纳米硬度技术的发展和应用

张志华  杨波

中国科学院力学研究所非线性力学国家重点实验室，北京 100080

摘要 近二十年来，主要在于检测材料表面微米和亚微米尺度力学性能的纳米硬度技术发展迅速。首先，概述硬度的定义、分类及其适用范围。然后，着重就纳米硬度技术的发展，重点介绍纳米压痕硬度的测量原理及其影响因素。连续扫描测量原理，高分辨率的载荷位移测量原理，几种常用压头的几何形状，试样表面的准备和确定，相关的测试方法，仪器校准和显微观察等问题。通过压痕实验可获得硬度、弹性模量、断裂韧性、存储模量和损耗模量、蠕变应力指数等。最后，简要介绍纳米硬度测量技术的发展和应用。

关键词 纳米压痕，纳米划痕，微力学性能，硬度，模量，断裂韧性，蠕变应力指数

1 引言

硬度 (hardness) 是评价材料力学性能的一种简单、有效的手段，已有近百年的应用历史。关于硬度的定义，目前尚未统一。从作用形式上，可定义为“某一物体抵抗另一物体产生变形能力的度量”；从变形机理上，可定义为“抵抗弹性变形、塑性变形和破坏的能力”或“材料抵抗残余变形和破坏的能力”。无论如何定义，在测量固体材料硬度时，总是会将一定形状和尺寸的较硬物体以一定的压力接触被测材料表面。硬度测量，不仅与材料的弹性模量、屈服强度、抗拉强度等力学性能有关，还与测量仪器本身的测量条件有密切关系。所以，硬度本身不是一个物理量，而是材料局部区域力学性能在特定条件下的整体表现。它是材料对外界物体机械作用 (压入或刻划) 的局部抵抗能力的一种表现，反映了固体物质凝聚或结合强弱的程度[1~4]。

为了比较各种固体材料的软硬，人们给出多种不同的硬度标准及其测量方法，归纳起来有：静态压痕硬度 (static indentation hardness)，动态压痕或回弹压痕硬度 (dynamic or rebound hardness)，划痕硬度 (scratch hardness)[2~3]。

静态压痕硬度测量，通过球体、金刚石锥体或其他锥体将力施加在被测材料上，使材料产生压痕，再根据施加载荷与所产生压痕面积或深度之间的关系，给出其硬度值。根据施加载荷的大小可分成：宏观硬度 (日本、美国和前苏联等定为 10N 以上，欧共体国家和国际机构则定为 2N 以上)，显微硬度 (上限：10N 或 2N；下限：10mN 左右)，纳米硬度 (一般在 700mN 以下，有的生产商为了便于研究者模拟显微硬度，配有 10N 载荷附件(www.mts.com/nano))。不同的静态压痕硬度标准及其测量方法见表 1，载荷范围见图 1[2~4]。

收稿日期：2001-06-01，修回日期：2002-05-08
* 国家自然科学基金 (10102021，10172086) 和中国科学院知识创新工程 (KJCX2-SW-L2，KGCX1-11) 资助项目
表 1 常用静态压痕硬度测量方法比较 [2]

<table>
<thead>
<tr>
<th>硬度实验</th>
<th>压头形状</th>
<th>压痕对角线直径 / μm</th>
<th>测量深度 / μm</th>
<th>载荷 / N</th>
<th>测量方法</th>
<th>表面制备</th>
<th>应用范围</th>
<th>备注</th>
</tr>
</thead>
<tbody>
<tr>
<td>布氏</td>
<td>2.5 mm 或 10 mm 直径的钢球</td>
<td>1000 ~ 5000  &lt; 1000</td>
<td>钢铁用 30 000，软金属低于 1000</td>
<td>显微镜下测压痕直径，显微镜下测表面积</td>
<td>由精确测量表面积直径，块状材料适合平整表面</td>
<td>使用轻载荷的球形压头，使表面的破坏程度减小至最小</td>
<td></td>
<td></td>
</tr>
<tr>
<td>洛氏</td>
<td>120° 金刚石锥体，或 1.99 mm 直径的球体</td>
<td>100 ~ 1 500 25 ~ 350</td>
<td>主载荷 600，上直接读取载荷数值，测定平面</td>
<td>显微镜下测压痕直径，显微镜下测表面积</td>
<td>由精确测量表面积直径，块状材料适合平整表面</td>
<td>从压痕深度测量压头，可使用测量压头比布氏硬度大。</td>
<td></td>
<td></td>
</tr>
<tr>
<td>表面洛氏</td>
<td>120° 金刚石锥体，或 1.99 mm 直径的球体</td>
<td>10 ~ 1 000 1 ~ 100</td>
<td>主载荷 150，300，450；副载荷 30</td>
<td>显微镜下测压痕直径，显微镜下测表面积</td>
<td>由精确测量表面积直径，块状材料适合平整表面</td>
<td>压痕尺寸和载荷均比洛氏大</td>
<td></td>
<td></td>
</tr>
<tr>
<td>维氏</td>
<td>136° 的金刚石正四棱锥</td>
<td>10 ~ 1 000 1 ~ 100</td>
<td>10 ~ 1 200，副载荷 0.25</td>
<td>显微镜下测压痕直径，显微镜下测表面积</td>
<td>由精确测量表面积直径，块状材料适合平整表面</td>
<td>对于表面性能变化的灵敏度低，压头小。</td>
<td></td>
<td></td>
</tr>
<tr>
<td>努氏</td>
<td>136° 的金刚石正四棱锥</td>
<td>10 ~ 1 000 0.3 ~ 30</td>
<td>2 ~ 40，副载荷 0.01</td>
<td>显微镜下测压痕直径，显微镜下测表面积</td>
<td>由精确测量表面积直径，块状材料适合平整表面</td>
<td>用于脆性材料的微结构及成分的研究</td>
<td></td>
<td></td>
</tr>
<tr>
<td>贝氏</td>
<td>中心线与镜面之间夹角为 85.3° 的三棱镜</td>
<td>0.01 ~ 10 0.001 ~ 1</td>
<td>实时测量压痕直径和载荷</td>
<td>显微镜下测压痕直径，显微镜下测表面积</td>
<td>超薄表面的纳米压痕硬度计</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

动态或回弹压痕硬度测量，是将一个具有标准重量与尺寸的物体从某一高度（具有一定的势能）下落到被测物体的表面，从其表面弹起，根据其回弹高度来确定被测试物体的硬度值。该方法主要用于金属材料 [2,3]。

划痕法是在小曲率半径的硬质钢球上施加一定的法向力，并使压头沿拟样表面刻划，通过试样表面的划痕来评价其硬度。它是测量块体材料或表面涂层材料抵抗划痕（scratch resistance）的能力、摩擦、变形和薄膜对基体粘着能力的方法 [2,3]。

各种硬度的定义不同，测定的方法和适用的材料各异，各种硬度值不能直接比较，但可以进行相应的比较。
通过实验进行比较，某些金属材料的硬度与其强度指标（如屈服极限、强度极限等）有一定的关系，通过较简单和经济的硬度测量即可推断材料的强度指标[3,4]。

宏观硬度和显微硬度测量适用于较大尺寸的试样，仅能得到材料的塑性性质。随着现代材料表面工程（气相沉积、溅射、离子注入、高能束表面改性、热喷涂等）、微电子、集成光微机电系统、生物和医学材料的发展，该类仪器在进行电子薄膜、各类涂层、材料表面及其表面的力学性能检测的理论手段。它不需要将表层从基体上剥离，可以直接给出材料表层力学性质的空间分布。例如，可检测出焊点及其附近材料的力学性能。由于实验准备简单，即使是材料大到可以用其它宏观方法检测，该方法仍然是一种可供选择的方法。


2 纳米压痕硬度

纳米压痕硬度计（nanoindentation, instrumented-indentation testing, depth-sensing indentation, continuous-recording indentation, ultra-low load indentation）是一类先进的材料表面力学性能测试仪器。该类仪器装有高分辨率的致动器和传感器，可以控制和监测压头在材料中的压入和退出，能提供高分辨率连续载荷和位移的测量，明显拓宽了传统显微硬度的能力。该技术的主要优点：可直接从载荷-位移曲线中实时获得接触面积，把人们从寻找压痕位置和测量残余面积的繁琐劳动中解放出来，这样可以大大地减小误差，非常适合于较浅的压痕深度；能完成多种力学性能的测试，只检测测量的是硬度（H）和弹性模量（E）；对不会导致压痕周围凸起（pile-up）的材料，如大多数陶瓷、硬金属和加工硬化的软金属，硬度和弹性模量的测量精度通常优于 10%。

2.1 压痕测试原理

2.1.1 确定硬度和弹性模量的原理和模型[5]

在加载过程中，试样首先发生弹性变形，随着载荷的增加，试样开始发生塑性变形，加载曲线呈非线性；卸载曲线反映了被测试物体的弹性恢复过程。通过分析加载卸载曲线可以得到材料的硬度和弹性模量等。图 2(a) 给出了典型的整个加载和卸载过程中的压痕载荷 (load or load on sample) 与位移 (displacement or displacement into surface) 之间关系的曲线。图 2(b) 为一轴对称压头 (tip) 在加载卸载过程中任一压痕剖面的示意图。在压头压入材料的过程中，材料经历了弹性和塑性变形，产生了同压痕形状相一致的压痕接触深度（contact depth）h。和接触圆半径（radius of the circle of contact）a。在压头退出过程中，仅弹性位移恢复。硬度和弹性模量可由最大载荷 Pmax，最大压入深度 hmax，卸载后的残余深度 hψ 和卸载曲线的端部斜率 S = dP/dh(称
为弹性接触刚度 elastic contact stiffness 中获得。

假设：试样为各向同性材料，其几何尺寸远大于压痕的尺寸，材料表面平整；不存在与时间相关的变形，即是蠕变和粘弹性。定义硬度 \( H(\text{hardness}) \) 和复合响应模量 \( E_r(\text{composites response modulus}) \)

\[
H = \frac{P}{A}
\]

式中 \( P \) 为在任意压痕深度的实时载荷，\( A \) 为在 \( P \) 作用下接触面积的投影面积。根据此定义，纳米压痕硬度是材料对接触载荷承受能力的量化。这里，请注意该定义与传统显微硬度定义 \( H = P_{\text{max}}/A_{\text{residual}} \) 的区别。对于塑性形变起主作用的过程，两种定义给出类似的结果；但，对于弹性形变为主的过程，两种定义给出完全不同的硬度，因为纯弹性接触过程，剩余接触面积非常小，传统的定义将导致硬度无穷大。式 (2) 来源于弹性接触理论 [6~8]，\( E_r \) 可以被用来解释压头和试样的弹性形变，\( \beta \) 为与压头形状有关的常数，对不同形状的压头，\( \beta \) 数值不同：圆锥压头 \( \beta = 1.000 \)，Berkovich (三棱锥) 压头 \( \beta = 1.034 \)，Vickers 压头 (四棱锥) \( \beta = 1.012 \)。被测试样的弹性模量可从下式中获得

\[
\frac{1}{E_r} = \frac{1 - \nu^2}{E_1} + \frac{1 - \nu_1^2}{E_1}
\]

式中 \( E, \nu \) 分别为被测材料的弹性模量和泊松比；\( E_1, \nu_1 \) 分别为压头的弹性模量和泊松比。对于金刚石压头，\( E_1 = 1141 \text{ GPa}, \nu_1 = 0.07 \)。这里，要计算出 \( E \)，必须先知道 \( \nu \)。粗略估计一下，当 \( \nu = 0.25 \pm 0.1 \)，对大多数材料的弹性模量 \( E \) 仅会产生 5.3% 的不确定度。所以，在不知道被测材料 \( \nu \) 的情况下，可取 \( \nu = 0.25 \)，这样不会引起较大的误差。

为了从载荷 - 位移数据计算出硬度和弹性模量，必须准确地知道弹性接触刚度和接触面积。通过卸载后的残余压痕照片获得投影接触面积，很不方便，对于纳米尺度的压痕，这是几乎不可能。只有通过连续载荷 - 位移曲线如图 2(a) 计算出接触面积。目前，Oliver-Pharr 法是计算接触面积最常用的方法。这种方法通过如下函数拟合载荷 - 位移曲线的卸载部分

\[
P = B(h - h_f)^m
\]

式中 \( B \) 和 \( m \) 是通过测量获得的拟合参数，\( h_f \) 为完全卸载后的位移。弹性接触刚度便可以根据式 (4) 的微分计算出

\[
\]
对于一条拟合的卸载曲线，式（5）不一定总能提供正确的描述。例如，对附着在基底上的薄膜材料，式（5）就不一定总是正确的。在这种情况下，根据整条卸载曲线拟合得到的参量常常导致非常大的误差。因此，确定接触刚度的曲线拟合通常只能取卸载曲线顶部的 25%~50%。

对于弹性接触，接触深度（hc）总是小于压入深度（即最大位移 h），由下式计算
\[
    h_c = h - \frac{P_{\text{max}}}{S}
\]
（6）
式中 ε 为与压头形状有关的常数。对于球头或棱锥形（如 Berkovich 和 Vickers）压头，ε = 0.75；锥形压头，ε = 0.72。虽然式（6）未源于弹性接触理论，但对塑性变形也符合得相当好。值得注意的是，式（6）不能说明凸起（pile-up）的塑性现象，因为假设接触深度总是小于压入深度。接触面积的投影 A 可由经验公式 \( A = f(h_c) \) 计算。对于一个理想的 Berkovich 压头，A = 24.56h_c^2。但实际使用的压头往往都会偏离理想情况。因此，实际压头的接触表面的投影面积一般表示为一个级数
\[
    A = 24.56 h_c^2 + \sum_{i=0}^{5} C_i h_c^{1/2i}
\]
（7）
其中 C_i 对不同的压头有不同的值，具体由实验确定。

如果知道接触表面的投影面积和接触刚度，硬度和弹性模量就可以由式（1）和式（2）求出。

2.1.2 凸起的影响

上面确定硬度和弹性模量的方法完全基于压痕接触的弹性模型。如果使用尖压头如 Berkovich 和 cube-corner 压头或球形压头，当载荷较大时，会发生弹性与塑性变形，有时塑性变形起重要作用。塑性模量不能单独解释。最重要的塑性现象是凸起，见图 3。这种现象在压痕过程中有时会发生。一旦发生，实际接触面积就会超过上面基于弹性接触模型的预测结果（如在纯弹性接触中材料会凹陷）。如果用式（1）和式（2）确定硬度和弹性模量，其值会偏高。压痕有限元模拟结果显示（按面积-深度关系将 Berkovich 和 Vickers 压头折算成圆锥头），硬度和弹性模量分别偏高 60% 和 30%（9,10）。

![图 3 压痕过程中的凹陷 (sink-in) 和凸起 (pile-up)](image)

目前，已能用有限元模拟的方法确定出最易发生凸起现象的材料类型和条件（9~11），屈服应力 / 弹性模量 (σ_y/E) 和加工硬化 (work-hardening) 性质是最重要的影响因素。一般说来，当 σ_y/E 很小且加工硬化很低甚至没有，凸起最严重。加工硬化能抑制凸起，因为在变形过程压头周材的表面硬化限制材料向表面上流动。

有限元模拟等效 Berkovich 和 Vickers 的圆锥压头结果显示，当 σ_y/E > 0.03 时，凸起现象不明显，与加工硬化特性无关。通常情况下，由于不知受试材料的 σ_y/E 值，因此该参数使用不方便。而 σ_y/E 和 h_f/h_{max} 之间的相关性，可通过直接测量 h_f/h_{max} 判断。模拟结果显示，当 h_f/h_{max} < 0.7，凸起最少，否则，凸起将与加工硬化相关。典型的 h_f/h_{max} < 0.7 材料是陶瓷和
硬金属\cite{12}。在实际情况下，如果根据 $h_f/h_{\text{max}}$ 或其他方法判断可能会在明显的凸起发生，就应该通过光学显微镜、扫描电镜或原子力显微镜观察实际的压痕凸起情况，以便确定真实的接触面积。如果再根据载荷，深度曲线来计算接触面积，就不可能得到精确的实验结果。

2.1.3 连续刚度的确定

上述确定接触刚度的方法是根据卸载曲线起始点的斜率来计算的，它只能得到最大压痕处对应的硬度和弹性模量。Petlica, Oliver 和 Pharr 等\cite{13-16} 提出在加载过程中连续计算接触刚度的测量原理。该原理是将相邻较高频率（如 45 Hz）的小简谐力叠加在准静态的加载信号上，测量压头的简谐响应，简谐力控制着位移的波动，维持的典型值为 1~2 nm。接触刚度的连续测量，实现了硬度和弹性模量随压痕深度变化的连续测量。这种方法被称为连续刚度测量法 (continuous stiffness measurement, CSM) 或固定频率的动态压痕测试 (dynamic indentation, FSDI)。该技术将压痕测试系统简化为图 4 所示的动力学模型。压杆质量 $m$ 由刚度为 $K$ 的两个叶片弹簧支撑，弹簧的特点为在叶片平面内刚度很高而在垂直方向上刚度很低。带有压头的压杆被线圈 - 磁铁装置驱动，压头上的准静态载荷由加载在探头由缓慢变化的电流控制。再叠加小简谐力，这可由锁定放大器的振荡器完成。位移由平衡电容器测量。所有的载荷都严格限制在一个自由度上。

图 4 压痕测量原理示意图及其动力学模型

图 4 的纳米压痕系统可以用一维简谐振子模型描述。其运动方程可表示为

$$m \ddot{Z} + D \dot{Z} + K Z = F(t)$$

（8）

式中 $D = D_{1} + D_{a}$ 为等效阻尼，其中 $D_{1}$ 和 $D_{a}$ 分别为系统和试样的阻尼； $K = (S^{-1} + K_{1}^{-1})^{-1} + K_{0}$，为等效刚度，其中 $K_{1}$ 为加载框架刚度； $F(t)$ 为总的载荷。假设载荷函数可表达为

$$F(t) = F_{0} e^{i\omega t}$$

（9）

产生的位移

$$Z(t) = Z_{0} e^{i(\omega t - \phi)}$$

（10）

式中 $F_{0}$ 为激励载荷幅值， $Z_{0}$ 为位移幅值， $\phi$ 为位移滞后载荷的相位角， $\omega = 2\pi f$ 为角频率。将式 (9) 和式 (10) 代入式 (8)，等同于

$$F_{0}/Z_{0} = \sqrt{K - m\omega^{2}} + (\omega D)^{2}$$

（11）

$$\tan \phi = \omega D / (K - m\omega^{2})$$

（12）

从上面这两个方程中可得到 $K$ 和 $D$。最后，可得试样的接触刚度和接触阻尼系数。
式中，\( K_f, K_s, m \) 和 \( D_0 \) 均为仪器本身的参数，而 \( \omega \) 为实验设置常数，\( F_0, Z_0 \) 和 \( \phi \) 则由实验测量。因此，根据式 (13) 和式 (14) 可确定出试样的接触刚度 \( S \) 和接触阻尼系数 \( D_\omega \)。

### 2.2 压痕测量技术

#### 2.2.1 载荷和位移的测量

纳米压痕硬度测量的商品化仪器主要由三部分组成：固定在刚性杆上特殊形状的压头，致动器提供动力，传感器测压头位移。

商品化仪器的差别主要表现在压头的驱动和位移的测量上。致动器主要有三种设计方式：

1. 电磁致动（www.mts.com/nano；www.csem.ch/instrum；www.micromaterials.co.uk），图 4 所示的设计原理是最典型的一种形式，压头的驱动是基于载流线圈在磁场中受力原理。磁场中的线圈通过电流时，产生的安培力为 \( F = kBI \)。如果上式中 \( k \)（与线圈结构、磁场参数等有关的常数）、\( B \)（磁感应强度）和 \( I \)（线圈匝数）均为常数，该式可改写为 \( F = KI \)。在标定 \( K \) 值后，可通过对称调节电流强度 \( I \)，实现对压载力大小的精确控制。目前，量程为 10 mN 时，分辨力可达 1 nN（www.mts.com/nano）。

2. 静电致动，由动极板和定极板组成的电容器提供静电力（www.hysitron.com）。


位移测量有两种设计方式：

1. 电容位移传感器（www.mts.com/nano；www.csem.ch/instrum；www.hysitron.com；www.micromaterials.co.uk），平衡板电容器的定极板和动极板间的间距为 \( d \)，电容的相对变化为 \( \Delta C/C = 2\Delta d/d \)。若压头行程为 2 mm，分辨力可达 0.002 nm（www.mts.com/nano）。


目前，在不断提高测试分辨率的同时，工作频率也在拓宽，范围已达 0.1~300 Hz。该类仪器除了能研究常见块体材料和各种薄膜外，还适用于研究超薄膜、聚合物的动态特性、软材料如软组织等（www.mts.com/nano）。

#### 2.2.2 压头几何形状的选择

选择合适的压头需要考虑许多因素。压头材料需选用具有很高硬度和弹性模量的金刚石，这样可以减小压头在测量过程中对位移的影响。按其形状可分为棱锥体压头（如 Vickers 压头、Berkovich 压头、Cube Corner 压头等）和光滑旋转体压头（如圆锥压头、球压头等）两大类。表 2 列出了该类压头各种常用压头的几何参数。表中，锥角 \( \psi \) 为中心线与锥面的夹角，棱锥体压头的等径锥体是按相同的面积 - 深度关系折算成圆锥体压头的锥角。三种典型压头形状见图 5。

一般认为，当与压头接触的试样材料为完全塑性时，所获得的硬度和屈服应力测量值之间成正比关系。在小载荷作用下，如果压头的面积函数接近于理想压头的面积函数，试验就能发生塑性变形。所以，压头面积对测量结果的影响较大。大多数压头的端部在很小尺寸范围内都是球形的。压头的端部半径越小，较浅的压痕深度就可获得理想的测量结果；反之，就需要较...
深的压痕深度，在加工磨制压头时，三个面容易交于一点，四个面容易导致压头顶端出现楔角（chisel edge），小半径的球形压头很难加工[17]。

### 表 2 常用压头的几何参数[17]

<table>
<thead>
<tr>
<th>参数</th>
<th>Vickers</th>
<th>Berkovich</th>
<th>Cube-corner</th>
<th>Cone (angle ψ)</th>
<th>Sphere (radius R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>中心线与面的夹角 (α)</td>
<td>68°</td>
<td>65.3°</td>
<td>35.2644°</td>
<td></td>
<td></td>
</tr>
<tr>
<td>边长 / 深度 (l/h)</td>
<td>4.9502</td>
<td>7.5315</td>
<td>2.4491</td>
<td></td>
<td></td>
</tr>
<tr>
<td>投影面积 (A)</td>
<td>24.504 h²</td>
<td>24.56 h²</td>
<td>2.5981 h²</td>
<td>πα²</td>
<td>πα²</td>
</tr>
<tr>
<td>体积 / 深 度 体积 / 面 积 (V/h)</td>
<td>8.1831 h³</td>
<td>8.1783 h³</td>
<td>0.8657 h³</td>
<td>0.067 A³/2</td>
<td>0.067 A³/2</td>
</tr>
<tr>
<td>体积 / 面积 (V/A)</td>
<td>0.927</td>
<td>0.908</td>
<td>0.5774</td>
<td></td>
<td></td>
</tr>
<tr>
<td>等效棱角 (ψ)</td>
<td>70.2996°</td>
<td>70.32°</td>
<td>42.26°</td>
<td>h tan ψ</td>
<td>(2h/h - h²)³/2</td>
</tr>
</tbody>
</table>

![Berkovich压头](www.hysitron.com) ![Cube-Corner压头](www.hysitron.com) ![Conical压头](www.hysitron.com)

图 5 三种典型压头形状 (www.hysitron.com)

在显微硬度中，常用 Vickers 压头，但压头顶端的楔角将会导致在不同尺度下压头几何形状下不能自相似。目前，最好的加工水平为楔角长度约 1 μm。随着压痕深度的减小，楔角引入的误差会逐步增大。为克服该现象，设计了三棱锥体并与 Vickers 压头有相同面积 - 深度关系 (A = 24.5 h²) 的 Berkovich 压头。Berkovich 压头是纳米压痕硬度计中最常用的，它可以加工得很尖，而且几何形状在很小尺度内保持相似，适合于尺度的压痕实验。目前，该类压头的加工水平：端部半径 ≤ 50 nm，典型值约 40 nm，中心线和面的夹角精度为 ±0.025°。大多数理论分析和有限元模拟的压头形状假设是轴对称的，棱锥体压头不满足此假设。三维有限元模拟结果显示，实际差别很小。总结，在纳米压痕硬度测量中，Berkovich 压头是一种理想的压头。优点包括：易获得好的加工质量，很小载荷就能产生塑性，能减小摩擦的影响，能和 Vickers 压头的结果直接对比。图 6 为该压头在三种材料上的压痕形状，残余压痕形状取决于材料特性。

![二氧化硅](www.mts.com/nano) ![镍](www.mts.com/nano) ![铝](www.mts.com/nano)

图 6 Berkovich 压头在三种材料上的压痕形状 (www.mts.com/nano)

Cube-corner 压头 因其三个面相互垂直，像立方体的一个角，故取此名称。压头越尖，就会在接触区内产生理想的应力和应变。目前，该种压头主要用于断裂韧性（fracture toughness）的研究[18～20]。它能在脆性材料的压痕周围产生很小的规则裂纹，这样的裂纹能在相当小的范围内用来估计断裂韧性。

球形压头 球形压头的接触不同于棱锥体的 Berkovich 和 Vickers 压头接触的应力 - 应变场。球形压头的初始接触应力小，仅产生弹性变形。当球形压头压入表面后，弹性变形开始向塑性变形发展。理论上可以用来确定屈服应力和加工硬化，并可以从单个压痕实验数据中再现整个...
单轴应力应变曲线，这在球半径的球形压头中得到了成功地应用。在微米尺度，由于难以获得高质量的金刚石球形压头，使用受到限制。这也是在大多数小尺度实验中选用 Berkovich压头的原因。

锥形压头：圆锥具有尖的自相似几何形状，从模型角度常利用它的轴对称特性，纳米压痕硬度的许多模型均基于圆锥压痕。由于加工出尖的圆锥金刚石压头，它在小尺度实验中很少使用。

2.2.3 试样表面的准备和确定

目前，纳米压痕硬度计的载荷和位移的测量分辨率已达到相当高的水平。在2.1.1节中，模型假设材料表面是平的，计算弹性模量和硬度所需的接触刚度和面积依赖于接触深度。实验之前，需对某些试样表面进行抛光（如果是机械抛光，试样会发生表面硬化），可按 ASTM E-380规定制样。事实上，试样的物理表面具有一定粗糙度，且会附着几个纳米厚的水膜等，为小压痕深度的测量，如何精确确定压头和试样的接触点成为决定测试结果的主要问题。

图7(a)是典型低弹性模量高聚物的初始加载曲线，压头在逼近试样表面的过程中，在约340 nm距离范围内出现负载荷。这是由于表面的黏着特性将压头拉近试样表面所致。高聚物表面的吸附作用，导致接触点难以确定。图7(b)为压头靠近熔融硅表面的载荷曲线，熔融硅较硬，具有良好吸湿特性。压头逼近试样表面时，有明显的吸附区域。作为参考，实际的接触点通常采取载荷曲线重新回到零点的位置。这是一种确定表面的典型方法。还有其他方法，如在压头的周围装有一个直径为几nm的圆环结构，通过此结构判断压头和试样的接触（www.csem.ch/instrum）。

2.2.4 相关技术

控制方式：目前，已发展了四种控制方式：常载荷率 $P' = \text{const}$，常位移率 $h' = \text{const}$，常 $P' / P = \text{const}$，常应变率 $h'' / h = \text{const}$。要根据不同实验目的具体确定控制方式。如果材料是应力率敏感的，所测材料的硬度将依赖于 $h'' / h$。要获得体材料的硬度值，必须被控制。后两种控制方式之间，存在下面关系

$$ h' / h = P' / (2P) $$  (15)

环境要求：振动和温度的变化会对位移测量的不确定度和误差。为减小振动，商品化仪器均带有减振系统，仪器尽量放置在振动低的环境内。温度的变化，会导致试样和实验系统的膨胀和收缩。仪器应放置在室温稳定的实验环境中，温度控制在±0.5°C。

仪器校准：通常包括仪器校准或刚度校准、压头面积函数校准和热漂移校准。仪器刚度或刚度的校准和压头面积函数的校准由生产商完成。压头使用一段时间之后，面积函数也需使用者重新校准。热漂移校准是修正位移测量过程中压头和试样热膨胀的，每次压痕实验均可由仪器自带的软件自动控制完成。

2.3 压痕硬度的应用


2.3.1 硬度和弹性模量

张志华，杨玉敏使用 MTS Nano Indenter XP 进行了压痕实验。试样分别为二氧化硅（SiO₂）、钛膜（Ti Film）、类金刚石膜（DLC Film）。

SiO₂ 是 MEMS 中的常用材料。其表面光滑，抗氧化，非晶，各向同性，无加工硬化，中等范围的力学特性，典型的陶瓷行为。在卸载时有较大的弹性恢复，无明显时间相关性。故常被选为该类仪器的标准试样。从图 8 中可以看出，SiO₂ 的压痕模量和硬度随压痕深度变化。

图 8

使用离子溅射技术，在氮化硅陶瓷上沉积约 1000 nm 的钛膜。从图 8 中可以看出，在压痕深度约在 50~300 nm 范围内，硬度和压痕模量保持稳定，这是薄膜的力学性质。随着压痕深度
的增加，硬度和压痕模量不断增大，这是基体对薄膜的影响所致。当压痕深度超过1000 nm时，硬度和压痕模量随压痕深度增加的趋势变缓，薄膜对基体的影响不明显。这反映了软膜硬基体材料力学性能变化的规律。

硬的 DLC 膜常作为保护膜使用。气相沉积 DLC 膜。从图 8(a) 中可以看出，在压痕深度为700 nm 左右时加载曲线上有一平台。对一般金属材料，加载曲线上多为连续变化的。经分析可知，这一平台对应于膜厚，是由膜的粘着失效 (adhesion failure) 造成的。从图 8(b) 和图 8(c) 中可以看出，在测试范围内，硬度和压痕模量随压痕深度而减小。当压痕深度超过700 nm时，压痕模量接近于 200 GPa 即钢的弹性模量。这反映了硬膜软基体材料力学性能变化的规律。

从上面的实验结果看出，CSM 技术能检测出硬度和弹性模量随压痕深度变化的连续规律。所以，该技术是研究基体对薄膜力学性质影响规律的有效手段。在显微硬度测量中，一般认为压痕深度控制在 10% 以内，基体材料对薄膜的力学性质没有影响 [10]。依据此经验，我们能大致得到薄膜、基体和界间的力学性质。

需要说明的是，Oliver-Pharr 确定接触面积的方法对薄膜并不总是很理想的。对于硬基体软膜，基体约束薄膜塑性流动，凸起有时会过低估计实际的接触面积，导致硬度和弹性模量偏大。对于软基体硬膜，凹痕有时会过高估计实际的接触面积，导致硬度和弹性模量偏小。对于过分的凸起或凹陷，必须发展新的方法描述真实的接触面积 [17]。

2.3.2 断裂韧性

断裂韧性是指材料抵抗断裂的能力，是用 Berkovich 压头获得材料的硬度 H 和弹性模量 E，然后用 Cube-corner 压头在材料中产生径向裂纹，见图 9。断裂韧性与径向裂纹长度的关系为

$$K_c = a \left( \frac{E}{H} \right)^{1/2} \left( \frac{P}{C^3/2} \right)$$

式中 P 为使用 Cube-corner 压头的最大作用荷载，C 为径向裂纹长度，a 为与压头形状相关的经验系数。由于 Cube-corner 压头要比 Berkovich 压头尖得多，易于在压头周围材料中产生较大的应力和应变，便于压痕裂纹的形成和扩展。

在较小载荷作用下，Cube-corner 压头可以在许多脆性材料中产生带有径向裂纹的亚毫米压痕，特别适合于检测薄膜或材料微小区域的断裂韧性。

2.3.3 时间相关材料的特性

上面对硬度和弹性模量的讨论基于材料的响应和时间无关，或近似如此。大多数金属和陶瓷材料在室温条件下的压痕实验具有这种特性。实际上，材料的响应与时间相关，而且相关的程度和特性明显依赖于温度。除了聚合物在室温条件下的黏弹性外，金属和陶瓷的在升温过程中时间相关的蠕变也是很重要的现象。目前，描述时间相关现象的方法不断地发展。

2.3.3.1 载荷刚度材料的硬度和弹性模量

对于与时间相关的黏弹性和蠕变变形，能导致压头在卸载时位移的增加。更有甚者，当载荷维持不变或减小时位移继续增加，这将会产生负斜率的卸载曲线。即式 (5) 为负，式 (6) 将违反 “对于弹性接触，接触深度总是小于总的压入深度” 的假设。由于连续刚度测量模型利用的是加载时的实验数据，不受卸载时负刚度的影响。因此，该技术可以用来研究与时间相关的黏弹性和蠕变变形的材料。
2.3.3.2 黏弹性

连续刚度测量技术能描述聚合物这类黏弹性材料的时间相关特性。使用力和位移随时间变化的幅值和相位，可以测量描述弹性的存储模量（storage modulus $E'$）和描述内摩擦和阻尼的损耗模量（loss modulus $E''$）[28,29]。在最简单形式中，把接触处理成一个刚度为 $S$ 的弹簧和与之并联的接触阻尼系数为 $D_w$ 的黏壶。存储模量和损耗模量分别表示为

$$
E' = \frac{\sqrt{\pi} S}{2 \beta \sqrt{A}}
$$

$$
E'' = \frac{\sqrt{\pi} D_w \omega}{2 \beta \sqrt{A}}
$$

2.3.3.3 蠕变参数的测量

对蠕变（creep）响应明显的材料，纳米压痕技术能量化蠕变速率。对单轴拉伸蠕变实验，温度和应力与蠕变速率的关系为

$$
\dot{\varepsilon} = \alpha \sigma^n \exp(-Q_c / RT)
$$

式中 $\alpha$ 为材料常数， $\sigma$ 为应力， $n$ 为蠕变应力指数， $Q_c$ 为活化能， $R$ 为气体常数， $T$ 为温度。对大多数金属， $n$ 的典型范围为 3～5。通过类比的方法，可以直接获得压痕蠕变的等效表达式。例如对平行管，记录压头位移随时间的变化，定义对棱锥体和锥形压头适用的压痕应变率

$$
\dot{\varepsilon}_t = \frac{\dot{h}}{h}
$$

平均接触压力 $H = P/A$ 可等效于应力，类比式（19），压痕蠕变实验为

$$
\dot{\varepsilon}_t = \alpha_t H^n \exp(-Q_c / RT)
$$

式中 $\alpha_t$ 为材料常数。实验发现，该式对描述一些材料的蠕变是合适的 [30～33]。目前，压痕蠕变的精确计算仍然在低熔点金属，这类金属在室温下蠕变。例如铜，通过控制不同的压痕应变率 $h'/h$，可获得系列的硬度 - 压痕深度数据，见图 10(a)。图 10(b) 为压痕应变率 - 硬度的对数曲线，每个数据点是五个压痕实验的平均值，数据拟合线的斜率为 7.3 ($R = 0.999$)，这就是钢的压痕蠕变应力指数，与文献 [35] 的室温拉伸蠕变实验获得的应力指数 7.6 接近。

3 纳米划痕硬度

在材料工程和许多工业应用中，了解材料或机器部件的抗磨损性能和表面变形机理是非常重要的。划痕实验过程包含了许多变形和破坏过程。以前，划痕实验不能提供定量和重复的结果。作为一种定性的方法，应用受到限制。随着实验技术的发展，又重新受到重视。

目前，纳米划痕硬度计主要针对研究材料抵抗正压力为小载荷如 1～100 mN 范围内的划刻能力而设计的，已能定量分析材料表面的摩擦性能和描述划痕变形机理。同以前的划痕计相比，提高了载荷和位移的分辨率，并改进了对临界载荷的确定方式 [2]。
3.1 划痕测量技术


棱锥体压头和光滑旋转体压头的选择根据材料性质而定。压头形状的轻微变化会导致实验结果的明显差异。即使同一台纳米划痕仪器，不同的压头会得到明显差异的结果。所以要求压头的加工形状尽量保持一致，以便提供定量、可信和重复的实验数据 [33]。

棱锥体压头的棱面和试样表面之间的夹角为攻角 (attack angle)，见图 11(a)。可用棱销前和面朝前两种方式进行划痕实验。见图 11(b)。两种方法的区别在于，在同样的正压力和划刻速度下，用面朝前划痕实验的沟槽深度和两边隆起的高度均大于棱销前划痕实验。

![图 11](image)

划痕能导致两种不同的失效方式：塑性变形和断裂。实验时，仪器提供线性增加的法向力，匀速移动试样，可以绘制法向力、表面粗糙度、划痕深度、残留划痕深度、切向力和摩擦系数随切向位移或时间连续变化的数据，见图 12。还能给出划痕的宽度和边缘凸起高度 (pile-up height) 等数据，见图 11(b) 和图 13。

3.3 划痕硬度的应用

纳米划痕硬度计不仅可以研究材料的摩擦磨损行为，还被广泛应用于薄膜的粘着失效 (adhesion failure) 和粘弹行为 [38]。对刻划材料来说，载荷和压入深度是重要的参数，而且残余划痕的深度、宽度、凸起的高度在研究接触压力和实际摩擦也是重要的参数。尤其是时间变化和温度影响的聚合物的粘弹松弛行为决定着划痕的残余形貌。目前，该类仪器已广泛应用于各种电子薄膜、汽车喷漆、胶卷、光学镜头、磁盘、化妆品 (指甲油和口红等) 的质量检测 (www.mts.com/nano; www.csem.ch/instrum; www.micromaterials.co.uk)。

应用一 绝缘膜具有存储静电的功能，已被广泛应用于印刷电路板、太阳能电池、触摸屏、各种芯片和传感器中。它们的力学性质受到广泛的关注。

在 100nm 厚硬质绝缘膜 / 钨基体上进行划痕实验，球形压头的半径为 2μm。图 12(a) 为划痕形貌的光学照片，图中框内的形貌表明划痕中的薄膜已剥离。图 12(b) 为图 12(a) 中薄膜基底失效微区的扫描电子显微照片。为划痕残余形貌。图 12(c) 为失效微区某划痕截面的原子力显微表面轮廓图。图 12 中，1，2 和 4 点的纵坐标分别低于 200nm，说明在压头划动过程中划痕两侧材料也分别下降。图 12(d) 为划痕长度 1mm 的测量结果，法向力 $F_N$ 从 0 mN 线性
连续增加到 10 mN，切向力 $FT$，压入深度 $PD$，残余压入深度 $RD$ 同在 $FN = 4.2$ mN（对应的划痕位置为 0.4 mm）时发生明显变化，说明薄膜开始剥落，此时的 $FN$ 定义为临界失效载荷。在 0.4~1.0 mm 之间，$FT$ 和 $PD$ 曲线的波动实时反映了薄膜剥离和压入钨基体的情况，$RD$ 反映了薄膜剥离和钨基体弹性恢复后的压痕深度（引自 www.csem.ch/instrum 提供的使用 Nano-Scratch Tester 的实验结果）。

应用二 喷漆是附着在汽车表面上的一层聚合物，对日常各种变化的温度、湿度和机械损伤，如雨冲和路面上各种微小颗粒的撞击等作用下，要求喷漆至少保持 5 年的高光泽度。划痕的残余深度与喷漆光泽度有密切关系，定量测量划痕横剖面（cross-profilometry）的工具是研究材料的划痕硬度和黏弹塑性松弛行为的重要手段 [37]。目前，该技术已成为评价汽车喷漆质量的重要方式 [38,39]。图 13 显示该技术评估汽车喷漆由于受时间和温度影响刻痕实验结果，反映了该漆随时间和温度变化的自修复能力。

4 结束语

目前，纳米硬度技术已广泛应用于薄膜、喷涂、材料表面改性等领域，并不断向微电子、集成电路电极系统、生物材料、医疗器械等领域延伸，已成为检测材料微小区域力学性能的有效手段。未来的几年内，人们希望能在如下三方面得到发展 [17]：

（1）在仪器方面，希望常规显微硬度计具有纳米硬度计的实时测量功能，提供经济的兼顾显示和纳米硬度计功能的新一代仪器，与新技术相结合，如用原子力显微镜观察纳米尺度的三维压痕和凸起形象，建立基于激光干涉的位移测量方法，提高测量精度和减少仪器柔性与热漂。
移对测量结果的影响。

(2) 在测量原理和分析方面，人们期待着测量和分析技术的新发展。基于连续刚度测量技术的新方法可能会拓宽测量高聚物黏弹性能的频率范围。有限元模拟将为材料力学性能数值测量和分析的有力工具，例如分析凸起现象、基底对薄膜的影响，从球形压痕实验数据中建立单轴应力-应变关系。

(3) 新的实用需求 超薄膜特性的研究，如检测磁盘存储器仅几个纳米厚保护膜的性能。在这种尺度范围内，由于表面污染和表面力将使接触现象变得异常复杂，需要建立新的测量和分析方法。

参考文献
2. 曲敬信. 测试技术. 北京: 化学工业出版社. 1998. 556~557
3. 中国大百科全书·力学卷. 北京: 中国大百科全书出版社. 1985. 549~551
DEVELOPMENTS AND APPLICATIONS OF NANO-HARDNESS TECHNIQUES*

Zhang Taihua  Yang Yemin

State Key Laboratory of Nonlinear Mechanics (LNM), Institute of Mechanics,
Chinese Academy of Sciences, Beijing 100080, China

Abstract  Great strides have been made over the past two decades in the development of techniques for studying the mechanical properties of materials on the micron and submicron levels. In this review paper, first, several definitions, classifications and their scopes used for hardness are introduced. Secondly, developments for nanoindentation techniques are summarized, especially the measuring principles of nanoindentation, CSM(Continuous Stiffness Measurement), high resolution load/displacement sensors, several typical tips, sample preparation, reference to the point of contact, experimental methods, instrument calibration and so on. With nanoindentation systems, quantitative data can be obtained: hardness, elastic modulus, fracture toughness, storage and loss modulus, stress exponent. Finally, advances in nanoscratch techniques are briefly discussed.

Keywords  nanoindentation, nanoscratch, micromechanics properties, hardness, modulus, fracture toughness, stress exponent

*The project supported by the National Natural Science Foundation of China (10102021, 10172086) and the Key Project from Chinese Academy of Sciences(KJCX2-SW-L2, KGCX1-11)