometry.

As early as 1945, Bishop et al. [109] suggested that the stress distribution under a conical indenter might be approximated by that of a spherical or cylindrical cavity. Using known solutions from the newly developed plasticity theory for spherical and cylindrical cavities, the authors established a relationship between the mean pressure under the indenter and Young’s modulus and yield strength, as well as work-hardening behavior. For elastic-perfectly plastic solids, the pressure at which the cavity expands depends on the ratio of Young’s modulus to yield strength and Poisson’s ratio according to

$$p_Y = \frac{2}{3} \left( 1 + \ln \frac{E}{3(1-v)Y} \right).$$

(28)

In a series of experiments with a variety of materials Marsh [110] found that his hardness results followed a pattern very close to that of equation above, but the constants were somewhat different. An extension of the spherical cavity was made by Johnson in the early 1970s [111]. He pointed out that the radial displacement of material lying on the elastic–plastic boundary must accommodate the volume of materials displaced by the indenter during indentation. He derived an equation for conical indenters of semi-angle $\theta$ and Poisson’s ratio 0.5:

$$p_Y = \frac{2}{3} \left( 1 + \ln \frac{E \cot \theta}{3Y} \right).$$

(29)

This equation has been used often to provide an approximate, analytical description of conical indentation in elastic-perfectly plastic solids, though questions remain about the validity of the expanding cavity model for indentation problems, such as the simplification of the stress-field [112] and the neglect of the piling-up and sinking-in effects [51].

Based on slip-line field solution for indentation in a rigid-plastic solids by a frictionless rigid wedge and experimental observations of indentation in metals with elastic-perfectly plastic behavior, Tabor [113] established a relationship between hardness, $H$, and yield stress, $Y$,

$$H = 3Y.$$  

(30)
For most metals experiencing work-hardening, the indentation process itself produces an increase in the yield or flow stress in the materials. The plastic strain will consequently vary over the deformed region so that the amount of work hardening will vary from point to point. Although the strain distribution under an indenter is complex, Tabor suggested that there is a "representative" or "average effective" strain such that the indentation hardness is

$$H = 3Y_0,$$

where $Y_0$ is the uniaxial flow stress at some specified strain value. For a Vickers indenter, this representative strain is between 0.08 and 0.1 and the indentation hardness is then 3 times the yield stress at this strain value. Tabor's relationship appears in nearly every textbook on mechanical behavior of solids. Yet long-standing concerns remain. For example, "what is the physical meaning of the average effective strain?" "Where is the location under the indenter that materials actually experience the representative strain?" [114] "Can Tabor's relationship be applied to materials other than metals?" It has also been recognized by Tabor [115] that the hardness versus yield stress relationship is inconsistent with Johnson's spherical cavity model, for the latter Eq. (29) does not suggest a constant ratio between $H$ and $Y$. Thus, it seems necessary to examine these well known relationships further.

The scaling relationships established in Sections 4.1 and 4.2 provide an opportunity for systematic investigation of these questions. Using Eqs. (21), (25) and (26) the ratio of hardness to initial yield strength is given by

$$H = \cot^2 \theta \pi \frac{\Pi_2(\frac{Y}{E}, v, n, \theta)}{(\frac{Y}{E})^2 \Pi_0^2(\frac{Y}{E}, v, n, \theta)},$$

or, equivalently,

$$H = \frac{\cot^2 \theta}{\pi} \frac{\Pi_2((\frac{Y}{E}, v, n, \theta)}{\Pi_0^2((\frac{Y}{E}, v, n, \theta)},$$

These two equations show that hardness, $H$, is independent of indentation depth $h$ for conical indentation in the class of elastic–plastic solids that obey Eq. (10). In particular, this conclusion is valid for both elastic-perfectly plastic solids (i.e., $n = 0$) and for work-hardening solids (i.e., $n > 0$). It is not surprising since there is no length scale involved in either the materials model or indenter geometry as discussed in Sections 4.1 and 4.2. Indeed, the common origin of the depth independent hardness and the square-depth dependence of loading curves is the absence of a length parameter in the model for elastic–plastic deformation (Eq. (10)) and in describing the indenter geometry.

Furthermore, finite element calculations of $\Pi_2$ and $\Pi_0$ would allow a systematic investigation of the relationships between hardness and $\frac{Y}{E}$, $v$, $n$, and $\theta$. As an example, we plot $H/Y$ versus $Y/E$ in Fig. 11 for $\theta = 68^\circ$ and $v = 0.3$. It is apparent that, over the practically relevant range of $Y/E$, the ratio $H/Y$ is not a constant. The hardness, $H$, depends on $E$, $Y$, and $n$. As expected, work-hardening has a greater effect on the hardness value for small ratio of $Y/E$. For large ratios of $Y/E$, the hardness value approach 1.7 times the initial yield strength, $Y$, and is insensitive to $n$. For small ratios of $Y/E$, the hardness value can be many times that of the initial yield strength, $Y$.

Following Tabor's idea of representative strain, we evaluated $H/Y_0$, where $Y_0 = K\varepsilon_0^0$ is the representative stress at a representative strain $\varepsilon_0$. We noted an approximate scaling relationship between $H/(K\varepsilon_0^0)$ and $Y/E$, if the value for strain, $\varepsilon_0$, is taken to be 10%. In Fig. 12, we plot $H/$
(0.10^n K) against Y/E. All the data points shown in Fig. 11 lie approximately on a single curve. Thus, the concept of representative strain seems applicable. Furthermore, this analysis suggests that the concept of representative is not necessarily linked to the actual location of areas under the indenter where the strain value is that given by the representative strain.

It is also evident from Fig. 12 that H/Y0 is a function of Y/E and is, therefore, not a constant over the wide range of Y/E (Fig. 12). For Y/E < 0.02, H/Y0 is about 2.4 to 2.8 (Fig. 12) which is consistent with Tabor’s relationship Eq. (31). For Y/E > 0.06, however, H/Y approaches 1.7 (Fig. 11). Thus, Tabor’s relationship Eq. (31) should be modified as

$$H = \begin{cases} 
2.8Y_0, & \text{where } Y_0 \text{ is the yield stress at } 10\% \text{ of strain, for } \frac{Y}{E} \to 0.0 \\
1.7Y, & \text{for } \frac{Y}{E} \to 0.1.
\end{cases} \quad (34)$$

It is also instructive to interrogate Eq. (33) by plotting H/E versus Y/E (Fig. 13). This figure shows that, for given $E$ and $v$, there are multiple choices of $Y$ and $n$ that produce the same hardness values.

![Fig. 11. Scaling relationship between $H/Y$ and $Y/E$ for several values of $n$ [53].](image1)

![Fig. 12. Scaling relationship between $H/(K_{\epsilon_0}^n)$ and $Y/E$, if the value for strain, $\epsilon_0$, is taken to be 10% [53].](image2)
Thus, hardness is a multiple function of materials mechanical properties. The practical implication is that one can achieve a desired hardness value by judicially engineering $E$, $Y$, and $n$. Furthermore, hardness is indenter geometry dependent as evident from several classic experiments [116] and recent calculations [117].

This section demonstrates that dimensional analysis together with finite element calculations can be used to unveil the relationships between hardness, mechanical properties, and indenter geometry.

4.4. Unloading-curve, initial unloading slope, contact area, and elastic modulus

For indentation in purely elastic solids, unloading curves would retrace loading curves. They do not, therefore, contain any new information. In contrast, unloading curves for indentation in elastic–plastic solids are thought to possess more information and have attracted much study. In particular, the initial unloading slope has been used for the determination of elastic modulus and hardness, for example, the Oliver–Pharr procedure [13]. While this procedure is routinely used, the underlying assumptions are not always understood by the practitioners. Thus, basic questions remain. For examples, should unloading curves follow the square-law depth dependence? Does the Oliver–Pharr procedure depend on the elastic unloading assumption for the whole unloading curve? What is the range of validity of the procedure? In this section, some of these and related questions are studied with the aid of dimensional analysis.

Since unloading takes place after loading during which the indenter reaches a maximum depth, $h_m$, there is one additional independent parameter, $h_m$, in the unloading equation. Thus, the force, $F$, is now a function, $f_U$, of seven independent governing parameters: $E$, $v$, $Y$, $n$, $h$, $\theta$, and $h_m$,

$$ F = f_U(E, v, Y, n, h, h_m, \theta). $$

(35)

Dimensional analysis yields

$$ \Pi_j = \Pi_j(\Pi_1, \Pi_2, v, n, \theta), $$

or equivalently,

$$ F = \frac{Yh^2}{E} \Pi_1 \left( \frac{Y}{E}, \frac{h}{h_m}, v, n, \theta \right), $$

(36)

where $\Pi_j = F/Eh^2$, $\Pi_1 = Y/E$, $\Pi_2 = h/h_m$, $v$, $n$, and $\theta$ are all dimensionless. In contrast to loading, Eq. (36) shows that the force, $F$, is, in general, no longer simply proportional to the square of the
indenter displacement, \( h \), it also depends on the ratio, \( h/h_m \), through the dimensionless function \( \Pi' \). In general, we should not, therefore, expect unloading curves to follow square depth dependence, independent of whether unloading is purely elastic or not.

A special point of interest on an unloading curve is the final indentation depth, \( h_f \), at which the force on the indenter first becomes zero during unloading. Using the general expression for unloading Eq. (36), we can formally solve for \( h_f \), i.e.,

\[
F = Eh_m^2 \Pi' \left( \frac{Y}{E}, \frac{h_f}{h_m}, v, n, \theta \right) = 0. \tag{37}
\]

Evidently, \( h_f/h_m \) is a function of \( Y/E, v, n \), and \( \theta \). Consequently, \( h_f/h_m \) is independent of indentation depth \( h_m \) or load \( F_m \) (the force at the maximum indentation depth, \( h_m \)).

Many people have studied indentation unloading curves for a variety of materials \([13,118–138]\). It has been suggested \([13]\) that the unloading curves could be described by

\[
F = b(h - h_f)^m. \tag{38}
\]

Some authors treated \( b \) and \( m \) as numerical fitting parameters. Others believed them to be materials constants. The exponent \( m \) was usually found between 1.2 and 1.7, although \( m \) equals 2 was also hypothesized. Recently, Pharr et al. \([123,124]\) advanced the concept of an “effective shape” of indenters that takes into account the observation that the unloaded shape of indents is not perfectly conical or pyramidal but exhibits a subtle convex curvature. They analyzed elastic reloading and found that the exponent \( m \) is related to the shape of the deformed surface, thus providing the parameter, \( m \), a physical meaning. This idea has been further developed by Woirgard et al. \([125,126]\). Fundamentally, however, the primary difference between loading and unloading curves is the difference in contact conditions, i.e., initially flat surfaces versus deformed surfaces. The presence of the length parameter, \( h_m \), in the unloading equation is a consequence of this change in contact condition, leading to the exponent \( m \) differing from 2.

We now consider the initial unloading slope \( df/dh \). Taking the derivative with respect to the indenter displacement and evaluating it at \( h_m \), Eq. (36) becomes:

\[
\left. \frac{df}{dh} \right|_{h=h_m} = Eh_m \left[ \Pi'_\gamma \left( \frac{Y}{E}, 1, v, n, \theta \right) + 2\Pi'_\gamma \left( Y/E, 1, v, n, \theta \right) \right]. \tag{39}
\]

As a result, the dimensionless quantity,

\[
\frac{1}{Eh_m} \left. \frac{df}{dh} \right|_{h=h_m} = \Pi'_\gamma \left( \frac{Y}{E}, 1, v, n, \theta \right) + 2\Pi'_\gamma \left( Y/E, 1, v, n, \theta \right) \equiv \Pi_\delta \left( \frac{Y}{E}, v, n, \theta \right), \tag{40}
\]

is independent of \( h \) and is a function, \( \Pi_\delta \), of \( Y/E, v, n \), and \( \theta \). Eq. (40) shows that the initial unloading slope is proportional to \( h_m \) or \( F_m^{1/2} \).

Using the above conclusion we now examine the validity of Eq. (38) as a general expression for unloading curves. Differentiating Eq. (38) with respect to indenter displacement \( h \), we obtain:

\[
\left. \frac{1}{h_m} \frac{df}{dh} \right|_{h=h_m} = bhm_m^{m-2} \left( 1 - \frac{h_f}{h_m} \right)^{m-1}. \tag{41}
\]

Since \( h_f/h_m \) is independent of \( h_m \), we note that \( (1/h_m)(df/dh)|_{h=h_m} \) depends on \( h_m \), unless \( m = 2 \), which contradicts the fact that \( m \) does not equal to 2 in general. Alternatively, the parameter \( b \)
must be proportional to \( h_m^{-(m-2)} \), in order to make \( (1/h_m)(dF/dh) \) depth independent. Thus, the parameter \( b \) depends on indentation depth. Of course, we may avoid this problem by applying Eq. (38) only to a given indentation depth \( h_m \) and treat parameter \( b \) as a fitting parameter. In other words, Eq. (38) is not a generally valid equation for describing unloading curves for all indentation depths. It is only an interpolation formula for given \( h_m \). The parameter \( b \) is, therefore, not a material constant.

As another application of the scaling relationships, we investigate the relationship between the initial unloading slope, the contact radius, \( a \), materials properties, and indenter geometry. Using Eqs. (25) and (40) we obtain

\[
\frac{1 - v^2}{E_a} \left. \frac{dF}{dh} \right|_{h=h_m} = \tan \theta \frac{\Pi_d(Y/E, v, n, \theta)}{\Pi_p(Y/E, v, n, \theta)}. \tag{42}
\]

Eq. (42) is compared with the well-known expression for unloading slopes:

\[
\frac{1 - v^2}{E_a} \left. \frac{dF}{dh} \right|_{h=h_m} = 2. \tag{43}
\]

Eq. (43) was initially derived for indentation into elastic solids [6,12,13]. Using the \textit{infinitesimal} theory of continuum mechanics, we have recently generalized Eq. (43) to the initial unloading in elastic–plastic solids using indenters with axisymmetric smooth profiles [127]. Furthermore, we have shown that Eq. (43) holds true even for materials with work-hardening and residual stresses [127].

Eq. (42) may be considered a special case of Eq. (43), where the right hand side of Eq. (42) is 2. This observation suggests a systematic way of examining the unloading slope equation through either experiments or numerical analysis. Indeed, several groups have studied unloading slope equation using finite element analysis [52,53,128–136]. For example, we have evaluated \([1 - v^2]/E_a)(dF/dh)\) for \( \theta = 68^\circ \) and \( v = 0.3 \). The results are plotted against \( Y/E \) in Fig. 14. It is apparent from Fig. 14 that the quantity \([1 - v^2]/E_a)(dF/dh)\) is roughly independent of \( Y/E \) and \( n \). The numerical value, obtained from averaging all the cases presented in Fig. 14, is given by \([1 - v^2]/E_a)(dF/dh)\) \( h=h_m \approx 2.17 \pm 0.05 \). The key point is that \([1 - v^2]/E_a)(dF/dh)\) is approximately a constant independent of piling-up or sinking-in of the surface profiles around the indenter.

Fig. 14. Results from finite element calculations showing that \([1 - v^2]/E_a)(dF/dh)\) is approximately a constant greater than 2.0 over a wide range of \( Y/E \) and \( n \) [53].
Several authors [93,133–136] have calculated for different indenter angles and found that 
\[
\frac{(1 - \nu^2)}{Ea} \left( \frac{dF}{dh} \right)_{h=h_m}
\] also has a weak dependence on \( \theta \) and on \( \nu \). The \( \theta \) and \( \nu \) dependence can be understood by the fact that Eq. (43) was derived using linearized boundary conditions and infinitesimal theory of continuum mechanics, finite element calculations take into account nonlinear effects, including large strain and moving contact boundaries. Therefore, the small difference in numerical values for 
\[
\frac{(1 - \nu^2)}{Ea} \left( \frac{dF}{dh} \right)_{h=h_m}
\] between that given by Eqs. (42) and (43) is not unexpected. In fact, Tankaka and Koguchi [133] and Bolshakov and Pharr [134] have shown that 
\[
\frac{(1 - \nu^2)}{Ea} \left( \frac{dF}{dh} \right)_{h=h_m}
\] is about 2.09 and 2.16, respectively, after taking into account the finite radial displacements of points along the surface of contact. While their analyses were for conical indentation in elastic solids [133,134], Eq. (42) can be generalized to elastic–plastic solids with work-hardening. Taking into account these corrections, the unloading slope equation is typically written as
\[
\frac{dF}{dh} \bigg|_{h=h_m} = \beta \frac{2E}{\sqrt{\pi(1-\nu^2)}} \sqrt{A}. \tag{44}
\]
where \( \beta \) is found to be somewhat greater than 1 for a Berkovich indenter. Oliver and Pharr have recently [138] reviewed the improvements made to their procedure since the publication of their seminal paper in 1992. They proposed that \( \beta = 1.05 \) based on their analysis of available results from experiments and finite element calculations that \( 1.0226 \leq \beta \leq 1.085 \). Thus, the elastic constant, \( E/(1 - \nu^2) \), can be calculated from the initial slope of unloading curves, provided that either the contact depth, \( h_c \), radius, \( a \), or area, \( A \), is known. The question is then how to accurately determine \( h_c, a, \) or \( A \) under load for all cases of piling-up and sinking-in of surface profiles around the indenter. 

The most widely used the method for estimating the contact area under load is the procedure proposed by Oliver and Pharr [13] for estimating the contact depth from the initial unloading slope. The above dimensional and finite element analysis provide an opportunity to methodically evaluate this procedure.

Based on the results of Sneddon [87] on the shape of the surface outside the area of elastic contacts for an indenter of conical and paraboloid of revolution, Oliver and Pharr developed an expression for, \( h_c \), at the indenter displacement, \( h_m \),
\[
h_c = h_m - \zeta \frac{F_m}{(dF/dh)_m}. \tag{45}
\]
where \( F_m \) and \( (dF/dh)_m \) are the respective load and the initial slope of the unloading curve at the indenter displacement depth, \( h_m \). The numerical value of \( \zeta \) is 0.72 for conical indenter, 0.75 for the paraboloid of revolution, and 1.0 for flat punch. Although Eq. (45) was derived from solutions to elastic contact problems, it has been used to estimate contact depth for general elastic–plastic solids. The implicit assumption is that unloading is elastic at least during the initial stage of unloading. This assumption warrants a close inspection, because the contact condition prior to unloading is not the same as that for elastic contacts due to plastic deformation during loading.

Applying the Oliver and Pharr procedure to the loading–unloading curves obtained from finite element calculations, we evaluated [104] the contact depth using Eq. (45) and plot it in terms of \( h_c/h_m \) in Fig. 15a–d for elastic–perfectly plastic solids \( (n = 0.0) \) and for elastic–plastic solids with increasing degree of work-hardening \( (n = 0.1, 0.3 \) and 0.5). For comparison, the values of \( h_c/h_m \) that were directly obtained from finite element calculations (shown in Fig. 14) are also shown in Fig. 15a–d.
It is apparent from Fig. 15 that the Oliver and Pharr procedure for estimating the contact depth under load is valid when the ratio of $Y/E$ is large (e.g., $>0.05$ for $0.0 < n < 0.5$). This is expected since this procedure is based on Sneddon’s analysis of surface profiles for elastic contacts. Thus, the procedure may be used with confidence for highly elastic materials, including hard coating materials such as diamond-like carbons, carbides, and nitrides.

For materials with a wide range of $Y/E$ (e.g., $10^{-4}$ to $10^{-2}$), such as metals, the Oliver and Pharr procedure should be used with caution. For example, the procedure is approximately correct for most $Y/E$ values if the work-hardening exponent is approximately 0.3 (Fig. 15c). However, the procedure underestimates the contact area for elastic-perfectly plastic solids over most $Y/E$ values (Fig. 15a). The error is most significant when piling-up occurs, i.e., $h_c/h_m > 1$. In fact, the contact depth, $h_c$, estimated using Eq. (45) is always less than 1. It should also be noted that Eq. (45) could also overestimate contact area for materials with a large work-hardening exponent, e.g., $n = 0.5$ (Fig. 15d). Similar conclusions have also been obtained by Bolshakov and Pharr [135]. Thus, the Oliver–Pharr procedure should be used with caution when piling up occurs.

The piling up problem is most severe when $Y/E$ is small and $n$ is close to zero, which is also the condition for little elastic recovery or rebound upon unloading, i.e., $h_f/h_m \approx 1$. Consequently, the contact area under load is well approximated by that of residual indents. The shape and size of the indents after unloading can be measured using techniques such as atomic force microscopy, scanning electron microscopy, and optical profilometry. Following this line of thinking, we have developed a

Fig. 15. The values of $h_c/h$ obtained using the Oliver–Pharr procedure and those directly from finite-element calculations for (a) elastic–perfectly plastic solids and (b)–(d) elastic–plastic solids with increasing work-hardening exponent (b) $n = 0.1$, (c) 0.3, and (d) 0.5 [104].
method [139] that consists of imaging surface profiles and measuring load–displacement curves, circumventing the problem of piling-up (Fig. 16).

4.5. Work of indentation

By integrating the loading and unloading curves, the work of indentation can be readily obtained. Specifically, the area under the loading curve is the total work, \( W_{\text{tot}} \); the area under the unloading curve is the reversible work, \( W_u \); and the area enclosed by the loading and unloading curve is the irreversible work, \( W_p \) (see Fig. 17) of indentation. Obviously,

\[
W_{\text{tot}} = W_u + W_p.
\] (46)

Less obvious, however, are their physical meaning and significance. How do reversible and irreversible work relate to the mechanical properties of solids? Do they provide new information about indentation-induced elastic and plastic deformation? The equations for loading and unloading curves obtained in previous sections using dimensional analysis and finite element calculations can help answer these questions.

From Eq. (21), the total work done by the indenter, \( W_{\text{tot}} \), to cause elastic and plastic deformation when the indenter reaches a maximum depth, \( h_m \), is given by

\[
W_{\text{tot}} = \int_0^{h_m} F \, dh = \frac{E h_m^3}{3} \Pi_2 \left( \frac{Y}{E}, v, n, \theta \right).
\] (47)

Thus, the total work is proportional to \( h_m^3 \).
The work done by the solid to the indenter during unloading, \( W_u \), can be expressed, using Eq. (36), as

\[
W_u = \frac{1}{3} E h_m^3 \Pi_u \left( \frac{Y}{E}, x, v, n, \theta \right),
\]

where the integral, \( \int_{h_f/h_m}^{h_m} x^2 \Pi_y \frac{Y}{E}, x, v, n, \theta \) dx, is obviously a dimensionless function, designated as \( \Pi_u \), of \( Y/E, v, n, \) and \( \theta \), since \( h_f/h_m \) is a function of these parameters (see, Eq. (37)). Consequently, the work done on the indenter during unloading is also proportional to \( h_m^3 \). Note that the \( h_m^3 \) dependence is derived without resorting to assuming any explicit form for the unloading curves (e.g., Eqs. (19) and (38)).

The ratio of irreversible work to total work for a complete loading–unloading cycle, \( (W_{tot} - W_u) / W_{tot} \), can then be written as

\[
\frac{W_{tot} - W_u}{W_{tot}} = 1 - \frac{\Pi_u(Y/E, v, n, \theta)}{\Pi_y(Y/E, v, n, \theta)}.
\]

This ratio is, therefore, independent of \( h_m \).

Eqs. (49), (37) and (33) show that \( (W_{tot} - W_u)/W_{tot}, h_f/h_m, \) and \( H/E \) are all functions of \( Y/E, v, n, \) and \( \theta \). It is, therefore, instructive to investigate the connections among them. Indeed, we observed a correction between \( W_p/W_{tot} \) and \( h_f/h_m \) (Fig. 18) [140]. Initially, this correlation was found using data from finite element calculations for a given indenter geometry (i.e., a fixed \( \theta \)). Later, it was noted that the same correlation holds true for conical indenters of various angles (i.e., a range of \( \theta \) values). Furthermore, this correlation is approximately linear for \( h_f/h_m > 0.4 \) or for \( W_p/W_{tot} > 0.2 \). An interpolation formula, obtained using least square curve fitting to a large number of data from finite element calculations, was given in [140]:

\[
\frac{W_p}{W_{tot}} = \left( 1 + \gamma \right) \frac{h_f}{h_m} - \gamma, \quad \text{for} \quad \frac{h_f}{h_m} > 0.4,
\]

where \( \gamma = 0.27 \).
Several authors have previously explored correlations between $W_p/W_{tot}$ and $h_f/h_m$. Lawn and Howes [141] studied the elastic recovery effect in the indentation of several ceramic materials and steels. By assuming that the respective loading and unloading curves are given by $F = Ah^2$ and $F = B(h^2 - h_f^2)$, where $A$ and $B$ are constants, a relationship between $W_p/W_{tot}$ and $h_f/h_m$ was obtained,

$$
\frac{W_p}{W_{tot}} = \frac{1}{C_0} - \frac{3}{C_0} \left( \frac{h_f}{h_m} \right)^2 + 2 \left( \frac{h_f}{h_m} \right)^3.
$$

Eq. (51) is plotted in Fig. 18.

Loubet et al. [9] and Menčík and Swain [41] have also explored the relationship between $W_p/W_{tot}$ and $h_f/h_m$. The experimental results of Menčík and Swain for two types of steels and glass are also included in Fig. 18. By assuming that the respective loading and unloading curves are given by Eqs. (19) and (38), a relationship between $W_p/W_{tot}$ and $h_f/h_m$ is then given by

$$
\frac{W_p}{W_{tot}} = \frac{\alpha + 1}{m + 1} \left( \frac{h_f}{h_m} \right) - \frac{\alpha - 1}{m + 1}.
$$

Eqs. (50) and (52) are the same, if $\alpha = 2.0$ and $m = 1.362$.

Thus, several methods exist that corroborate a correlation between $W_p/W_{tot}$ and $h_f/h_m$. This relationship is independent of indenter geometry ($\theta$). It is also explicitly independent of the details of the materials properties, and the stress distribution under the self-similar indenters. The advantage of the approach based on dimensional analysis and finite element calculation is that it does not rely on the assumptions about the shape of the indentation loading and unloading curves.

This one-to-one correspondence suggests that $W_p/W_{tot}$ and $h_f/h_m$ contain the same information about the mechanical properties of materials so far as indentation tests are concerned. The measurement of one leads to another. In practice, however, the determination of $W_p$ and $W_{tot}$ can be made more accurately than that of $h_f$, since the former is from the integration of loading–unloading curves and the latter is from the estimate of a single point on the unloading curve. A potential application of this correlation is to predict surface deformation based on the work of indentation and vise versa.
From knowing \( \frac{W_{\text{tot}}}{C_0 W_u}/W_{\text{tot}} \) and \( H/E^* \) as functions of \( Y/E, n, \) and \( \nu \), we have recently \cite{142} established an approximately linear correlation between \( \frac{W_{\text{tot}}}{C_0 W_u}/W_{\text{tot}} \) and \( H/E^* \), first for a given indenter geometry (i.e., \( \theta = 68^\circ \)) (Fig. 19) and later for conical indenters of a range of angles \( 60^\circ \leq \theta \leq 80^\circ \) \cite{140}:

\[
\frac{H}{E^*} = \kappa \frac{W_u}{W_{\text{tot}}},
\]

(53)

where \( \kappa \) is a function of \( \theta \). From this relationship, the ratio \( H/E^* \) can be obtained readily by integrating the loading–unloading curves to obtain \( W_u/W_{\text{tot}} \). The ratio \( H/E^* \) is of significant interest in tribology. This ratio multiplied by a geometric factor is the “plasticity index” that describes the deformation properties of a rough surface in contact with a smooth surface \cite{143}. When the plasticity index is much less than unity, the deformation of asperities is likely to be entirely elastic. When it is significantly greater than one, the deformation is predominantly plastic. In fracture mechanics, the quantity of \( H/E^* \) appears in various fracture toughness equations \cite{49,144}. A few methods have been proposed to obtain this ratio, including that by Marshall et al. \cite{144} from the measurement of the diagonals of Knoop indentation. Recent experimental work suggests that the wear resistance of hard coatings is better correlated with \( H/E^* \), instead of hardness alone \cite{145}. The present method of obtaining \( H/E^* \) from \( W_u/W_{\text{tot}} \) is expected to facilitate the development of wear resistant coatings. It also provides a linkage between wear resistance and \( W_u/W_{\text{tot}} \), leading to an energy based method of predicting wear resistance.

Using the definition of hardness (Eq. (26)) and the equation for the initial unloading slope (Eq. (43)), it is readily shown that the ratio \( H/E_{\text{initial}}^\circ \) can be obtained from the initial slopes of unloading curves \cite{146,95,142} without the need of knowing the contact area. Together with the fact that \( H/E_{\text{initial}}^\circ \) can be obtained from \( W_u/W_{\text{tot}} \), the values of \( H \) and \( E^\circ \) can be obtained from the work of indentation and the initial unloading slope. This method of determining \( H \) and \( E^\circ \) does not require knowing the contact area, thus circumventing the difficulty caused by piling-up and sinking-in seen earlier in this review. Indeed, it has been observed that this energy-based method appears to provide the most reliable values for hardness and modulus for carbon nitride films that exhibit large elastic recovery upon unloading \cite{147,148}.

Although Eq. (53) was obtained for indentation in homogeneous solids. Recent work by Malzbender et al. \cite{149} suggested that it may also be applicable to thin films and coatings deposited on substrates, provided that fracture and delamination events do not influence the energy dissipation.
A number of authors have also, over the years, discussed work of indentation from several different perspectives. Stilwell and Tabor [150] examined the elastic energy released when the indenter is withdrawn and showed that it accounts quantitatively for elastic rebound observed when a conical indenter strikes an elastic–plastic solid. Ternovskii et al. [5], in one of the earliest papers on instrumented indentation, related the ratio of irreversible work to total work to an “interatomic bond parameter”. Sakai derived an equation relating the irreversible work to hardness [93]. Rother et al. developed a model for indentation size effect based on an energy principle [151]. Gubicza et al. [152] suggested a method of obtaining $H$ from $W_u/W_{\text{tot}}$. Bull and co-workers [153–155] extended the work of indentation concept to the modeling of composite hardness of multilayer thin films and coatings. Suresh and co-workers also found relationships among $(W_{\text{tot}} - W_u)/W_{\text{tot}}, h_f/h_m$, and $H/E^*$ using finite element calculations [156]. Recently, Malzbender [157] has compared experiments with several equations that pertain these relationships. Analogous relationships have also been found for spherical indentation, leading to a new method of obtaining elastic modulus and hardness from spherical indentation experiments [158].

This section showed that dimensional analysis could be used together with extensive finite element calculations to uncover previously unknown relationships, thus providing a deeper understanding of indentation in elastic–plastic solids.

### 4.6. Probing stress–strain relationships

We have seen from previous sections that hardness is a function of the mechanical properties, as well as indenter geometry. There can be multiple combinations of mechanical properties, such as $E$, $Y$, and $n$, that give rise to the same hardness values even when the indenter geometry is the same. Furthermore, hardness usually cannot be directly used in modeling unless it is converted to a yield strength. Thus, it is highly desirable to go beyond the measurement of hardness. Probing the local stress–strain behavior of materials by indentation has in recent years become a focal point of interest in the indentation literature [156,159–191]. One of the main questions we discuss in this section is whether it is possible to obtain the fundamental mechanical behavior from instrumented indentation experiments, an issue evidently was taken for granted until recently [161]. Dimensional analysis has made this question tractable. We have seen that “in principle” a unique loading and unloading curve can be obtained either from calculations or experiments for a given set of material properties (i.e., $E$, $v$, $Y$, and $n$) and indenter geometry ($\theta$). In practice, this “uniqueness” is, of course, subject to numerical or experimental errors. But, we have no doubt that the load–displacement curve for the same sample measured using the same indenter and instrument is reproducible everyday and anywhere. Likewise, the calculated load–displacement curves should be the same using numerical methods based on the same theory. Thus, we say that this “direct problem” of generating indentation loading–unloading curves from given material properties and indenter geometry is unique.

We now consider the reverse process of deducing the material properties, i.e., $E$, $v$, $Y$, and $n$, from a loading–unloading curve generated by a prescribed indenter $\theta$. The uniqueness of this “inverse problem” should not be taken for granted. Indeed, the non-uniqueness has been demonstrated with help from dimensional analysis either by examples [161] or by statistical methods [166].

Fig. 20a and b are some of the examples that illustrate the non-uniqueness of this inverse problem [161]. They were constructed using finite element calculations with input parameters selected based on the scaling functions for the loading curves (Eq. (21)), initial unloading slopes (Eq. (40)), final depth (Eq. (37)), or work of indentation (Eqs. (50) and (53)). These examples
demonstrated that for each indenter (i.e., fixed $\theta$) and given $n$ and $E$ values, there is a range of values of $Y$ and $n$ for which the calculated loading and unloading curves are virtually indistinguishable.

The sensitivity of the inverse process was further examined through a series of finite element calculation and statistical methods [166]. Different magnitudes of normally distributed noise were superimposed on a calculated force–displacement curve to simulate hypothetical data sets of loading–unloading curves for specific values of $E$, $Y$ and $n$. The sensitivity of the parameter confidence intervals to noise was determined using the $\chi^2$-curvature matrix, statistical Monte Carlo simulations, and a conjugate gradient algorithm that explicitly searches the global parameter space. All three approaches demonstrate that 1% noise levels preclude the accurate determination of $Y$ and $n$ based on a single force–displacement curve. An example is shown in Fig. 21. Therefore, loading and unloading curves from conical indentation alone cannot uniquely determine stress–strain relationships.

Since pyramidal indenters are also geometrically self-similar, scaling relationships such as those for conical indenters exist [53]. Furthermore, since previous numerical work has shown that the loading and unloading curves are the same for pyramidal and conical indentation, provided that the volume-to-depth relationships are the same for the two types of indenters [167,168]. Consequently, loading and unloading curves from pyramidal indentation alone also cannot uniquely determine stress–strain relationships. This non-uniqueness is not surprising, considering that many features of loading–unloading curves are related to each other, such as $W_d/W_{tot}$, $h_f/h_m$, and $H/E^*$, making the

Fig. 20. For a given indenter geometry (i.e., fixed $\theta$) there exist overlapping loading and unloading curves: (a) for highly elastic solids (i.e., large $Y/E$) and (b) for highly plastic solids (i.e., small $Y/E$) [161].
information content in a loading–unloading curve obtained from a single conical or pyramidal indentation redundant and, thus, limited.

Recognizing this limitation, a number of authors have proposed methods of inferring stress–strain relationships by making additional measurements. For example, O’Dowd and co-workers [169,170] developed a procedure to obtain the stress–strain relationship from indentation tests based on dimensional analysis and finite element calculations. Their method consists of extracting information from loading–unloading curves and from the surface profile measurements of the contact area after unloading. A similar procedure has been advanced by Mata and Alcala [171]. Matsuda [172] has proposed a method of predicting stress–strain curves of elastic–plastic materials by measuring Vickers hardness and the magnitude of piling-up and sinking-in. Futakawa and co-workers demonstrated a method for determining the constitutive equation of elastic–plastic materials by using plural indenters with different apex angles [173]. They suggested that the number of the indenters with different apex angles should equal the number of unknown constants in the constitutive model. Several variations of the multiple indenter approach using conical or pyramidal indenters have also appeared in the literature [174–176].

Although the non-uniqueness conclusion was reached for ideally sharp conical and pyramidal indenters, it should be applicable to those indenters with rounded tips, provided that the indentation depth is sufficiently large and the loading and unloading curves approach those obtained using the ideally sharp indenters. When the indentation depth is comparable to the radius of the rounded tip, however, indentation loading–unloading curves are approximately the same as that of spherical indentations. The question then becomes whether stress–strain curves can accurately be obtain by spherical indentation.

The key feature of spherical indentation or conical indentation with a rounded tip is the presence of a length scale, \( R \), the indenter radius. Dimensional analysis shows that the loading curve is given by

\[
F = Eh^2 \Pi \left( \frac{Y}{E}, v, n, \frac{h}{R} \right).
\]

Fig. 21. Three sets of force–displacement curves that differ by less than 0.5% can be found for widely differing values of \((E, Y, n)\) by utilizing a gradient search algorithm [166].
The presence of $h/R$ in the above equation breaks the $h^2$ dependence of loading curves. It offers the possibility of a large number of equations for extracting $E$, $Y$, $v$, and $n$:

$$\frac{F_i}{Eh_i^2} = \Pi \left( \frac{Y}{E}, v, n, \frac{h_i}{R} \right), \text{ for } i = 1, 2, 3, \ldots . \tag{55}$$

Presently, there are several proposed methods [177–192] for estimating stress–strain relationships from spherical indentation load–displacement curves using such optimization and statistical techniques as response surfaces, Kalman filters, and neural networks. Using these methods, a set of values of $E$, $Y$, $v$, and $n$ may be identified so that the calculated load–displacement curves match that of experimentally measured ones. However, since measurement errors in $F_i$ and $h_i$ can cause errors in the estimation of $E$, $Y$, $v$, and $n$, it is insufficient to report the best estimates of $E$, $Y$, $v$, and $n$ without reporting their errors. Detailed sensitivity analyses of the effects of various errors would thus be necessary. Such robust techniques for ascertaining stress–strain relationship as well as their confidence levels will likely be available in the future.

5. Indentation in power law creep solids using conical and pyramidal indenters

5.1. Background

Indentation measurements have long been applied to the study of strain-rate and temperature effects on deformation behavior of materials. Earlier experimental work was done primarily by measuring hardness as a function of time and temperature [193]. Later, instrumented indentation techniques, where load, displacement, and time could be controlled or monitored, were developed to determine the creep properties of bulk materials. An example is the impression creep test using a cylindrical flat punch for examining metals and polymers [194–199] (see [199] for a comprehensive review on impression creep tests). With the advent of nanoindentation, measurements of strain rate and temperature effects in thin films and nanostructured materials became possible [200–205]. These measurements are important for a number of industrial applications, including interconnects for electronic devices, materials for lead-free solder joints, and metals exhibiting superplasticity for low cost forming of complex parts, as well as for fundamental understanding of deformation mechanisms at small length scales.

Several types of indentation experiments have been proposed to gain insight into the strain-rate dependent properties of materials using self-similar indenters. Examples include the use of constant loading rate, $\dot{F}$; constant displacement rate, $\dot{h}$; or the keeping of parameters such as $\dot{h}/h$ or $\dot{F}/F$ constant. The “indentation strain rate” is usually defined as $\dot{h}/h$ [200–202]. The strain-rate dependence of a measured property, such as hardness, is then expressed in terms of $\dot{h}/h$. This definition allows quantitative comparison between conventional creep or relaxation tests and indentation creep tests. It has been found that the creep exponent obtained from these macro- and micro-tests are comparable for metals [200–202] and ceramics [204,205] where the power-law creep model applies. A number of theoretical and numerical studies of indentation in strain-rate dependent solids have also appeared recently in the literature [206–209]. In previous sections, dimensional analysis was applied to indentation in elastic–plastic solids without strain-rate effects. In this section, we apply dimensional analysis to conical and pyramidal indentation in materials dominated by strain-rate effects in the form of power-law creep [55]. We show that the parameter $\dot{h}/h$ can indeed be chosen to represent indentation strain rate. The scaling relationships for load–displacement curves
and hardness are obtained. These relationships are then applied to several types of indentation creep experiments in which either \( h, \dot{F} \), or \( \dot{F}/F \) is kept constant. The predictions are consistent with experiments reported in the literature.

5.2. Load–displacement curves and hardness

We consider a three-dimensional, rigid, self-similar indenter (\( \theta \)) indenting normally into a homogeneous solid with power-law creep (Eq. (11)). We assume that the friction coefficient at the contact surface between the indenter and the solid is zero. By assuming Eq. (11), the effects of elasticity, such as Young’s modulus and Poisson’s ratio, are precluded.

For an isotropic solid obeying the creep rule given in Eq. (11), the two dependent variables, \( F \) and \( A_c \), are functions, \( f \) and \( g \), of all the independent governing parameters, \( b, m, \) indenter displacement (\( h \)), rate of indenter displacement (\( \dot{h} \)), and indenter angles (\( \theta \)):

\[
F = f(b, m, h, \dot{h}, \theta),
\]

\[
A_c = g(b, m, h, \dot{h}, \theta).
\]

Eqs. (56) and (57) are implicitly dependent on time, \( t \), since \( h \) and \( \dot{h} \) are dependent on time and \( t = \int_{h(0)}^{h(t)} (dh/\dot{h}) \).

Among the five governing parameters, three of them, \( b, h \) and \( \dot{h} \), have independent dimensions. The dimensions of \( F \) and \( A_c \) are then given by \( \frac{F}{C_{138}} = \frac{b}{C_{138}} \frac{h}{C_{138}} \frac{\dot{h}}{C_{138}} m \frac{h}{C_{138}}^2 \) and \( \frac{A_c}{C_{138}} = \frac{h}{C_{138}}^2 \) respectively. Applying the \( \Pi \)-theorem in dimensional analysis, we obtain:

\[
F = b \left( \frac{\dot{h}}{h} \right)^m \Pi_\alpha(m, \theta),
\]

\[
A_c = h^2 \Pi_\beta(m, \theta),
\]

where \( \Pi_\alpha = F/b(\dot{h})^m h^{2-m} \) and \( \Pi_\beta = A_c/h^2 \), \( m \), and \( \theta \) are all dimensionless. Consequently, the hardness under load is

\[
H = \frac{F}{A_c} = b \left( \frac{\dot{h}}{h} \right)^m \frac{\Pi_\alpha}{\Pi_\beta} \equiv (b \Pi_\gamma) \left( \frac{\dot{h}}{h} \right)^m,
\]

where \( \Pi_\gamma \equiv \Pi_\alpha/\Pi_\beta \). To simplify notation, \( \Pi_i \equiv \Pi_i(m, \theta) \) for \( i = \alpha, \beta, \gamma \) in the following.

This equation shows that the strain-rate dependence of hardness is contained in the parameter, \( \dot{h}/h \). Comparing with Eq. (11), we observe that, aside from the pre-factor, the power-law dependence of hardness, \( H \), on \( \dot{h}/h \) in indentation experiments is the same as that of stress, \( \sigma \), on strain-rate, \( \dot{\varepsilon} \), in uniaxial creep tests. Thus, the parameter, \( \dot{h}/h \), can indeed be chosen, aside from a time-independent pre-factor, to represent indentation strain-rate.

When the force, instead of displacement, is the independent variable, Eq. (58) may be integrated to obtain:

\[
h(t) = \left( \frac{2}{m} \right)^{m/2} (b \Pi_\alpha)^{-(1/2)} \left[ \int_0^t F^{1/m}(t) \, dt \right]^{m/2},
\]

with initial condition \( h(0) = 0 \).
An equation similar to Eq. (58) was first proposed by Grau et al. [205] based on the assumptions that \( H = b' \dot{\varepsilon}^m \) and \( \dot{\varepsilon} = \frac{\dot{h}}{h} \). The present derivation is based on the self-similarity that exists in the problem of indentation in power-law creep solids using self-similar indenters. The assumptions in the derivation of Grau et al. are shown to be the consequence of this self-similarity. Furthermore, the parameter \( b \) is not, in general, the same as \( b' \) suggested by Grau et al. [205].

In the following, the above equations are applied to several types of indentation conditions in which either \( \dot{h} \), \( F \), or \( F/F \) is kept constant.

5.2.1. Constant displacement rate, \( \dot{h}_c \), condition

When \( \dot{h}_c \) is constant, the force and hardness are, according to Eqs. (58) and (60),

\[
F = (b \Pi_s) \left( \frac{\dot{h}}{h} \right)^m h^2, \\
H = (b \Pi_l) \left( \frac{\dot{h}_c}{h} \right)^m.
\]

These equations show that during loading the force is proportional to \( h^{2-m} \) and is no longer proportional to \( h^2 \). The square dependence is characteristic of indentation using self-similar indenters in elastic–plastic solids without strain-rate dependence as we have seen in Section 4.1. The hardness decreases with indentation depth. The creep exponent, \( m \), can be obtained from either the indentation loading curve or from a graph of \( \ln(H) \) versus \( \ln(\dot{h}/h) \). The expression for hardness (Eq. (63)), scaled by its value at the maximum indenter displacement \( h_m \), is illustrated in Fig. 22.

5.2.2. Constant loading rate, \( \dot{F}_c \), condition

Substituting \( F = \dot{F}_c t \) in Eq. (61), we obtain

\[
h(t) = \left( \frac{\dot{F}_c}{b \Pi_s} \right)^{1/2} \left( \frac{2}{m + 1} \right)^{m/2} t^{(m+1)/2}.
\]

Fig. 22. Indentation in power-law creep solids with constant-displacement-rate: hardness decreases with indentation depth [55].
Consequently, the force, using Eq. (58), is

\[ F = (b \Pi_x)^{1/(m+1)} \left[ \frac{\dot{h}}{\dot{F}} \frac{m+1}{2} \right]^{m/(m+1)} H^{2/(m+1)}. \]  

(65)

The indentation strain-rate can be expressed, using Eq. (64), as

\[ \frac{\dot{h}}{h} = \frac{m+1}{2} \frac{\dot{F}_c}{F}, \]  

(66)

Using Eq. (60), the hardness becomes

\[ H = (b \Pi_x) \left( \frac{\dot{h}}{h} \right)^m = (b \Pi_x) \left( \frac{m+1}{2} \right)^m \left( \frac{\dot{F}_c}{F} \right)^m. \]  

(67)

These equations show that during loading the force is proportional to \( h^{2/(m+1)} \) and is no longer proportional to \( h^2 \). The hardness decreases with increasing indentation load. The creep exponent, \( m \), can be obtained from either the indentation loading curve, a graph of ln(H) versus ln(\( \dot{h}/h \)), or a graph of ln(H) versus ln(\( \dot{F}/F \)). The expression for hardness (Eq. (67)), scaled by its value at the maximum indenter displacement \( h_m \), is illustrated in Fig. 23.

5.2.3. Constant loading rate over load, \( \dot{F}/F \), condition

Since \( \dot{F}/F = \lambda \) is a constant, the force is given by \( F = F_0 e^{\lambda t} \), where \( F_0 \) is the force at \( t = 0 \). Substituting into Eq. (61), we obtain a solution:

\[ h(t) = \frac{1}{\sqrt{b \Pi_x}} \left( \frac{2}{\lambda} \right)^{m/2} F_0^{1/2} (e^{\lambda t/m} - 1)^{m/2}, \]  

(68)

and for large \( t > m/\lambda \),

\[ h(t) \approx \left( \frac{2^m F_0}{\lambda^m b \Pi_x} \right)^{1/2} e^{\lambda t/2}. \]  

(69)

Fig. 23. Indentation in power-law creep solids with constant-loading-rate: hardness decreases with indentation depth [55].
Consequently, the indentation strain-rate is given by

\[ \frac{\dot{h}}{h} = \frac{\dot{\lambda}}{2} (1 - e^{-\dot{\lambda}/m})^{-1} \approx \frac{\dot{\lambda}}{2} = \frac{1}{2} \frac{\dot{F}}{F}. \] (70)

Thus, the indentation strain-rate, \( \dot{h}/h \), is half of \( \dot{F}/F \) after a transient period on the order of \( ml/\dot{\lambda} \).

Using Eqs. (58) and (60), the respective indentation loading curve and hardness may be written,

\[ F = F_0 e^{\dot{\lambda} t} \approx (b \Pi_x) \left( \frac{\dot{\lambda}}{2} \right)^m h^2, \] (71)

\[ H = (b \Pi_x) \left( \frac{\dot{h}}{h} \right)^m \approx (b \Pi_x) \left( \frac{\dot{\lambda}}{2} \right)^m. \] (72)

Eq. (72), scaled by the hardness value at \( h_m \), is shown in Fig. 24. Clearly, hardness reaches a steady state value when \( \dot{F}/F \) is kept constant. Correspondingly, the loading force is again proportional to \( h^2 \). The hardness increases with \( (\dot{F}/F)^m \). The creep exponent, \( m \), can be obtained from either a graph of \( \ln(H) \) versus \( \ln(\dot{h}/h) \), or a graph of \( \ln(H) \) versus \( \ln(\dot{F}/F) \).

The results of the above-analysis are consistent with experiments reported in the literature. For example, numerous authors have shown a linear dependence between \( \ln(H) \) and \( \ln(\dot{h}/h) \) for all three loading conditions considered above (i.e., either \( \dot{h}, \dot{F} \), or \( \dot{F}/F \) is kept constant). Furthermore, the creep exponent, \( m \), has been obtained from the slope of the straight lines in the graph of \( \ln(H) \) versus \( \ln(\dot{h}/h) \). The creep exponent, \( m \), has also been obtained from indentation loading curves by Grau et al. [205] using either constant \( \dot{h} \) or \( \dot{F} \) experiments and equations similar to Eqs. (62) and (65). Recently, Ma et al. [210,211] performed a series of Berkovich indentation experiments with constant loading rates on solder alloys that have known power-law creep behavior and demonstrated the applicability of Eqs. (65) and (67). The constant loading rate experiments have also been performed to show that indentation creep was relatively insignificant in the CNx fullerene-like coatings by demonstrating that the indentation loading curves are proportional to the square of indenter displacement over a range of indentation depths [212].

![Fig. 24. Indentation in power-law creep solids with constant-loading-rate-over-load: hardness reaches a steady-state value which is depth independent [55].](image-url)
Several authors have reported an “indentation size effect” in constant $h$ or $F$ experiments as predicted by the above equations (Eqs. (63) and (67)): hardness decreases with increasing indentation depth or load [203,213]. It has also been demonstrated recently by Lucas and Oliver [203] that, in constant $F/F$ experiments, the indentation strain-rate reaches a “steady state” and is given by $0.5F/F$ (Fig. 25). This observation agrees with Eq. (70). Furthermore, these authors showed that the steady state hardness is independent of indentation depth and is proportional to $(F/F)^m$ (Fig. 26). These observations are also in agreement with Eq. (72). Thus, a connection is established between indentation in power-law creep solids and indentation size effects. It is also evident that conical or pyramidal indentation experiments under either constant $F/F$ or $h/h$ condition may be used to distinguish power-law creep effects from other sources of indentation size effects.

Fig. 25. A plot of indentation strain rate vs. displacement for five different constant $F/F$ experiments conducted on indium showing constant indentation strain rates, $h/h$ equal to $0.5 F/F$ [203].

Fig. 26. A plot of hardness vs. displacement for five different constant $F/F$ experiments conducted on indium showing indentation stead-state behavior in response to constant $F/F$ load ramp [203].
Hence, dimensional analysis can be very helpful in gaining a basic understanding and much insights into indentation in power-law creep solids using self-similar indenters.

6. Indentation in linear viscoelastic solids using conical and pyramidal indenters

6.1. Background

Instrumented indentation is playing an increasing role in the study of small-scale mechanical behavior of “soft matters”, such as polymers, composites, biomaterials, and food products. Many of these materials exhibit viscoelastic behavior, especially at elevated temperatures. Modeling of indentation into viscoelastic solids thus forms the basis for analyzing indentation experiments in these materials. In fact, theoretical studies of contacting linear viscoelastic bodies can be traced to the mid 1950s. Lee [214], Radok [215], and Lee and Radok [216] have analyzed several viscoelastic contact problems using the elastic-viscoelastic correspondence principle. Their approach was to replace the materials elastic constants in the elastic solutions of contact problems by the corresponding differential operators in the viscoelastic constitutive equation. Their “tentative” solutions obtained using this approach satisfied all the necessary boundary conditions under certain loading conditions. More general solutions were obtained by Hunter [217], Graham [218,219], Yang [220] and Ting [221,222] using techniques such as Laplace transform, Riemann–Stieltjes integrals, and convolution integrals. These early studies produced load–displacement equations for rigid spherical or conical indenters indenting viscoelastic solids. In recent years, Cheng et al. [223] analyzed flat-punch indentation of viscoelastic materials. Larsson and Carlsson [224] applied finite element analysis to both flat circular punch and spherical indentation in linear viscoelastic materials. There is also an effort to simplify the derivations using more intuitive approaches [225–228]. Examples include the approach by Sakai and co-workers [225–227] using the concept of “representative strain” and “representative stress”. In this section, we provide a concise derivation of the load–displacement relationship for conical indentation in linear viscoelastic solids using a scaling approach. This new derivation does not invoke sophisticated mathematical tools nor rely on concepts of representative strain or stress. We will then discuss several typical loading procedures for measuring linear viscoelastic properties and provide examples of loading–unloading curves obtained from finite element calculations.

6.2. A new derivation of the load–displacement relationship

We consider a rigid, smooth conical indenter with half-angle $\theta$ indenting the viscoelastic solid that can be described by the constitutive relationship given in Eq. (13). We assume that the displacement field is self-similar and, following dimensional analysis,

$$u_i = h\Pi_l\left(\frac{x_i}{h}, \frac{\dot{\lambda}}{G}, \theta\right),$$

where $h$ is the indenter displacement and $x_i, l = 1, 2, 3$ are the spatial coordinates. Furthermore, we assume time independent Poisson’s ratio, $\nu$, and note that $\omega \equiv \dot{\lambda}(t)/G(t) = 2\nu/(1 - 2\nu)$ is also time independent. Consequently,

$$u_i = h\Pi_l\left(\frac{x_i}{h}, \nu, \theta\right).$$
Now the time, \( t \), appears implicitly only through \( h(t) \) in Eq. (74).

Using the definition of strain for small deformation and Eq. (74), we have

\[
\varepsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \frac{1}{2} \left( \frac{\partial \Pi_i}{\partial \xi_j} + \frac{\partial \Pi_j}{\partial \xi_i} \right),
\]

(75)

where \( \xi_i = (x_i/h) \). Substituting Eq. (75) in Eq. (13), we obtain,

\[
h \frac{\partial \sigma_{ij}}{\partial x_j} = - \int_0^t G(t-\tau) \frac{\bar{h}}{h} \left( \xi_i \frac{\partial}{\partial \xi_i} + 1 \right) \frac{\partial}{\partial \xi_j} [2\varepsilon_{ij} + \omega \delta_{ij} \varepsilon_{kk}] d\tau.
\]

(76)

Thus, the equilibrium condition, \( \frac{\partial u_i}{\partial x_j} = 0 \), is satisfied, if

\[
\frac{\partial}{\partial \xi_j} [2\varepsilon_{ij} + \omega \delta_{ij} \varepsilon_{kk}] = 0.
\]

(77)

Now, Eq. (77) is simply the equilibrium condition for a linear elastic indentation problem with the ratio of Lamé's constants equal \( \lambda/G = \omega \). Thus, if \( \varepsilon_{ij} \), a time independent strain field, is the solution of the static linear elastic indentation problem, then the stress field given below represents the solution for the stress field of the indentation of a linear viscoelastic material,

\[
\sigma_{ij} = - \int_0^t G(t-\tau) \frac{\bar{h}(\tau)}{h(\tau)} \frac{\partial}{\partial \xi_i} [2\varepsilon_{ij} + \omega \delta_{ij} \varepsilon_{kk}] d\tau.
\]

(78)

In particular the normal stress on the surface of the specimen, i.e., \( \xi_3 = 0 \), is given by

\[
\sigma_{33} = - \int_0^t G(t-\tau) \frac{\bar{h}(\tau)}{h(\tau)} \frac{\partial}{\partial \xi_3} [2\varepsilon_{33} + \omega \varepsilon_{kk}] d\tau, \quad \alpha = 1, 2.
\]

(79)

We note that the stress field is axially symmetric and that the surface stress \( \sigma_{33} \) vanishes beyond the contact area with radius \( h_c \tan \theta \) which is proportional to \( h \tan \theta \). Hence we may integrate \( \sigma_{33} \) over the entire surface, \( A \), of \( \xi_3 = 0 \) to obtain the indentation load \( F(t) \),

\[
F(t) = - \int \int_A \sigma_{33} \, dx_1 \, dx_2 = \int_0^t G(t-\tau) \bar{h}(\tau) \left\{ \int \int_A \xi_3 \frac{\partial}{\partial \xi_3} \frac{2\varepsilon_{33} + \omega \varepsilon_{kk}}{\partial \xi_3} \, d\xi_1 \, d\xi_2 \right\} d\tau.
\]

(80)

The surface integral is obviously a function of \( v \) and \( \theta \), and will be denoted by \( K(v, \theta) \),

\[
F(t) = K(v, \theta) \int_0^t G(t-\tau) \bar{h}(\tau) h(\tau) \, d\tau = \frac{1}{2} K(v, \theta) \int_0^t G(t-\tau) \, dh^2(\tau).
\]

(81)

In the limit of infinite relaxation time the above result reduces to the linear elastic solution that is well known in the literature [87], namely

\[
F = \frac{4G \tan \theta}{\pi(1-v)} h^2.
\]

(82)

We find \( K(v, \theta) \) by comparing Eqs. (81) and (82) and arrive at,

\[
F(t) = \frac{4 \tan \theta}{\pi(1-v)} \int_0^t G(t-\tau) \, dh^2(\tau),
\]

(83)
or equivalently,
\[
F(t) = 8 \tan \theta \frac{\pi(1 - v)}{\pi(1 - v)} \int_{0}^{t} G(t - \tau) h(\tau) \frac{dh(\tau)}{d\tau} d\tau = - \frac{8 \tan \theta}{\pi(1 - v)} \int_{0}^{t} G(\tau) h(t - \tau) \frac{dh(t - \tau)}{d\tau} d\tau. \tag{83t}
\]

The load–displacement relationship can therefore be obtained if the viscoelastic properties of materials, \(G(t)\) and \(v\), are known using the above equations. Conversely, the viscoelastic properties may be obtained from measured \(F(t)\) versus \(h(t)\) relations by solving an integral equation.

Eq. (83) is a special case of a more general expression derived first by Graham [218] and Ting [221] using more sophisticated techniques. They showed that Eq. (83) is valid for loading where the contact area is a monotonically increasing function of time [218–222]. Under the same condition of monotonically increasing contact area, the ratio of contact depth to indenter displacement is the same as that the purely elastic case, namely,
\[
\frac{h}{h} = \frac{2}{\pi}. \tag{84}
\]

The equations for unloading where the contact area decreases monotonically to zero have also been derived [218,221]. The agreement between Eq. (83) and the earlier results of Graham and Ting suggests that the assumption of a self-similar displacement field (Eq. (73)) is appropriate. An alternative derivation of Eq. (83) has recently been put forward by Saikai and co-workers [225–227]. Similar expressions have been derived for indentation in linear viscoelastic solids using spherical and more complex shapes of indenters [218–222]. These equations have been applied to obtain the loading–unloading curves for spherical indentation in linear viscoelastic solids with a constant Poisson’s ratio under specific loading conditions [221,229].

6.3. Obtaining viscoelastic properties from loading curves

In this section, we discuss the methods of determining linear viscoelastic properties from loading curves using conical or pyramidal indenters under various loading conditions. Under a quasi static constant indenter displacement rate condition, where \(h(t) = v_{0}t\) for \(t \geq 0\) and \(h(t) = 0\) for \(t < 0\), the load–displacement relationship becomes,
\[
F(t) = 8v_{0}^{2} \tan \theta \frac{\pi(1 - v)}{\pi(1 - v)} \int_{0}^{t} G(t - \tau) \tau d\tau = \frac{8v_{0}^{2} \tan \theta}{\pi(1 - v)} \int_{0}^{t} G(\tau)(t - \tau) d\tau. \tag{85}
\]

Taking the second derivative of \(F(t)\) with respect to \(t\), we obtain
\[
\frac{d^{2}F(t)}{dt^{2}} = \frac{8v_{0}^{2} \tan \theta}{\pi(1 - v)} G(t). \tag{86}
\]

This method of obtaining \(G(t)/(1 - v)\), requiring knowledge of the second order derivative of the load from conical or pyramidal indentation, was proposed and experimentally tested recently by Shimizu et al. [225].

Under a quasi static constant indenter displacement condition, where the displacement is described by a Heaviside step function, \(h(t) = h_{0}H(t)\), or equivalently,
\[
\begin{align*}
\text{for } t \geq 0 & \quad h(t) = h_{0}, \\
\text{for } t < 0 & \quad h(t) = 0.
\end{align*}
\]
Eq. (83) becomes, using the fact that $\frac{dh}{dt} = h_0 \delta(t)$, where $\delta(t)$ the Dirac delta function,

$$F(t) = \frac{8 \tan \theta}{\pi(1 - v)} \int_0^t G(t - \tau) h(\tau) h_0 \delta(\tau) \, d\tau = \frac{4h_0^2 \tan \theta}{\pi(1 - v)} G(t).$$

(87)

Therefore, by measuring the load as a function of time in the quasi static constant indenter displacement experiment the viscoelastic property, $G(t)/(1 - v)$, can be obtained directly from $F(t)$.

In practice, this constant indentation displacement condition may be realized in several ways, for example, by ramping the displacement from 0 to a finite value using a constant displacement rate, $h(t) = v_0 t$, for $0 \leq t \leq \delta t$, and then enforcing the constant indentation displacement condition, $h(t) = v_0 \delta t \equiv h(0^+)$,

$$F(t) = \frac{8 \tan \theta}{\pi(1 - v)} \int_0^t G(t - \tau) h(\tau) \frac{dh(\tau)}{d\tau} \, d\tau = \frac{8 \tan \theta}{\pi(1 - v)} \int_0^{\delta t} G(t - \tau) h(\tau) \frac{dh(\tau)}{d\tau} \, d\tau$$

$$+ \frac{8 \tan \theta}{\pi(1 - v)} \int_0^{\delta t} G(t - \tau) h(\tau) \frac{dh(\tau)}{d\tau} \, d\tau = \frac{8 \tan \theta}{\pi(1 - v)} \int_0^{\delta t} G(t - \tau) \frac{dh(\tau)}{d\tau} \, d\tau$$

$$= \frac{4h_0^2(0^+) \tan \theta}{\pi(1 - v)} G(t).$$

(88)

Eqs. (87) and (88) are of the same form and consequently $G(t)/(1 - v)$ can be obtained from knowing $F(t)$ and $h(0^+)$. In the previous section on indentation in power-law creep solids, we have seen that indentation strain-rate could indeed be characterized by $(dh/dt)/h$. Is the concept of indentation strain rate, defined as $(dh/dt)/h$, useful to indentation in linear viscoelastic solids using conical or pyramidal indenters? Is there advantage of performing indentation experiments with constant $(dh/dt)/h$? These questions are discussed as follows. We consider constant strain rate,

$$\frac{dh/dt}{h} = \frac{1}{\tau_0},$$

where $\tau_0$ is a constant. This condition is satisfied when $h(t) = h(0) \exp(t/\tau_0)$, where $h(0)$ is a constant.

In practice, this constant indentation strain rate condition may also be realized by ramping the displacement from 0 to a finite value using constant displacement rate, $h(t) = v_0 t$, for $0 \leq t \leq \delta t$, and then enforcing the constant indentation strain rate condition:

$$F(t) = \frac{8 \tan \theta}{\pi(1 - v)} \int_0^t G(t - \tau) h(\tau) \frac{dh(\tau)}{d\tau} \, d\tau = \frac{8 \tan \theta}{\pi(1 - v)} \int_0^{\delta t} G(t - \tau \tau) \, d\tau$$

$$+ \frac{8 \tan \theta}{\pi(1 - v) \tau_0} \int_0^{\delta t} G(t - \tau) h(\tau) \frac{h(\tau)}{d\tau} \, d\tau = \frac{4h_0^2(0^+) \tan \theta}{\pi(1 - v)} G(t)$$

$$+ \frac{8 \tan \theta}{\pi(1 - v) \tau_0} \int_0^{\delta t} G(t) h^2(t - \tau) \, d\tau,$$

(89)
where \( h(0^+) = v_0 \delta t \), as \( \delta t \to 0 \). Taking the first derivative of \( F(t) \) with respect to \( t \) and noticing
\[
1/\tau_0 = d h(t - \tau) / d \tau = -d h(t - \tau) / d \tau,
\]
we obtain
\[
\frac{dF(t)}{dt} = \frac{4h^2(0^+) \tan \theta}{\pi(1 - v)} \frac{dG(t)}{dt} + \frac{8 \tan \theta}{\pi(1 - v)} G(t) h^2(0^+) + \frac{8 \tan \theta}{\pi(1 - v)} \frac{2}{\tau_0} \int_0^t G(\tau) h(t - \tau) \frac{dh(t - \tau)}{d \tau} d \tau
\]
\[
= \frac{4h^2(0^+) \tan \theta}{\pi(1 - v)} \frac{dG(t)}{dt} + \frac{2}{\tau_0} G(t) + \frac{2}{\tau_0} F(t).
\]  
(90)

Solving this differential equation for \( G(t) \)
\[
\frac{dG(t)}{dt} + \frac{2}{\tau_0} G(t) = -\frac{\pi(1 - v)}{4h^2(0^+) \tan \theta} \left( \frac{dF(t)}{dt} - \frac{2}{\tau_0} F(t) \right),
\]  
(90)'

with the initial condition that the relation between \( G(0) \) and \( F(0) \) must satisfy the elastic relation, we obtain:
\[
G(t) = \frac{\pi(1 - v)}{4h^2(0^+) \tan \theta} \left( F(t) - \frac{4}{\tau_0} \int_0^t F(\tau) \exp \left[ \frac{-2(t - \tau)}{\tau_0} \right] d \tau \right).
\]  
(91)

Thus, the viscoelastic property, \( G(t)/(1 - v) \), can be obtained since \( \tau_0, h(t), F(t), \) and \( \int_0^t F(\tau) \exp \left[ -2(t - \tau)/\tau_0 \right] d \tau \) are either known or can be measured.

More general loading procedure may be used to determine the mechanical behavior of viscoelastic solids, \( G(t) \), from indentation. Applying Laplace transform to Eq. (81), we obtain,
\[
F^*(s) = K(v, \theta) G^*(s) W^*(s),
\]  
(92)

where \( W^*(s) \) is the Laplace transform of \( h(\tau) h(\tau) \). Let \( V(t) \) denote the inverse Laplace transform of \( 1/ W^*(s) \), then
\[
G(t) = \frac{1}{K(v, \theta)} \int_0^t F(t - \tau) V(\tau) d \tau.
\]  
(93)

6.4. Examples: “standard” model for linear viscoelastic solids

6.4.1. Loading curves for conical indentation in “standard” model solids with a constant Poisson’s ratio

As an illustration, we apply the above analysis to a three-parameter “standard” model for linear viscoelastic solids (Fig. 27) [82–84], assuming a constant Poisson’s ratio, \( v \). The relaxation modulus in shear for the model is given by [84],
\[
G(t) = \frac{G_1 G_2}{G_1 + G_2} + \frac{G_1^2}{G_1 + G_2} \exp \left( -\frac{t}{\tau_s} \right) = \begin{cases} G_1, & \text{for } t \ll \tau_s, \\ \frac{G_1 G_2}{G_1 + G_2}, & \text{for } t \gg \tau_s, \end{cases}
\]  
(94)
where $\tau_s = \eta(G_1 + G_2)$. A plot of $G(t)/G_1$ versus $t/\tau_s$ is shown in Fig. 28 for selected values of $G_1/G_2$. We note that $G(0) = G_1$ and $G(\infty) = G_1G_2/(G_1 + G_2)$ for $t = 0$ and $t = \infty$, respectively.

Under a quasi static constant indenter displacement rate condition, where $h(t) = v_0t$ for $t \geq 0$ and $h(t) = 0$ for $t < 0$, the load–displacement relationship becomes

$$F(t) = \frac{4\tan \theta}{\pi(1-v)}(v_0\tau_s)^2 \left[ \frac{G_1G_2}{G_1 + G_2} \left( \frac{t}{\tau_s} \right)^2 + 2 \frac{G_1^2}{G_1 + G_2} \left( \frac{t - \tau_s}{\tau_s} + \exp \left( - \frac{t}{\tau_s} \right) \right) \right]$$

$$= \begin{cases} 
\frac{4G_1\tan \theta}{\pi(1-v)}h^2(t), & \text{for } t \leq \tau_s \\
\frac{4\tan \theta}{\pi(1-v)}G_1G_2h^2(t), & \text{for } t \geq \tau_s.
\end{cases} \tag{95}$$

The force is, therefore, proportional to the square of indenter displacement in both the short and long time approximation, where the model solid is purely elastic with shear modulus given by $G_1$ and $G_1G_2/(G_1 + G_2)$, respectively. The relaxation time, $\tau_s$, characterizes the transition from these two purely elastic states. All three parameters, $G_1$, $G_2$, and $\eta$ can, thus, be measured from a single loading curve obtained with an arbitrary $v_0$ by either fitting the loading curve to Eq. (95) or by taking the second derivative of the loading curve with respect to time and calculate $G(t)$ using Eq. (86).
Under constant indentation strain rate condition, Eq. (89) becomes

\[
F(t) = \frac{4h^2(0^+) tan \theta}{\pi(1 - v)} G(t) + \frac{8 tan \theta}{\pi(1 - v)} \frac{h^2(t)}{\tau_0} \int_0^t G(\tau) \exp\left(\frac{-2\tau}{\tau_0}\right) d\tau = \frac{4h^2(0^+) tan \theta}{\pi(1 - v)} G(t)
\]

\[
+ \frac{4 tan \theta}{\pi(1 - v)} h^2(t) \left[ \frac{G_1 G_2}{G_1 + G_2} (1 - e^{-2/\tau_0}) + \frac{1}{1 + \tau_0/(2\tau_s) G_1 + G_2} (1 - e^{-2(1/\tau_0 + 1/(2\tau_s))t}) \right] 
\]

\[
= \begin{cases} 
\frac{4h^2(0^+) tan \theta}{\pi(1 - v)} G_1 + \frac{8 G_1 tan \theta}{\pi(1 - v)} h^2(t) \left( \frac{t}{\tau_0} \right), & \text{for } t \ll \min(\tau_0, 2\tau_s) \\
\frac{4h^2(0^+) tan \theta}{\pi(1 - v)} G_1 G_2 \left( \frac{G_1 G_2}{G_1 + G_2} + \frac{1}{1 + \tau_0/(2\tau_s) G_1 + G_2} \right), & \text{for } t \gg \max(\tau_0, 2\tau_s) 
\end{cases}
\] (96)

When \( t \ll \min(\tau_0, 2\tau_s) \), the force is not proportion to the square of indenter displacement even though the response is elastic (i.e., only \( G_1 \) appears in the equation above). However, the force \( F(t) \) is proportional to the square of indenter displacement when \( t \gg \max(\tau_0, 2\tau_s) \) since \( h(0^+) \) is much less than \( h(t) \). The coefficient, \( G_1 G_2/(G_1 + G_2) + (1/(1 + \tau_0/(2\tau_s) G_1 + G_2)) \), is a function of both viscoelastic properties (i.e., \( G_1 \), \( G_2 \), and \( \tau_s \)) and indentation strain rate \( (1/\tau_0) \). For very small indentation strain rate (i.e., \( \tau_0 \gg \tau_s \)), \( F(t) \approx (4/\pi(1 - v))(G_1 G_2 tan \theta/(G_1 + G_2)) h^2(t) \). Thus, the response is dominated by the long time behavior of the viscoelastic solid. For very large indentation strain rate, i.e., \( 2\tau_s > \tau_0 \), \( F(t) \approx (4G_1 tan \theta/(\pi(1 - v))) h^2(t) \). This implies that for very high indentation strain rate the mechanical response is dominated by \( G_1 \).

Eq. (96) shows that there are two time constants, \( \tau_s \) and \( \tau_0 \), in the problem of indentation in a three parameter standard linear viscoelastic model solid with constant indentation strain rate. These two independent time constants can both influence indentation loading curves. Unlike the case of indentation in power-law creep solids, the constant indentation strain-rate method does not seem to offer much advantage over other loading-conditions discussed in this section for obtaining viscoelastic properties from indentation loading curves.

### 6.4.2. Finite element calculations of loading–unloading curves for conical indentation in a “standard” model solid

In this section, we provide numerical examples obtained from finite element calculations of conical indentation in a linear viscoelastic solid. The main purpose is to illustrate the features of loading–unloading curves under various loading conditions, such as constant displacement rate, constant loading rate, and constant indentation strain rate. We will also discuss whether the initial unloading slope equation, Eq. (43), is applicable to indentation in viscoelastic solids. The numerical examples also help demonstrate the importance of specifying experimental conditions, such as loading and unloading conditions and the duration of experimental observations. Although analytical equations exist for calculating the complete loading and unloading curves for special cases such as constant Poisson’s ratio, commercial finite element analysis software packages are now available for solving problems of indentation in viscoelastic solids, including time-dependent Poisson’s ratio and finite friction at the interface between indenter and viscoelastic solid.

We consider a frictionless, rigid conical indenter with half angle \( \theta = 70.3^\circ \) indenting an isotropic linear viscoelastic solid. We assume a three-parameter shear relaxation modulus of the form given by Eq. (94) with \( G_1 = 234.60 \text{ MPa}, G_2 = 25.78 \text{ MPa}, \tau_s = \eta/(G_1 + G_2) = 0.99 \text{ s}, \) and a time-independent bulk modulus, \( K = 687.62 \text{ MPa}. \) Consequently, Young’s modulus and Poisson’s ratio
are all time-dependent and are given by $E(t) = 9KG(t)/(3K + G(t))$ and $\nu(t) = E(t)/(2G(t) - 1)$, respectively. Specifically, their values at $t = 0$ and $t = \infty$ are as follows: $G(0) = 234.60$ MPa and $G(\infty) = 23.23$ MPa; $E(0) = 631.94$ MPa and $E(\infty) = 68.90$ MPa; $\nu(0) = 0.3468$ and $\nu(\infty) = 0.4833$. The parameters of this fictitious model solid are used for illustration purposes of the effects of various indentation loading conditions. Because of linearity, the results may be scaled to represent other materials of the same general type when the dimensionless parameters, such as $G_1/G_2$, $G_1/K$, and $t/\tau_s$, are equal. Finite element calculations were carried out using the isotropic linear viscoelasticity model implemented in ABAQUS [90]. The model assumes that shear and volumetric behaviors are independent in multiaxial stress states and the time dependent stress states are described by Eq. (12) with independent relaxation moduli for deviatoric and volumetric behaviors. The finite element mesh is the same as that used in [158].

A set of calculations was carried out for symmetric triangular displacement functions shown in Fig. 29a where the final indentation depth was the same and the time duration, $\tau$, for the loading and unloading segments ranged from less than to greater than $\tau_s$ (i.e., $\tau/\tau_s$ is 0.505, 1.010, 5.051, 10.10, and 50.51). The calculated loading and unloading curves are shown in Fig. 29b together with two load–displacement curves for indentation in purely elastic solids calculated using Eq. (82) for $G(0)/(1 - \nu(0))$ and $G(\infty)/(1 - \nu(\infty))$, respectively. The two purely elastic cases represent the viscoelastic response under infinitely fast and infinitely slow indentation conditions where the material behaves purely elastically with the corresponding sets of elastic constants. The viscoelastic indentation response is bounded by the two elastic cases. At a given indentation depth, the greater
the indenter displacement rate, \( v_0 \), the larger the force, \( F \), which is consistent with Eq. (95) although this equation cannot be used to quantitatively describe the loading curves because of the time-dependent Poisson’s ratio. The ratio of irreversible work to total work (i.e., the ratio of the area between the loading–unloading curve and the area under the loading curve) increases and then decreases with increasing \( v_0 \). This can be understood because the indentation response is predominately elastic when \( v_0 \) is either very large or very small.

A set of calculations was also carried out for symmetric triangular force functions shown in Fig. 30a where the final indentation load was the same and the time duration, \( \tau \), for the loading and unloading segments ranged from less than to greater than \( \tau_s \) (i.e., \( \tau/\tau_s \) is 0.505, 1.010, 5.051, 10.10, and 50.51). The calculated loading and unloading curves are shown in Fig. 30b together with two load–displacement curves for indentation in purely elastic solids representing infinitely fast (i.e., \( G(0)/(1 - v(0)) \)) and infinitely slow (i.e., \( G(\infty)/(1 - v(\infty)) \)) indentation responses. For a given load during loading, the indentation depth increases with decreasing loading rate. When unloading is sufficiently slow, the indentation depth continues to increase after the force reaches the maximum, resulting in a “bulge” or “nose” in the unloading curve, which has also been observed experimentally [230–233]. This bulge is the consequence of the continuing forward movement of the indenter during the early stage of unloading. Because of the forward movement of the indenter, the maximum contact area occurs after the force maximum. This delayed maximum contact area.
behavior (shown in Fig. 30a) was predicted by the analytical theories of Graham [218] and Ting [221,222]. In contrast, there are no bulges on the unloading curves under displacement control conditions (see, for example, Fig. 29b) because the indenter is retracted and the force on the indenter cannot increase during displacement-controlled unloading. However, finite element results suggest that the delayed maximum contact area behavior, though much less pronounced, may occur even under constant displacement rate unloading conditions.

We now consider indentation under the condition of constant indentation strain rates. It has been suggested that the constant indentation strain rate condition can be realized by performing load-controlled indentations with constant values of \((dF/dt)/F\) for loading [203,230]. The reduced elastic modulus and hardness values are often calculated from the initial unloading slope using the Oliver and Pharr procedure. Very little attention, however, has been given to the unloading rate. To illustrate the effects of unloading rates, we have calculated a set of loading–unloading curves with loading–unloading history shown in Fig. 31a: exponential loading, \(F(t) = 0.001 \mu N \exp(t/1.0 \text{ s})\) (i.e., \((dF/dt)/F = 1.0 \text{ s}^{-1}\)), and linear unloading from the same maximum indentation load within a time period \(\tau = 0.01, 0.1, 1.0, \text{ and } 10 \text{ s}\), corresponding to \(\tau/\tau_s = 0.0101, 0.101, 1.010, \text{ and } 10.10\), respectively. The loading–unloading curves shown in Fig. 31b demonstrate that the initial unloading slope and the shape of unloading curve are functions of the unloading condition when the loading condition is kept the same. The finite element results show that unloading curves converge to a limiting case for very fast unloading (i.e., for \(\tau/\tau_s < 0.1\)) with the initial unloading slope given by Eq. (43), namely,

\[
\frac{dF}{dh} = \frac{4G(0)}{1 - v(0)}a,
\]

Fig. 31. Exponential loading and linear-unloading from the same maximum indentation load within a time period \(\tau = 0.01, 0.1, 1.0, \text{ and } 10 \text{ s}\) (a) and calculated loading–unloading curves (b).
where \( a = h_c \tan(\theta) \) is the contact radius. The numerical results demonstrate that one must be cautious in applying methods such as the Oliver–Pharr procedure (i.e., Eqs. (44) and (45)) because the force and initial unloading slope are both functions of loading and unloading conditions and histories. To further illustrate this point, we consider indentation in linear-viscoelastic solids with a constant Poisson’s ratio, \( \nu \). Suppose unloading takes place at \( t = t_m \) with a constant unloading rate of \( (dh/dt)|_{t_m} = -v_0 \), we have, using Eq. (83) for \( 0 \leq t \leq t_m + \Delta t \) and \( \Delta t \to 0 \),

\[
\frac{F(t_m + \Delta t) - F(t_m)}{\Delta t} = \frac{8 \tan \theta}{\pi(1 - \nu)\Delta t} \left[ \int_0^{t_m+\Delta t} G(t_m + \Delta t - \tau) \frac{dh(\tau)}{d\tau} d\tau - \int_0^{t_m} G(t_m - \tau) h(\tau) \frac{dh(\tau)}{d\tau} d\tau \right]
\]

\[
= \left. \frac{8 \tan \theta}{\pi(1 - \nu)\Delta t} \left[ \int_0^{t_m} G(t_m - \tau) h(\tau) \frac{dh(\tau)}{d\tau} d\tau + \Delta t \int_0^{t_m} \frac{dG}{dh} h(\tau) \frac{dh(\tau)}{d\tau} d\tau \right] h(\tau) \frac{dh(\tau)}{d\tau} d\tau + G(0) h(t_m^+) \Delta t - \int_0^{t_m} G(t_m - \tau) h(\tau) \frac{dh(\tau)}{d\tau} d\tau \right]
\]

\[
= \left. \frac{8 \tan \theta}{\pi(1 - \nu)} \left[ \int_0^{t_m} \frac{dG}{dh} \bigg|_{\eta = t_m-\tau} h(\tau) \frac{dh(\tau)}{d\tau} d\tau - G(0) h(t_m^+) v_0 \right] \right]. \tag{97}
\]

Here, the use of Eq. (83) for analyzing initial unloading slope may be justified because the contact area has not yet decreased as \( \Delta t \to 0 \). The initial unloading slope is then given by, using \( dF/dh = (dF/\,dt)/(dh/dt) \) and Eq. (84),

\[
\frac{dF}{dh} = \left. \frac{4 \tan \theta}{(1 - \nu)} \left[ G(0) h_c(t_m^+) - \frac{2}{\pi v_0} \int_0^{t_m} \frac{dG}{dh} \bigg|_{\eta = t_m-\tau} h(\tau) \frac{dh(\tau)}{d\tau} d\tau \right] \right]. \tag{98}
\]

The second term on the right-hand side is a function of loading condition and history. Unless this second term is negligible compared to the first term, which can be achieved by using fast unloading (i.e., large \( v_0 \)), the relationship between unloading slope and contact area, Eq. (43), is no longer valid. In practice, we may use several unloading rates, as we did here in finite element calculations, to test whether unloading is sufficiently fast. Once this limiting case is reached, Eq. (43) can be used to determine \( G(0)/(1 - \nu(0)) \) or \( E(0)/(1 - \nu^2(0)) \).

A number of authors have suggested adding a hold-period between loading and unloading to avoid the occurrence of the “bulge” on the unloading curves [230–233]. It has been shown recently that the initial unloading slope is nevertheless, a function of holding history and unloading condition [233–235]. A correction formula for the initial unloading slope has been proposed by Ngan et al. so that Eq. (43) is applicable for initial unloading [233–235]. It seems, however, that the hold-period may be unnecessary if unloading can be made sufficiently fast.

It is evident from the above discussion that only the instantaneous response of viscoelastic property can be obtained from the initial unloading equation. For a complete determination of linear viscoelastic properties, the equations developed in Sections 6.3 and 6.4.1 may be used for cases of constant Poisson’s ratios. More generally, however, finite element calculations together with statistical methods are necessary for extracting viscoelastic properties from indentation experiments. Several methods have recently appeared in the literature for determining viscoelastic and viscoplastic properties from indentation measurements [236–239].
6.5. Steady state oscillatory conditions

It is often desirable to obtain the mechanical response of viscoelastic solids under steady state oscillatory conditions. If the material is linear viscoelastic and is subjected to an applied shear strain of magnitude $g_0$ and frequency $\omega$

$$\gamma(t) = \gamma_0 e^{i\omega t},$$  

then the steady state shear stress response can be written as,

$$\sigma(t) = \sigma^* e^{i\omega t} = \left(\frac{\sigma^*}{\gamma_0}\right)\gamma_0 e^{i\omega t},$$  

(100)

where $\sigma^*$ is complex function of $\omega$. The complex quantity in parenthesis is called the complex relaxation modulus in shear analogous to the relaxation modulus in shear defined in the quasi-static loading cases. The complex relaxation modulus is typically written as [82–84]

$$\frac{\sigma^*}{\gamma_0} = G'(\omega) = G'(\omega) + iG''(\omega),$$  

(101)

where $G'(\omega)$ is the storage modulus and $G''(\omega)$ is the loss modulus.

A number of authors [240–247] have in recent years proposed experimental hardware and analysis procedures for determining the storage and loss modulus from instrumented indentation experiments using, for example, DMA, nanoindenters, or atomic force microscope with controlled oscillatory loads or displacements. Taking the advantage of the high lateral resolution, many of these techniques allow mapping of storage and loss modulus. It should be noted, however, that mathematical relations exist between relaxation modulus $G(t)$ and complex relaxation modulus $G'(\omega)$ and $G''(\omega)$. Specifically, these relations are as follows [82–84],

$$G'(\omega) = \omega \int_0^\infty G(t) \sin(\omega t) \, dt \quad \text{and} \quad G''(\omega) = \omega \int_0^\infty G(t) \cos(\omega t) \, dt.$$  

(102)

These relations are Fourier sine and cosine transforms and can be used to compute the real and imaginary parts of complex modulus for different frequencies $\omega$ when the relaxation modulus $G(t)$ is known. Conversely, using Fourier inverse transformation, the relaxation modulus, $G(t)$, can be evaluated from either $G'(\omega)$ or $G''(\omega)$,

$$G(t) = \frac{2}{\pi} \int_0^\infty \frac{G'(\omega)}{\omega} \sin(\omega t) \, d\omega \quad \text{or} \quad G(t) = \frac{2}{\pi} \int_0^\infty \frac{G''(\omega)}{\omega} \cos(\omega t) \, d\omega.$$  

(103)

However, these relations have not yet been explored in the context of indentation measurements. Indentation experiments on the same viscoelastic solid using both quasi-static and oscillatory conditions are highly desirable, because they would allow consistency check among various methods. These experiments can help provide a better understanding of the pros and cons of different indentation methods for probing viscoelastic materials. This understanding will help increase the use of indentation for probing mechanical behavior of soft matters which has had a long and renewed interest across several scientific disciplines, including materials science, chemistry, and biology.
7. Summary

We have provided an overview of the basic concepts of scaling and dimensional analysis and applied them to the modeling of instrumented indentation in elastic–plastic solids, power-law creep solids, and linear viscoelastic solids. Using this scaling approach, we have answered many of the basic questions posted in the Introduction section of this review.

For indentation in elastic–plastic solids, we derived equations for loading–unloading curves; identified parameters controlling piling-up and sinking-in; established relationships between hardness, mechanical properties, and indenter geometry; examined methods for estimating contact area from initial unloading slopes; discovered relationships between hardness, elastic modulus, and the work of indentation; and identified the limitations and possibilities of deducing stress–strain relations from load–displacement curves.

For indentation in power-law creep solids, we showed that the parameter, \( \dot{h}/h \), can indeed represent the indentation strain rate. We established equations describing loading curves and hardness, and applied them to several types of experimental conditions: (1) constant displacement rate, (2) constant loading rate, and (3) constant indentation strain rate. In particular, we showed that loading forces are no longer proportional to the square of indenter displacement and hardness decreases with increasing depth in both constant displacement rate and constant loading rate experiments. In contrast, the loading force is proportional to the square of indenter displacement and hardness reaches a steady-state value in constant indentation strain rate experiments. These results help establish a connection between power-law creep and “indentation-size effect”.

For indentation in linear viscoelastic solids, we provided a new derivation of the relationship between load and indenter displacement, and examined methods for obtaining viscoelastic relaxation modulus from loading curves under conditions such as constant displacement rate, constant indenter displacement, and constant indentation strain rate. Equations for loading curves were derived for indentation in a three-parameter “standard” viscoelastic solid. These equations can be used to help select indentation conditions for determining viscoelastic properties from experiments. We have also provided numerical examples of indentation in a viscoelastic solid that demonstrate the effects of loading and unloading conditions on indentation responses.

Taking together these results help form the basis for understanding indentation in much more complex materials systems. In particular, the scaling approach may help detect and confirm indentation size effects that have attracted much attention ever since the early days of the development of hardness tests [1–3,49]. Historically, indentation size effects refer to the variation of hardness as a function of indentation depth or load. It was noted by Brinell that the hardness values of steels depended on the diameter of spheres used in his tests [1]. Since then, indentation size effects have been reported in many materials, including metals [2], ceramics [49], and polymers [248–250]. The hardness values are usually observed to decrease with increasing indentation load or depth, though opposite trend of increasing hardness with depth has also been reported in some materials. The depth scale, where substantial size dependency is observed, ranges from a few nanometers to tens of micrometers and even millimeters. Presently, many factors are believed responsible for indentation size effects that may be categorized as materials-dependent or measurement-specific (see, Ref. [251] for a recent review). Specifically, the presence of a surface film, surface roughness, surface energy, composition variation, crystal isotropy, fracture, rounding of the indenter tip, friction, and loading conditions have all been proposed as possible causes. Recent years have seen increased interest and activities in applying strain-gradient plasticity models to explain indentation size effects in crystalline [252–261] and amorphous metals [262], as well as in polymeric materials [263]. Others took an energy balance approach by including surface-energies due to fracture [7,264–268].
surface energy contribution from compressing rough surfaces at the indenter-material interfaces has also been proposed to explain indentation size effects for a wide range of materials [269–271]. As discussed in Sections 4.1 and 4.3 size effects are consequences of the presence of some characteristic length scales. The variation of hardness with depth is one of many possible manifestations of such length scales. In particular, the analysis of the shape of indentation loading-displacement curves should help identify scale and size effects. By measuring the shape of load–displacement curves, it is also possible to distinguish length scales present in materials from other effects such as creep or viscoelasticity as seen from Sections 5 and 6. This method of detecting size effects would also avoid errors in the estimation of hardness at small scales due to problems in determining contact area. In addition to hardness, the analysis of the load–displacement curves can help determine whether size effects exist in other types of deformation phenomena, as demonstrated in recent studies of the microscopic shape memory and superelastic effects in shape memory materials [272,273]. Thus, a paradigm shift may be called for in the study of indentation size effects from hardness measurements to the direct examination of the shape of indentation load–displacement curves with dimensional analysis as a guide.

We conclude this review with the opening sentences of Lord Rayleigh’s 1915 paper on “The Principle of Similitude” [61]: “I have often been impressed by the scanty attention paid even by original workers in physics to the great principle of similitude. It happens not infrequently that results in the form of ‘laws’ are put forward as novelties on the basis of elaborate experiments, which might have been predicted a priori after a few minutes’ consideration”. We hope that this review serves to demonstrate that elaborate experiments and computer simulations guided by scaling and dimensional analysis can help provide deep insights into problems of interest to materials scientists and engineers.

Acknowledgements


References