Review Article: Stress in thin films and coatings: Current status, challenges, and prospects

Grégory Abadias, Eric Chason, Jozef Keckes, et al.

COLLECTIONS

This paper was selected as Featured

ARTICLES YOU MAY BE INTERESTED IN

Intrinsic stress in sputtered thin films
Journal of Vacuum Science & Technology A 9, 2431 (1991); https://doi.org/10.1116/1.577295

Consistency and reproducibility in atomic layer deposition
Journal of Vacuum Science & Technology A 38, 020804 (2020); https://doi.org/10.1116/1.5140603

Influence of apparatus geometry and deposition conditions on the structure and topography of thick sputtered coatings
Journal of Vacuum Science and Technology 11, 666 (1974); https://doi.org/10.1116/1.1312732
Review Article: Stress in thin films and coatings: Current status, challenges, and prospects

Grégory Abadias
Département de Physique et Mécanique des Matériaux, Institut Pprime, UPR 3346, CNRS-Université de Poitiers-ENSMA, SP2MI, Téléport 2, F86962 Futuroscope-Chasseneuil, France

Eric Chason
School of Engineering, Brown University, Providence, Rhode Island 02912-9104

Jozef Keckes
Department of Materials Physics, Montanuniversität Leoben und Erich Schmid Institute for Materials Science, Austrian Academy of Sciences, Leoben 8700, Austria

Marco Sebastiani
Engineering Department, University of Rome “ROMA TRE,” Via della Vasca Navale 79, Rome 00147, Italy

Gregory B. Thompson
Department of Metallurgical and Materials Engineering, The University of Alabama, P.O. Box 870202, Tuscaloosa, Alabama 35487-0202

Etienne Barthel
Laboratoire Sciences et Ingénierie de la Matière Molle, ESPCI Paris, PSL University, Sorbonne Université, CNRS, F-75005 Paris, France

Gary L. Doll
Mechanical Engineering Department, The University of Akron, Akron, Ohio 44325-0406

Conal E. Murray
IBM T.J. Watson Research Center, Yorktown Heights, New York 10598

Chris H. Stoessel
Eastman Chemical Co., 3788 Fabian Way, Palo Alto, California 94303

Ludvik Martinu
Department of Engineering Physics, Polytechnique Montréal, Montreal, Quebec H3T 1J4, Canada

(Received 3 November 2017; accepted 9 February 2018; published 5 March 2018)

The issue of stress in thin films and functional coatings is a persistent problem in materials science and technology that has congregated many efforts, both from experimental and fundamental points of view, to get a better understanding of how to deal with, how to tailor, and how to manage stress in many areas of applications. With the miniaturization of device components, the quest for increasingly complex film architectures and multiphase systems and the continuous demands for enhanced performance, there is a need toward the reliable assessment of stress on a submicron scale from spatially resolved techniques. Also, the stress evolution during film and coating synthesis using physical vapor deposition (PVD), chemical vapor deposition, plasma enhanced chemical vapor deposition (PECVD), and related processes is the result of many interrelated factors and competing stress sources so that the task to provide a unified picture and a comprehensive model from the vast amount of stress data remains very challenging. This article summarizes the recent advances, challenges, and prospects of both fundamental and applied aspects of stress in thin films and engineering coatings and systems, based on recent achievements.

---

a)Author to whom correspondence should be addressed; electronic mail:
gregory.abadias@univ-poitiers.fr

b)Electronic mail: eric_chason_phd@brown.edu
c)Electronic mail: jozef.keckes@gmail.com
d)Electronic mail: marco.sebastiani@uniroma3.it
e)Electronic mail: gthompson@eng.ua.edu
f)Electronic mail: etienne.barthel@espci.fr
g)Electronic mail: gd27@uakron.edu
h)Electronic mail: conal@us.ibm.com
i)Electronic mail: cstoessel@eastman.com
j)Electronic mail: ludvik.martini@polymtl.ca
I. INTRODUCTION

The presence of stress in thin films and functional coatings constitutes a major concern in many technological applications as excessive residual stress levels can dramatically affect the performance, reliability, and durability of material components and devices. Worst case scenarios lead to film cracking for layers subjected to tensile stress, or peeling off, buckling, or blistering in the case of compressive stress. Residual stress distributions can significantly impact the adhesion and the fracture toughness of thin films, the ductility of bulk metallic glasses (BMGs), the performance of optoelectronic and aerospace components, the thermo-mechanical behavior of stacks in through silicon via (TSV) 3D integrated devices, and the resonant frequency and lifetime of micro- and nanoelectromechanical systems (MEMSs and NEMSs) or cause the emergence of hillocks and whiskers in metal interconnects at the origin of short circuit failures, to cite a few examples. Alternately, stress (or strain) can have a beneficial influence on the physical properties of thin layers and nanostructures, e.g., conductivity, piezoelectricity, magnetic anisotropy, and magneto-elastic coupling or enhancement in charge carrier mobility in silicon-based semiconductor technology. Therefore, there is significant motivation to understand the origin of stress in thin films as it can directly affect the design, processing, and lifetime of advanced materials and components. Various research strategies to tailor and control the stress state are currently devised in diverse applications ranging from micro/optoelectronic devices, MEMS/NEMS (thermal sensors and actuators), optical components (lenses, mirrors, filters, etc.) to protective and functional coatings intended to impart thermal, mechanical, tribological, environmental, electrical, magnetic, or biological functions. From a technological point of view, this has generated an intense research and innovation activity over the last few decades related to the measurement of residual stresses, more recently, including the determination of intrinsic (growth) stress due to thin film deposition conditions, which have gained interest thanks to the potential offered by in situ and real-time diagnostics.

Regardless of the technique of film preparation [physical or chemical vapor deposition (PVD/CVD), electrochemical deposition, etc.], thin films can develop large intrinsic stresses in the course of their growth process. For many materials, the stress magnitude in thin layer forms and nanostructures can typically exceed the tensile strength of their respective bulk counterpart. Indeed, the density of defects that are frozen into a film during deposition can be 2 orders of magnitude higher than that produced by the severest cold-work treatment of a bulk material.

The concern about stress dates back over a century ago with the observation by Gore that “the inner and outer surfaces of electrodiposits were in unequal states of cohesive tension,” causing the plate on which they were deposited to bend. This led Stoney, some years later, to derive a relationship between the film stress and the amount of substrate bending, which can be readily quantified from the measurement of the substrate curvature, or equivalently, its deflection. Since the pioneering work of Stoney in the early 1900s, considerable work has been done, especially after the 1960s with the achievement of high vacuum conditions during the PVD processes, enabling one to separate various parameters of film growth which influence the intrinsic film stress while minimizing the influence of the impurity content due to chemical reactivity with residual gas components. Despite the fact that many trends have been identified, in particular the common stress behaviors depending on material mobility, the stress development during thin film growth exhibits complex dependences with respect to the growth rate, temperature, film microstructure, and morphology so that a complete and unifying picture of our understanding of stress has not yet been achieved. The situation gets even more intricate in the case of multicomponent systems, where alloying effects such as solute atom segregation or phase transformation may come into play. Deposition processes in which the temporal and energetic profiles in terms of energy and fluxes of specific species (such as ions) can be manipulated are opening new avenues for stress...
tailoring. This explains why stress is a hot topic amongst the materials science and engineering community, calling for more systematic testing grounds as well as new achievements gained from multiscale stress modeling.

From a more practical point of view, the reliable assessment of residual stresses on a micron, or submicron, scale is a strategic challenge for the robust design and reliability of a wide range of micro- and nanosystems. Protective thin films and coatings often possess complex gradients of phases, microstructure, and residual stresses, which result from (1) intentionally varying deposition conditions, (2) self-organization phenomena like competitive grain growth and diffusion along grain boundaries (GBs) and/or on the surface, (3) postdeposition mechanical and thermal loads caused, for example, by friction between coating and machined surfaces, and (4) the effect of the surrounding environment leading to gas or vapor sorption. These gradients decisively influence the functional properties of thin films and coatings, such as hardness, toughness, oxidation resistance, wear behavior, adhesion, and durability. Therefore, in order to optimize the functional properties, it is necessary to assess the depth variation of residual stresses with respect to the microstructure with a nanoscale resolution. The evolution of the average in-plane stress can be measured by wafer curvature techniques during deposition, providing information on the stress profile throughout the film thickness.23 In recent years, several high-resolution residual stress measurement techniques have been implemented, such as microfocus synchrotron x-ray diffraction (XRD),46 holographic transmission electron microscopy (TEM),47 micro-Raman spectroscopy,48 and electron backscattered diffraction.49 Unfortunately, most of these methods have limitations in terms of instrument accessibility, applicability to amorphous (or strongly textured) materials, perfect lattice rectification, and nonhomogeneous stress states. Nonetheless, the measurement of residual stress in submicron volumes still poses a challenging task, especially in the case of strongly textured, complex multiphase, nanocrystalline, or amorphous materials and films.

This article, which gathers contributions from several leading groups working in the field, encompasses a critical viewpoint on the current state of knowledge on stress in thin films and coatings. It gives a nonexhaustive overview and a snapshot on the progress of both fundamental and applied research in this field, based on recent outcomes presented during the Stress Workshop entitled “Stress Evolution in Thin Films and Coatings: from Fundamental Understanding to Control,” which was jointly organized by the Advanced Surface Engineering Division of the American Vacuum Society and the Society of Vacuum Coaters, in Chicago, in October 2016. It is mostly focused on polycrystalline thin films so that the issue of strained surfaces and nanostructures, as well as epitaxial layers, will not be discussed here. The reader interested in such aspects is referred to the articles/monographs of Koch,50 Ibach,51 Brovko et al.,52 and Fluri et al.53 Similarly, for more extensive reviews of this work, particularly on the proposed stress models, it is advisable to go through the recent tutorial of Chason and Guduru.23 Still, stress in amorphous thin films represents an important part of technological interest especially in relation to optical coatings (OCs) and thin films on plastics.51,54–56

The present paper is divided into several parts. Sections II and III are dedicated to the fundamental aspects of how thin film stress can be measured and interpreted. In Sec. II, we outline the evaluation methods for stress determination, including wafer curvature, XRD, and focused ion beam (FIB)-based techniques. Section III describes recent experimental findings on stress evolution during thin film growth and their understanding based on a kinetic model, recently extended to account for energetic particle bombardment during deposition (Sec. III A). Selected examples are also provided to highlight chemical alloying effects and phase transformation (Sec. III B), as well as the interplay between the stress gradient and texture development (Sec. III C). Section IV addresses the ways how to deal with and how to tailor and manage stress in specific areas of applications. These include the methods and approaches to assess and control stress in microfabricated systems such as microelectronic components (Sec. IV A) or issues related to the effect of deposition processes and film growth conditions in the context of most frequent amorphous and polycrystalline OCs and optical interference filters (OIFs), as well as strategies to mitigate excessive stress by both experimental and design approaches (Sec. IV B). Section IV C gives an overview of problems and solutions related to stress in the context of inorganic coatings on plastic substrates, including polymer webs and injection molded components, while Sec. IV D describes case studies related to stress build-up and stress compensation approaches in the context of protective coatings for aeronautical, aerospace, and related manufacturing components. Complementary stress engineering strategies, based on interfacial and alloying design or pulse management in sputtering processes, are proposed in Sec. IV E, followed by an overview of the present status of understanding and interpreting the formation of cracks and their propagation in relation to stress (Sec. IV F). Finally, Sec. V summarizes the still open questions related to stress, and it suggests pathways and outlooks for future developments.

II. EVALUATION METHODS FOR STRESS DETERMINATION

A. Nondestructive method: Wafer curvature

An important method for measuring stress in thin films is the wafer curvature. The principle behind the method is to measure the curvature induced in the substrate due to stress in the film. Because it is nondestructive and can be used in real-time, it has been one of the workhorse techniques for quantifying stress in thin films.23,33,57,58 In this section, we describe how the curvature is related to the film stress and make clear some underlying assumptions used in interpreting the measurements. We will also discuss some methods that have been developed for measuring stress using the wafer curvature.
The well-known Stoney equation relates the average stress in the film \( (\bar{\sigma}) \) to the measured curvature \( (\kappa) \),
\[
\kappa = \frac{6 \bar{\sigma} h_f}{M_s h_s^2}.
\] (1)

The values \( M_s \) and \( h_s \) refer to the biaxial modulus and the thickness of the substrate, respectively. The product \( \bar{\sigma} h_f \) is termed the stress-thickness and has units of force/length; note that the average stress cannot be determined from the curvature without knowing the film thickness \( (h_f) \). The average stress in the Stoney equation is assumed to be equi-biaxial and laterally uniform. This ignores, for instance, that in polycrystalline films, the stress may be different at/near the GBs from that in the middle of the grain\(^{23,60}\) or that in patterned films (such as metallic interconnect lines), the stress may be nonuniform over the surface.\(^{61}\)

Even assuming that the stress is laterally uniform, it may still vary through the thickness of the film. Then, the thickness-averaged stress can be computed by integrating the in-plane stress at height \( z \) from the substrate, \( \sigma_{xx}(z) \), over the film thickness \( (h_f) \),
\[
\bar{\sigma} = \frac{1}{h_f} \int_0^{h_f} \sigma_{xx}(z) dz.
\] (2)

In this discussion, we are ignoring the presence of any surface or interfacial stresses that may also contribute to the curvature.\(^{62-64}\) Because of the thickness averaging, a single curvature measurement does not provide any information about the depth-dependence of the stress in the film. The curvature from a film with a nonuniform stress through the thickness is equivalent to a film of the same thickness with a uniform stress of the average value. A film that induces no curvature in the substrate may in fact have a large stress gradient in it. The lack of curvature only shows that the tensile and compressive stresses integrated over the thickness are equal. Since large stress gradients may lead to cracking even though the nominal stress appears to be small, it is important to know the distribution of stress throughout the film’s thickness.

The depth-dependence of the stress can be explored by measuring the evolution of the stress-thickness as the film is deposited. The time derivative of the stress-thickness is given by
\[
\frac{d(\bar{\sigma} h)}{dt} = \sigma_{xx}(h) \frac{dh}{dt} + \int_0^h \frac{\partial \sigma_{xx}(z)}{\partial t} dz,
\] (3)

where \( h \) is the thickness at time \( t \). The first term on the right corresponds to the effect of adding new layers to the surface with a stress of \( \sigma_{xx}(h) \); this is sometimes referred to as the incremental or instantaneous stress. The second term corresponds to a change in the stress of the layers that have already been deposited.

If we can assume that the stress does not change after deposition [i.e., the time derivative in the second term in Eq. (3) is zero], then the stress at each height in the film is the same as when it was deposited. In that case, the incremental stress can be determined from the derivative of the stress-thickness with thickness
\[
\sigma_{xx}(h) = \frac{d(\bar{\sigma} h)}{dh} = \frac{d(\bar{\sigma} h)}{dh}.
\] (4)

To illustrate how this works, a schematic representation of a film of thickness \( h_f \) with a distribution of in-plane stress is given in Fig. 1(a). The arrows represent the stress at different heights in the film. The arrows pointing away from the film correspond to tensile stress, and those pointing toward the film correspond to compressive stress. The corresponding evolution of the curvature during deposition is shown in Fig. 1(b) as a function of thickness with the value \( h_f \) shown by the vertical dotted line. The average film stress at this thickness is determined by dividing the measured curvature by the thickness, represented by the line between the measured value and the origin. The incremental stress at this thickness is represented by the slope of the stress-thickness at this point, represented by the tangent line on the curve. Note that at this thickness, the average stress is positive (tensile), while the incremental stress is negative (compressive). The data are taken from stress-thickness measurements during electron beam evaporation (EBE) of polycrystalline Ag on SiO\(_2\).

Recall that in equating the incremental stress with the slope of the stress-thickness, we assumed that the stress in the deposited layers does not change. The effect of the stress change in the deposited layers can be seen when the growth is stopped at a thickness of 100 nm indicated by the vertical line [Fig. 1(c)]. Since there is no more growth, the thickness of the film does not change after this point. However, it is
clear that the stress-thickness relaxes significantly. As discussed below, this may be due to various stress-induced processes like diffusion of atoms out of the GB, thermal expansion mismatch, or grain growth in the film.

Multiple techniques have been developed for measuring the wafer curvature. Some methods monitor the change in the shape of the cantilever by capacitance,66–68 microbalance,69 dilatometry, or interferometry.70 Others monitor the deflection of light beams reflected from the cantilever surface. Some monitor a single reflected beam71 or the spacing between multiple beams reflected from the surface.72,73,74 A benefit of the multibeam reflection method is that it reduces the sensitivity to sample vibration because the curvature is determined from the change in spacing between the reflected beams. Vibrations of the sample may change the direction of the reflected beams, but it changes them all by the same amount so that the relative spacing between the beams does not change. A multiple-beam optical stress sensor (MOSS) is therefore useful for in situ measurements in growth chambers where vacuum pumps may induce a large amount of mechanical vibration, as well as in liquid media, such as during electrochemical deposition.74–77,79 where convection (natural or forced) may induce perturbations.

B. X-ray diffraction techniques

1. Conventional laboratory XRD

Conventional laboratory XRD analysis performed in reflection geometry is routinely used to characterize average residual stresses and stress gradients in thin films, coatings, and near-surface regions.76–80 Residual stress is determined from the measurement of x-ray elastic strains, ε, and applying a constitutive equation in the form of generalized Hooke’s law, which requires the knowledge of the elastic properties of the film.79

The common procedure employs the so-called “sin²ψ method,” based on the determination of a set of lattice spacings dhkl using Bragg’s law from measured Bragg’s angle positions ψhkl of hkl reflections at various sample tilt angles ψ and azimuth angles φ, whereby ψ represents the angle between the diffraction vector Q and the sample normal z and φ is the rotation angle around z. In many situations, polycrystalline films exhibit a preferred crystallite orientation of rotational symmetry around the substrate normal, referred to as a fiber texture, so that it is sufficient to measure the strain solely as a function of the tilt angle ψ. In the absence of shear stress components (σij = 0 for i ≠ j) and assuming an equi-biaxial in-plane stress state (σ11 = σ22 = σ0), the quantity σ0 can be readily extracted from the slope of dhkl vs sin²ψ curve, according to the fundamental equation of x-ray residual stress analysis

\[ \varepsilon_{hkl} = \frac{d_{hkl} - d_0}{d_0} = 2S_{11}^e \sigma_1 + \frac{1}{2}S_{22}^e \sigma_1 \sin^2 \psi, \]  

where \( S_{11}^e \) and \( (1/2)S_{22}^e \) refer to the x-ray elastic constants (XECs) and \( d_0 \) is the unstrained lattice parameter. XEC quantities are introduced in the formalism to account for the elastic anisotropy of crystalline materials. For randomly textured materials, they are independent of ψ (and φ) and can be calculated for each hkl reflection from the single crystal elastic compliances \( s_{ij} \) using a grain interaction model.77,79,81

Indeed, in a diffraction experiment, only a given subset of crystallites contribute to the measured intensity so that a micromechanical model relating the strain (or stress) inside these crystallites to the macroscopic average stress to which the whole specimen is subjected is required. The most common approximations are the Reuss (all crystallites are subjected to the same stress),82 Voigt (all crystallites exhibit the same strain),83 and Neerfeld-Hill limits,84,85 with the last being the arithmetic mean of the Reuss and Voigt estimates.77

Note that for isotropic materials, the XECs reduce to the mechanical elastic constants, \( S_{11}^{mech} \) and \( S_{22}^{mech} \), defined as \( S_{11}^{mech} = -(1/4)E \) and \( (1/2)S_{22}^{mech} = (1 + \nu)/E \), where \( E \) and \( \nu \) are the Young’s modulus and Poisson ratio of the deposit, respectively.

The value of the unstrained lattice parameter \( d_0 \) can be obtained from the strain-free direction \( \psi^0 \), derived by stating \( \varepsilon_{hkl} = 0 \) from the fundamental Eq. (5), which yields \( \sin \psi^0 = \sqrt{2\nu/(1 + \nu)} \) for the case of isotropic materials in an equi-biaxial stress state. In practice, when the elastic constants of the deposit material are unknown (which is rather often the case for multicomponent alloys or complex compounds), the strain-free lattice parameter may be obtained from the intersection point of \( d_{hkl} \) vs \( \sin^2 \psi \) lines obtained for the same deposit in different stress states.86,87

However, it is important to note that the lattice spacing \( d_{hkl}^{psi} \) represents an average quantity, which depends on the actual stress depth gradient, \( \sigma(z) \). X-ray penetration depth, τ, chemical gradients influencing \( d_{hkl}^{psi} \), and XRD experiment geometry. In practice, different approaches can be implemented depending on the laboratory diffraction equipment. On a two-circle diffractometer, a simple procedure consists of measuring a series of 20 scans around a single hkl reflection at varying incident beam angles ω. In this asymmetric geometry, known as the Ω mode,79,88 the angle ψ is given by ψ = ω − θ. A second option is to use a glancing angle scan (at fixed ω) and determine the \( d_{hkl}^{psi} \) lattice spacing of various hkl reflections which correspond to different ψ values. This approach is well suited for thick polycrystalline coatings with random orientation of the crystallites, allowing for the selection of hkl reflections at high 2θ values with appreciable intensity, contributing to a higher accuracy in the measurement of \( d_{hkl}^{psi} \) values (misalignment issues of the diffractometer are minimized at high scattering angles). However, for thin films, Bragg reflections at high scattering angles are most often hardly measurable.79 A third option, for researchers having a four-circle diffractometer, is to measure a series of symmetric θ–2θ scans at various inclination ψ angles of the sample holder around the diffraction axis (known as the Ω mode,79,88 or side-inclination method76), allowing for a change in the angle ψ independent of the detector rotation. This geometry is the same as the one used for pole figure measurements and has one great virtue that the accessible tilt angle range is much larger than that in the Ω mode (for which |ψ| < 90°), starting from ψ = 0° to almost ψ = 90°. Moreover, the penetration depth remains nearly constant for a wide range of tilt angles (up to ψ ~ 30°–40°).
One should keep in mind, however, that the analysis outlined above is only valid for polycrystalline thin films/ coatings having a random orientation of grains. Thin films/coatings produced by PVD/CVD techniques rarely belong to this category, as they most often exhibit a crystallographic texture related to their specific columnar growth morphology so that they can no longer be treated as macroscopically elastically isotropic specimens. In this case, the stress analysis is more complex, and one has to use the concept of x-ray stress factors to relate stress and strain. Also, the available ψ directions which yield sufficient diffracted intensity are often restricted to small angular ranges near the intensity poles. However, rather simple expressions can still be derived for the most common fiber textures, \( \langle 100 \rangle \), \( \langle 110 \rangle \), \( \langle 111 \rangle \), and \( \langle 211 \rangle \), when considering crystals with cubic symmetry. The plots of the strain \( \epsilon_{\psi}^{hkl} \) in the distinct poles vs \( \sin^2 \psi \) do not generally fall on a straight line, even for a biaxial stress state. The only exceptions are reflections of types \( h00 \) and \( hhh \), as well as the \( \langle 111 \rangle \) texture.

Another approach is offered for the case of thin films with strong and sharp texture (as also found in epitaxial layers or cold-rolled materials), known as the crystallite group method. This method was introduced in 1982 by Willemse et al. and then adapted by Baron and Hauk to fiber-textured coatings. It is based on considering the highly textured film with a crystallographic growth texture along a given \( \langle uvw \rangle \) direction as a single crystal aligned along this corresponding ideal orientation and measuring the lattice spacings of various \( hkl \) planes in the same set of grains belonging to this \( \langle uvw \rangle \) ideal orientation (crystallite group), at specific angles \( \psi \) corresponding to the intensity poles. General expressions of the elastic strain \( \epsilon_{\psi}^{hkl} \) vs \( \sin^2 \psi \) have been reported by Clemens and Bain for materials with cubic symmetry and for equal or nonequal biaxial stress states. This methodology is a special case of the \( \sin^2 \psi \) technique discussed above; it differs by the fact that the texture in the film is explicitly taken into account rather than using XECs or stress factors to describe geometrical distributions of crystallites in the polycrystalline aggregate. Kim et al. have employed a similar procedure to determine the biaxial stress state in sputter-deposited AlCu films with a \( \langle 111 \rangle \) preferred orientation.

In addition to anisotropy issues in textured thin films, another possible cause for deviation from the linearity of \( \epsilon_{\psi}^{hkl} \) vs \( \sin^2 \psi \) plots is the presence of stress gradients along the film thickness, especially when the film thickness is on the same order as the penetration depth \( \tau \) at \( \psi = 0 \). Different concepts and methods have been proposed to determine the residual stress gradient \( \sigma(z) \) along the surface normal \( z \) in equibiaxially stressed polycrystalline thin films and coatings. Neglecting the chemical gradients, the measured x-ray elastic strain can be expressed as:

\[
\epsilon_{\psi}^{hkl}(\tau) = \frac{d\epsilon_{\psi}^{hkl}(\tau)}{d\epsilon_{\psi}^{hkl}} = \frac{1}{d\epsilon_{\psi}^{hkl}} \int_{0}^{\tau} \epsilon_{\psi}^{hkl}(z) \exp\left(-z/\tau\right) dz, \tag{6}
\]

where \( \zeta \) is the thin film depth and \( d\epsilon_{\psi}^{hkl}(z) \) is the depth-dependent measured lattice spacing. By varying the experiment geometry, e.g., by changing the angle \( \psi \), it is possible to determine x-ray elastic strains \( \epsilon_{\psi}^{hkl}(\tau) \) for various x-ray penetration depths \( \tau \). This information is then used to calculate the unknown residual stress depth profile \( \sigma(z) \), defined in the real space as a function of \( z \), by fitting its supposed usually polynomial dependence to the experimental \( \epsilon_{\psi}^{hkl}(\tau) \) dependence [Eq. (6)], expressed in the so-called Laplace space, by applying x-ray elastic constants. Since there are, however, infinitely many \( \sigma(z) \) dependencies, which can be fitted to the measured \( \epsilon_{\psi}^{hkl}(\tau) \) dependence, the inverse Laplace space approach can be used to evaluate residual stress gradients \( \sigma(z) \) only in simplified cases like shot-peened samples with linear and/or monotonic \( \sigma(z) \) dependencies. An illustration will be given in Sec. IV A for stress gradients in capped layers. In many other cases, \( \sigma(z) \) dependencies can be very complex, like stepwise or oscillatory, and can be even combined with the presence of chemical gradients in the sample, smearing the measured \( d\epsilon_{\psi}^{hkl} \) values. Therefore, there is a need to develop novel characterization approaches, which can be used to assess nanoscale \( \sigma(z) \) dependencies in gradient thin films and coatings.

### 2. Cross-sectional nanodiffraction using synchrotron facilities

In 2012, a novel experimental approach to characterize the nanoscale depth variation of residual stresses and the microstructure in thin films and coatings was introduced. Cross-sectional x-ray nanodiffraction is based on the application of synchrotron point (or pencil) x-ray nanobeams with a diameter (or thickness) down to 50 nm or even less to scan thin films at the cross-section in transmission (or reflection) diffraction geometries (Fig. 2).

Up to now, cross-sectional x-ray nanodiffraction experiments were performed at beamlines ID11 and ID13 of European Synchrotron Radiation Facility in Grenoble and at the beamline P03 of Petra III synchrotron source in Hamburg using monochromatic beams of energy \( E = 12–30 \text{ keV} \). For nanobeam focusing, Fresnel zone plates, nanofocusing parabolic refractive x-ray lenses, and multilayer Laue lenses were used. For the experiments, a thin lamella consisting of thin films or coating on the substrate with a thickness \( e \) in the beam direction in the range of \( \sim 10–250 \mu \text{m} \) is usually prepared using the FIB technique. It is important that the beam is aligned parallel to the substrate surface or to the interfaces between individual sublayers in the lamella by using the \( \phi \) rotation axis (Fig. 2). The diffraction data from the sample are collected using a charge-coupled device (CCD) area detector positioned behind the sample using a sample-detector distance of \( \sim 80–130 \text{ mm} \). In order to obtain diffraction data from different sample cross-sectional regions, the lamella is moved along the sample normal \( z \) with a step of the beam size or smaller. For each lamella position, two-dimensional (2D) diffraction data are collected using CCD. The diffraction data are then processed in order to obtain information about the position, shape, and nature of Debye–Scherrer (D-S) rings.
collected by the detector, e.g., using a FIT2D or similar software packages.

Every hkl Debye–Scherrer ring azimuthal position \( \delta \) represents the diffraction on (hkl) crystallographic planes oriented with their normal vector \( \mathbf{n}_{hkl} \) parallel to the diffraction vector \( \mathbf{Q}_{hkl} \) (Fig. 2). Therefore, Debye–Scherrer rings collected using the 2D detector at the thin film depth \( z \) can be used to evaluate the lattice spacing \( d_{hkl}^{\text{bkl}}(z) \) as a function of the ring azimuthal angle \( \delta \) using Bragg’s law by analyzing Bragg’s angle \( 2\theta_{hkl}^{\text{bkl}}(z) \) azimuthal dependencies on the detector. Depending on the diffraction statistics, usually 36 values of \( d_{hkl}^{\text{bkl}}(z) \) can be determined for 36 azimuthal angle \( \delta \) sections called azimuthal rings. Every \( d_{hkl}^{\text{bkl}}(z) \) value represents an x-ray probe volume-averaged lattice parameter for the diffraction vector \( \mathbf{Q}_{hkl} \) orientation defined by the angles \( \delta \) and \( \theta \) in Fig. 2. For every \( d_{hkl}^{\text{bkl}}(z) \), x-ray elastic strain at the thin film depth \( z \) can be determined as follows:

\[
v_{hkl}^{\text{bkl}}(z) = \frac{d_{hkl}^{\text{bkl}}(z) - d_{hkl}^{\text{bkl}}(0)}{d_{hkl}^{\text{bkl}}(0)}. \tag{7}
\]

The measured strain \( v_{hkl}^{\text{bkl}}(z) \) can be expressed as a function of unknown strain components \( v_{ij}^{\text{bkl}}(z) \) defined in the sample coordinate system with axes \( x, y, \) and \( z \) from Fig. 2 as follows:

\[
v_{hkl}^{\text{bkl}}(z) = \sin^2 \theta v_{11}^{\text{bkl}}(z) + \cos^2 \theta \sin^2 \delta v_{22}^{\text{bkl}}(z) + \cos^2 \theta \sin^2 \delta v_{33}^{\text{bkl}}(z)
+ \cos^2 \theta \cos^2 \delta v_{12}^{\text{bkl}}(z) - \sin 2\theta \cos \delta v_{31}^{\text{bkl}}(z)
+ \cos^2 \theta \sin 2\delta v_{32}^{\text{bkl}}(z) - \sin 2\theta \sin \delta v_{21}^{\text{bkl}}(z). \tag{8}
\]

By considering thin film x-ray elastic constants \((1/2)\varepsilon_{hkl}^{\text{bkl}}(z)\) and \(S_{hkl}^{\text{bkl}}(z)\), which depend on single crystal elastic constants, crystallographic texture, hkl reflection, and grain interaction mechanism, it is possible to write x-ray diffraction Hooke’s law as

\[
v_{hkl}^{\text{bkl}}(z) = S_{hkl}^{\text{bkl}}(z)[\sigma_{11}(z) + \sigma_{22}(z) + \sigma_{33}(z)]
+ \frac{1}{2} \varepsilon_{hkl}^{\text{bkl}}(z) \left[ \sin^2 \theta \sigma_{11}(z) + \cos^2 \theta \sin^2 \delta \sigma_{22}(z)
+ \cos^2 \theta \cos^2 \delta \sigma_{33}(z) \right]
+ \frac{1}{2} \varepsilon_{hkl}^{\text{bkl}}(z) \left[ -\sin 2\theta \cos \delta \sigma_{31}(z)
+ \sin 2\delta \cos^2 \theta \sigma_{32}(z) - \sin 2\theta \sin \delta \sigma_{21}(z) \right], \tag{9}
\]

where \( \sigma_{ij}(z) \) represents unknown stress components defined in the sample coordinate system. Usually, the stress state in thin films is equi-biaxial with \( \sigma_{11}(z) = \sigma_{22}(z) = \sigma(z) \), and the out-of-plane stress component as well as shear stress components can be neglected with \( \sigma_{33}(z) \approx 0 \) and \( \sigma_{ij}(z) \approx 0 \). Consequently, Eq. (9) can be simplified

\[
v_{hkl}^{\text{bkl}}(z) = 2\sigma(z)S_{hkl}^{\text{bkl}}(z) + \sigma(z) \frac{1}{2} \varepsilon_{hkl}^{\text{bkl}}(z) \left[ \sin^2 \theta + \cos^2 \theta \sin^2 \delta \right]. \tag{10}
\]

Since for small Bragg’s angles \( \theta \), the terms \( \sin^2 \theta \) and \( \cos^2 \theta \) in Eq. (10) go to zero and one, respectively, Eqs. (7) and (10) can be rewritten as

\[
\frac{\partial d_{hkl}^{\text{bkl}}(z)}{\partial \sin^2 \delta} \approx \frac{\sigma(z)}{2} \frac{\partial S_{hkl}^{\text{bkl}}(z)}{\partial \sigma_{hkl}^{\text{bkl}}(z)}. \tag{11}
\]

It means that for every hkl Debye–Scherer ring and the corresponding \( d_{hkl}^{\text{bkl}}(z) \) dependence, the in-plane residual stress is proportional to the Debye–Scherer ring ellipticity quantitatively expressed through \( \partial d_{hkl}^{\text{bkl}}(z)/\partial \sin^2 \delta \).

In practical cases, \( d_{hkl}^{\text{bkl}}(z) \) values are evaluated for 36 azimuthal positions \( \delta \) of the hkl Debye–Scherer rings, which are then plotted as a function of \( \sin^2 \delta \). The slope of the \( d_{hkl}^{\text{bkl}}(z) \) vs \( \sin^2 \delta \) dependence is proportional to the magnitude of the in-plane stress \( \sigma(z) \) [Eq. (11)]. An important advantage of this approach is that the unstressed lattice parameter \( d_{o}^{\text{bkl}}(z) \) has to be determined only approximately.

In the case of a thin lamella with the lamella thickness \( e \) comparable to or even smaller than thin film or coating thickness \( \xi \) (Fig. 2), the lamella preparation results in the relaxation of the stress component \( \sigma_{11}(z) \), which influences the measured \( v_{hkl}^{\text{bkl}}(z) \) as well as the evaluated stress \( \sigma(z) \). In that case, it is necessary to perform a finite element (FE) modeling of the stress state in the thin lamellae in order to recalculate the stress state in the unstressed sample as extensively discussed.
In a majority of cases, it is possible to prepare lamellae with ε ≫ ξ, and the FE analysis is not required.

Another advantage of cross-sectional x-ray nanodiffraction is the fact that besides the possibility to evaluate residual stress gradients, thin film and coating depth gradients of phases and microstructures (including texture and crystallite size) can also be determined directly in real space as a function of the coating depth z. This can be done by evaluating and comparing individual Debye–Scherrer rings collected at different depths (Fig. 2). The approach opens the possibility to analyze residual stress gradients in graded thin films with chemical gradients and correlate them with depth gradients of texture, crystallite sizes, and phases.

C. FIB-based methods

1. Method description and validation

In recent years, the development of FIB coupled with scanning electron microscopy (SEM) and digital image correlation (DIC) created new possibilities for the downscaling to the submicron scale of material removal residual stress measurement techniques (e.g., hole drilling, ring-core, and slitting methods). The novel FIB techniques involve a combination of FIB milling, SEM imaging, DIC analysis, and FE modeling. The method consists of incremental FIB milling, by using several possible milling geometries, combined with in situ high-resolution field emission gun SEM imaging of the relaxing surface, and a full field strain analysis by DIC. The through-thickness profile of the residual stress can also be obtained by comparing the experimentally measured surface strain with FE and/or analytical constitutive modeling.

FIB-DIC procedures have recently been established as the one method capable of achieving submicron spatial resolutions for the residual stress analysis in amorphous materials, coatings, and thin nanostructured layers. Kang et al. first proposed the downscaling of the slitting method measuring the stress release caused by milling of a slot on a thin diamondlike-carbon (DLC) film deposited on a glass substrate. Several other milling relaxation geometries have been proposed in the literature, including hole-drilling, ring-core, H-bar (also called the double slot), and four-slot geometries.

The ring-core geometry has recently received particular attention because of its main advantages of highest spatial resolution (<1 μm), the possibility of full 2D strain and stress mapping, and the fact that complete stress relaxation is achieved if the trench depth h is higher than the central stub’s diameter (d_m), thus allowing for the direct use of Hooke’s law to calculate the average residual stress from the measured strain relief. Additionally, the use of a ring (or pillar) milling geometry could even be used to measure the fracture toughness of the film, by using the nanoindentation pillar splitting method, which involves sharp nanoindentation over the pillar to induce fracture of the material at a certain critical load that can be correlated with the fracture toughness, by knowing the pillar’s radius. Since this section is focused on the use of the ring-core geometry to measure residual stress, we refer to previous papers for more details of the possible use of the same geometry to measure fracture toughness.

In Fig. 3, a schematic presenting the main steps of the method and examples of the possible typical dimensions of the core is reported. More details on SEM imaging strategies and DIC procedures are reported elsewhere.
Significant activities in terms of validation of the method have recently been performed, including comparisons with other standard measurement techniques (Fig. 4). In particular, the procedure has been validated against conventional XRD \((\sin^2 \psi)\), curvature (Stoney’s equation), and micro-Raman methods. In a recent paper, a critical comparison between the XRD-\(\sin^2 \psi\) and the FIB-DIC methods has been performed on a fully characterized CrN coating grown by cathodic arc deposition (CAD). While a very good agreement is found between the two measurements, the observed differences have been critically discussed. In particular, the uncertainty on the elastic modulus of the coating, which is needed for stress calculations in both cases, can be the primary explanation for the observed differences. Additionally, the presence of a texture in the film can be a source of stress calculation errors in both cases since it may affect both the linearity of the stress vs \(\sin^2 \psi\) fit and the relaxation strain distribution after FIB milling. The presence of a stress depth profile can also be considered as a possible source of discrepancy between the two approaches since this is not considered in the most conventional XRD-\(\sin^2 \psi\) method. In relation to the latter point, another main source of discrepancy could be the different probing volumes between the two techniques, which is usually equal to the entire film thickness in the case of FIB-DIC, while it is a function of the beam energy and angle of incidence in the case of XRD methods.

2. Choice of the proper FIB milling geometry

The choice of an appropriate milling geometry allows for a quantitative and detailed evaluation of the full in-plane stress tensor, the depth profile of residual stress with a submicrometer resolution, the stresses inside single grains with different crystal orientation, or even Poisson’s ratio of the material.

The selected geometry should be in compliance with the material’s requirements and the object of the measurement. From this point of view, the method offers multiple possibilities that may be used for gaining more comprehensive information about the residual stress distribution over the sample surface and in-depth. Table I reports a series of suggestions regarding selection of the milling geometry for different applications.

3. Residual stress depth profiling

The FIB-DIC method can also be used for residual stress depth profiling by performing the milling process in a
Table I. Summary with practical recommendations on the correct choice of the best milling geometry for different applications.

<table>
<thead>
<tr>
<th>Milling geometry</th>
<th>Characteristics</th>
<th>Suggested application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-slot</td>
<td>Higher displacement, for a given stress, over a large area. Strong strain gradient from the edge Possibility of stress depth profiling</td>
<td>Best choice in the case of very thin films and stress profiling on them</td>
</tr>
<tr>
<td>Double-slot</td>
<td>Uniform strain (linear displacement) in the area between the two slots. Fast and simple analysis. Information only from one direction. Good choice for homogeneous coatings with equal-biaxial stress</td>
<td>Homogeneous coatings with a thickness range of 1–10 µm, where equal-biaxial stress is expected</td>
</tr>
<tr>
<td>Hole drilling</td>
<td>Strong strain gradient from the edge. Possibility of stress depth profiling</td>
<td>Good option for depth profiling since it is a very well established procedure (from macroscale)</td>
</tr>
<tr>
<td>Ring-core</td>
<td>Highest spatial resolution (&lt;1 µm). Full 2D stress analysis. Possibility of stress depth profiling. It is the one geometry that gives surface full stress relaxation: analytical calculation of stress is possible for ( h/d_m &gt; 1 )</td>
<td>Coatings with a thickness range of 0.2–20 µm Best choice for residual stress mapping over heterogeneous and/or multiphase materials. Analysis on thicker films (&gt;20 µm) by cross-sectional profiles. Possibility of stress profiling</td>
</tr>
<tr>
<td>Four-slot</td>
<td>Combination between two-slots and the ring-core. The same advantages as the ring-core. Additionally, Poisson’s ratio can be obtained within the same experiment. The procedure is more time-consuming</td>
<td>Homogeneous coatings with a thickness range of 1–10 µm, also when nonbiaxial stress is present</td>
</tr>
<tr>
<td>Ion beam layer removal</td>
<td>Method based on cantilever bending after stress relaxation by FIB cutting. Fully analytical procedure, but complex and time-consuming experimental procedure</td>
<td>Best procedure for depth profiling for coatings with the range of 0.5–10 µm</td>
</tr>
</tbody>
</table>

stepwise fashion and by using an extended integral method for stress calculation. Such a method was originally proposed for hole-drilling and can be adapted to other milling geometries, provided that the basic influence functions \( A(H,h) \) are assessed by FE calculations

\[
\bar{\varepsilon}(h) = \int_0^h A(H,h) \cdot \sigma(H) \cdot dH, \tag{12}
\]

which can be converted into a matrix formulation by assuming that the calculation is divided into \( n \) calculation steps

\[
\bar{\varepsilon}_i = \sum_{j=1}^{n} A_{ij} \sigma_j \quad 1 \leq j \leq i \leq n, \tag{13}
\]

where \( A_{ij} \) is a triangular matrix of FE calculated calibration coefficients and \( \bar{\varepsilon} \) is the measured strain relief.

The main limitation of the integral method for residual stress depth profiling relies on the mathematical ill conditioning of the matrix inversion procedure to calculate the stresses, which involves significant calculation errors for \( z/d_m > 0.3 \), as discussed in previous papers. An attempt to solve such a limitation is to increase the stub’s diameter in order to keep \( z/d_m < 0.3 \) or even to perform tests with multiple diameters to gain information on residual stress distributions over a wider range of depths. Such an approach has been lately used to evaluate residual stress profiles in multilayer Cr-CrN coatings, where a significant effect of the residual stress depth-profile on scratch adhesion was observed and discussed.119

In a recent study, Korsunsky et al.120 have presented and validated an innovative non-integral approach for residual stress depth profiling, based on a novel calculation procedure that revises the classical integral method [Eq. (13)] by focusing on eigenstrain reconstruction121 to evaluate the residual stress profile. This new method can overcome the limits of previous approaches, thus allowing for residual stress depth profiling with <50 nm resolution.

4. Factors affecting method’s reliability

Several error sources have been identified for this method, most of which can be properly taken into account and corrected for. In particular, the quality of SEM images has been identified as the major source of experimental errors and data uncertainty. A recent paper by de Hosson and coworkers113 has presented a statistical method to check instabilities in SEM imaging, based on the correlation of digital images and the definition of four different parameters that can be used to evaluate image qualities and to remove bad images from the analysis.

Some works have focused on method development as well, showing that best SEM imaging conditions can be achieved by using the integration of a large number (usually 128) of micrographs acquired with a low dwell time.104 The same papers have shown strain measurement to be much more accurate and reliable along the fast SEM scan direction.104,113 The noise associated with the strain measurement was quantified to be in the range of 1.0–5.0 × 10⁻⁴ for SEM magnifications ranging from 10 000 to 50 000×112,113 A further improvement in terms of strain resolution can be achieved by using specifically designed surface patterning/-decoration111 and by applying automated electron (and ion) beam drift correction strategies.104

Another relevant source of calculation errors can be represented by the choice of the elastic parameters for calculating stresses from strains. This is a particularly critical
situation in the case of strongly oriented materials and/or stress analysis inside single grains with unknown orientation. A recent paper by Salvati et al.\textsuperscript{122} introduced a simple statistical procedure to quantify the uncertainty of residual stress evaluation due to elastic anisotropy effects in materials with unknown texture. It is demonstrated that the uncertainty in stress calculation can even reach 40% in the case of materials with a high anisotropy factor $A$ (Nickel alloy, $A = 2.83$), while it can be kept below 10% for aluminum ($A = 1.23$).

Finally, the major issue of the additional residual strains induced by FIB damage should be carefully considered. Recent molecular dynamics (MD) studies\textsuperscript{123} have reported a detailed and quantitative analysis of the defects and additional strains induced by a cascade of ion impacting a material (silicon) with known energy. The obtained results allowed quantifying the thickness of the damage layer and the amount of induced residual strains. In combination with validation experiments, it was then possible to evaluate the minimum pillar diameter for which the induced strain can be considered as negligible (i.e., lower than the usual strain error associated with DIC). The results for a 30 kV Ga$^+$ FIB showed that the minimum pillar diameter for Silicon is equal to 1 $\mu$m, thus meaning that FIB artifacts can be considered as negligible for a trench.

5. Industrially relevant applications

In recent years, the method was used for a very wide range of industrially relevant applications. Examples include (1) the use of this technique for high-resolution stress mapping and stress profiling in thin films and coatings, (2) strain mapping in MEMS and TSV structures, (3) the analysis of residual stress distribution inside single grains (or across GBs) in polycrystalline materials, (4) the investigation on stress concentration and/or redistribution in correspondence of fatigue cracks, (5) model validation for residual stress prediction in BMGs, or even (6) residual stress mapping in biomaterials and biological (dental) tissue.

The method has also been demonstrated to be extremely useful for the design and production of multilayered PVD coatings with a tailored residual stress profile and improved adhesion. In a recent paper, Renzelli et al.\textsuperscript{119} showed that multilayer Cr-CrN coatings with a controlled stress gradient could be produced by varying the applied bias voltage during deposition, as also highlighted by other recent publications. In their work, the authors demonstrated that reducing the interfacial residual stress can be extremely effective in improving the scratch adhesion, while maintaining constant the average compressive stress throughout the film. Residual stress profiles were selected through analytical modeling and then experimentally measured by the incremental FIB-DIC ring-core method.

D. Method validation and critical comparison of the available measurement techniques

In order to establish a critical comparison between different residual stress characterization methods, a multilayer chromium nitride (CrN) film was deposited on a Si(100) substrate by means of magnetron sputtering (MS). Deposition of CrN layers was made in an Ar + N$_2$ gas mixture at a temperature of 350°C. Three consecutive CrN sublayers, each 1 $\mu$m in thickness, were deposited by varying the energy of incident ions and without interrupting the film growth. A sequence of three bias voltages of $-40, -120,$ and $-80$ V (from the substrate to the external surface) was used, resulting in a total film thickness of 3 $\mu$m. In Fig. 5(a), an SEM micrograph of the film cross-section shows changes in the film morphology corresponding to the bias voltage transitions.

The residual stress gradient in the sample was characterized by three different techniques: (1) the conventional wafer curvature method based on the Stoney equation\textsuperscript{28,124} (2) cross-sectional x-ray nanodiffraction using a sampling step of 15 nm and an X-ray beam diameter of 30 nm\textsuperscript{96} and (3) FIB residual stress profiling, according to the new procedure that was recently proposed by Korsunsky et al.\textsuperscript{120}

The FIB method allows for both the assessment of the average residual stress in the film and the evaluation of the residual stress depth profile. Figure 5(b) shows a micropillar that was incrementally milled by FIB using a current of 48 pA, resulting in a relaxation strain profile reported in Fig. 5(c). In this profile [Fig. 5(c)], a clear transition from a slow negative relaxation strain in sublayer-1 to higher positive strains in sublayer 2 is observed. This corresponds to mild tensile stress states in sublayer 1, followed by a higher compressive stress in sublayer 2.

According to the stress calculation procedure reported in Sec. II C, the average residual stress analysis by FIB indicated a value of $-584 \pm 120$ MPa, which is in very good agreement with the result from the curvature method of $-429$ MPa.

As reported in other works,\textsuperscript{103,125} curvature measurements provide usually slightly lower stress values, in comparison with FIB-DIC. This observation can be explained by considering that FIB-DIC experiments acquire information from a smaller gauge volume (a few $\mu$m$^3$), while curvature measurements provide information from a larger scale and also include stress relaxation phenomena due to the presence of microdroplets and microcracks over the entire film surface.

A comparison between FIB-DIC and cross-sectional nano-diffraction is reported in Fig. 5(d), where a remarkably good agreement between both techniques is indicated. In particular, both methods demonstrate a clear transition from mild tensile stress in the sublayer near to the surface (where grain growth contributes to stress relaxation) to a significantly higher compressive stress in the second sublayer, which is associated with the higher bias voltage ($-120$ V) applied for sublayer 2. Then, a relatively small compressive stress is detected in the bottom sublayer, corresponding to the bias voltage of $-40$ V. Finally, we observe a sharp increase in the compressive stress in the film’s region near to the interface, corresponding to the early stages of film growth with a fine-grained microstructure.

Therefore, it can be concluded that both approaches provided very complementary data on the residual stress gradients in the graded CrN film.
III. STRESS IN POLYCRYSTALLINE FILMS: CURRENT MODELS AND SELECTED EXAMPLES

A. Modeling stress development during polycrystalline thin film growth

1. Nonenergetic deposition conditions

As noted above, there is a large literature quantifying the evolution of stress in numerous systems for many deposition methods and processing conditions. Because of the impact of stress on film performance and failure, there is a strong motivation for trying to understand it in terms of the underlying atomic-level processes occurring during film growth. In this section, we describe recent progress in developing a rate-equation based model to understand the dependence of stress on the temperature, growth rate, and evolving microstructure.

Many different kinetic processes occurring simultaneously during film growth can influence the stress, including deposition, attachment of atoms to terrace ledges, GB formation, and diffusion of atoms on the surface and into the GB. Some of these are shown schematically in Fig. 6. The deposited atoms can have low kinetic energy in nonenergetic processes such as evaporation or electrodeposition. In energetic deposition processes such as MS, the deposited species have much higher thermal kinetic energies that can modify the stress. For example, sputter deposition is commonly used to counteract large tensile stresses that develop in films of refractory materials if nonenergetic deposition is used. The impact of energetic particle bombardment on the intrinsic stress development will be addressed specifically in Secs. III A 2, IV B 2, and IV E.

---

Fig. 5. (Color online) Complementary results from x-ray nanodiffraction and FIB-DIC analyses of residual stresses in the 3 μm thick CrN thin film sputtered in steps using three bias voltages of −40, −120, and −80 V (a). A FIB incremental milling of a micropillar (b) using a current of 48 pA resulted in a relaxation of the strain profile (c), which was used to determine the FIB-DIC stress depth profile (d). For comparison, a stress profile evaluated from x-ray nanodiffraction with a sampling step of 15 nm is presented. The results presented here demonstrate that both nanodiffraction and FIB-DIC profiling techniques have become robust methods for stress profiling with a submicrometer spatial resolution, for which standardization and industrialization routes could be open.

Fig. 6. (Color online) Schematic of a section of the thin film around a GB illustrating kinetic processes that can influence stress. Reprinted with permission from E. Chason, Thin Solid Films 516, 1 (2012). Copyright 2012, Elsevier (Ref. 141).
After adsorption on the surface, the deposited atoms may be mobile if the diffusivity is sufficiently high. These atoms can meet other atoms and form clusters on the surface or diffuse to sinks such as terrace edges or GBs. The film’s microstructure also evolves as the film grows. Starting from a bare substrate, the deposited atoms cluster into islands that are initially not connected, assuming that the film does not wet the substrate. As the thickness increases, the isolated clusters start to interconnect and coalesce into a uniform film. This coincides with the formation of GBs between the islands. Ultimately, the film becomes relatively uniform and flat; depending on the material mobility, the grain size may continue to change with the thickness as it grows.

The measurements of the stress-thickness in Fig. 1(b) show that the film stress goes through different stages corresponding to the evolving microstructure. In the earliest stages, the shallow slope indicates that the incremental stress is small. At a thickness of ~10 nm, the slope starts to increase, indicating a tensile stress in the layers being deposited. At ~30 nm, the stress-thickness reaches a maximum and the incremental stress changes from tensile to compressive. After this, the incremental stress remains compressive, and the average stress ultimately becomes compressive.

These different regimes of stress evolution are correlated with the evolution of the film’s microstructure with thickness. The early low-stress stage corresponds to the film consisting of individual islands on the surface. The increasing tensile stress corresponds to the onset of coalescence, where the individual islands start to impinge on each other and form GBs between them. For metal films like Ag, the transition to compressive stress corresponds to the film becoming fully coalesced into a continuous film. This results in the existence of a maximum (tensile peak) in the film stress evolution with thickness. Recent findings, based on simultaneously coupling MOSS and surface differential reflectance spectroscopy (SIRS) during deposition of a series of high-mobility metal films, have demonstrated that the onset of film continuity coincides with the tensile peak. For materials with lower atomic mobility, the incremental stress may remain tensile and not become compressive, at least under conditions of low-energetic vapor flux.

The evolution of stress with thickness depends on the material and was described as type I or II by Abermann. The behavior shown for Ag in Fig. 1 is called type II; this is characterized by the incremental stress changing from tensile to compressive with thickness and relaxing when the growth is interrupted. These materials have relatively high atomic mobility or low melting points, like Al, Ag, or Au. Alternatively, in type I materials, the incremental stress remains tensile with thickness and does not relax when the growth is interrupted. These materials have relatively low atomic mobility or high melting points, such as Mo, Ta, and W.

The different stress behavior depends on the material, but it may also be modified by changing the temperature or growth rate. For instance, evaporated Fe films grown at low temperature show stress-thickness evolution like type I materials, but when the same material is grown at higher temperature, the behavior is like a type II material. In general, higher growth rates and lower temperatures tend to promote type I behavior, while lower growth rates and higher temperatures promote type II behavior.

The stress depends on the grain size, but its dependence is complicated. Koch et al. showed that the smaller grain size can lead to more compressive stress in the growth of a type II material. Similar behavior was found for electrodeposited Ni and Cu films at low growth rates where the smaller grain size led to more compressive stress. However, at high growth rates, the smaller grain size led to the stress becoming more tensile. This shows that stress cannot be understood without considering the interaction between the growth rate and the grain size. The model developed below is able to explain this complicated behavior.

There have also been numerous measurements of the stress evolution during relaxation when the growth is interrupted. This relaxation can be reversible if the growth is resumed shortly after the interruption; for longer times, there can also be an irreversible component. Measurements of the relaxation dependence on the grain size suggest that GBs play a role in relaxation as well as growth stress.

The stress measurements provide guidance about the underlying kinetic processes controlling it. The correspondence between the rise in the tensile stress and the onset of island coalescence suggests that GB formation plays a role. Based on this, Hoffman suggested a mechanism that considers the energy for creating new sections of GB between islands and for elastically deforming the islands. This analysis shows that adjacent islands will snap together to form new GB as long as the increase in strain energy is less than the decrease in interfacial energy. The maximum tensile stress is equal to

$$\sigma_T = \frac{M_f \Delta V}{L} \left( \frac{L}{c} \right)^{1/2},$$

where $M_f$ is the film’s biaxial modulus, $\Delta V$ is the change in interfacial energy, and $L$ is the grain size. Films with small $L$ are predicted to have more tensile stress at coalescence because of the large number of GBs, in agreement with recent experimental findings. Similar results for the stress at coalescence have been produced by others using different geometries.

The transition from tensile to compressive stress in type II materials is more controversial, and several mechanisms have been proposed. One mechanism suggests that the compressive stress is inherited from the stress in the individual islands before they coalesce. This is attributed to the effect of the surface stress inducing compression in the islands before they become firmly attached to the surface. However, it is difficult to reconcile this mechanism with the reversibility of the stress relaxation.

GBs provide symmetry-breaking sites where atoms can be added to the film without distorting the lattice. Nix and Clemens proposed that adding atoms at the GB can relax the tensile stress in the film. Spaepen suggested that other
sites where atoms can be trapped are at ledges on the growing surface. These mechanisms provide ways in which compressive stress can be generated in the film, but they do not explain why the stress should become compressive. Indeed, with the increasing thickness, the film can develop net compressive stress, indicating that it is not just relaxing the initial tensile stress. Because elastic strain energy increases the total energy of the system, the stressed film has higher energy than it would if there were no stress. Since generation of stress increases the system’s total energy, there must be a driving force for the insertion of atoms into the film to generate compressive stress.

Chason et al.\textsuperscript{65} have proposed that the driving force for atom insertion into the GB is the supersaturation of atoms on the surface. This is a consequence of the nonequilibrium nature of film growth which increases the chemical potential of atoms on the surface ($\delta\mu_i$) during growth relative to equilibrium. On the other hand, stress in the film changes the chemical potential of atoms in the GB by $-\sigma\Omega$, where $\Omega$ is the nominal volume of the atom. For compressive film stress (i.e., negative values of stress), this increases the chemical potential and tends to drive atoms back out of the GB. Since the chemical potential difference between the surface and GB determines the rate at which atoms diffuse into the GB, the system can reach a steady-state where there is stress in the film due to the supersaturation on the surface. The diffusion of atoms in and out of the GB due to the surface and GB chemical potentials is also consistent with the reversibility of the stress when the growth is stopped and then resumed.

These mechanisms of tensile stress generation (due to GB formation) and compressive stress (due to insertion into GBs) have been incorporated into kinetic models\textsuperscript{65,140} to compare with the measurements. In the model discussed below, we consider the evolution of stress in each layer of the film while it is growing. We assume that diffusion in the GB is slow, and so, the stress in the layer does not change after it is covered by the next layer. An alternate set of analytical equations can be derived if we assume that the GB diffusion is rapid, i.e., the stress is uniform through the thickness of the film.\textsuperscript{65,141} For these two kinetic limits, we can derive analytical expressions for the stress evolution. Other GB diffusion conditions can be considered by solving the equations numerically, but they will not result in simple analytical expressions.

The tensile stress in the $i$th layer is reduced by the insertion of atoms into the GB ($N_i$),

$$\sigma_i = \sigma_T - M_f \frac{N_i a}{L},$$

where $a$ is the nominal size of the atom $\sim \Omega^{1/3}$. Combining this with the rate of atom insertion produces an equation for the stress evolution

$$\frac{\partial \sigma_i}{\partial t} = - \beta D \frac{aL}{\sigma_i} (\sigma_T - \sigma_i),$$

where

$$\beta = \frac{4C_v M_f \Omega}{kT},$$

$C_v$ is the concentration of mobile atoms on the surface, $D$ is the effective diffusivity for transitions from the surface into the top of the GB, and $\sigma_i \equiv -\delta\mu_i/\Omega$.

This equation describes the stress in the layer at the surface, starting from the time when adjacent islands come together in the layer to form a new segment of GB. At this point, the stress is equal to $\sigma_T$. The stress decays exponentially with time until the top of the GB layer is covered over by the next layer. The amount of time that the $i$th layer is at the surface is equal to $a/(dh_{gb}/dt)$, where $dh_{gb}/dt$ is the rate at which the top of the GB is moving upward. The resulting stress in the layer is equal to

$$\sigma_i = \sigma_c + \left(\sigma_{\tau,0} \left(\frac{L_o}{L_i}\right)^{1/2} - \sigma_c\right) e^{-\beta D L / (dh_{gb}/dt)},$$

Here, we have explicitly shown the grain size dependence of the tensile stress $\sigma_T$ by assuming that it has the value $\sigma_{\tau,0}$ when the grain size is equal to $L_o$. In the steady-state, the GB grows at the same rate as the rest of the film and $dh_{gb}/dt$ equals the average growth rate, $R$, so that the steady-state stress during growth can be expressed as

$$\sigma_{\text{growth}} = \sigma_c + \left(\sigma_{\tau,0} \left(\frac{L_o}{L}\right)^{1/2} - \sigma_c\right) e^{-\beta D R L / R}.$$  

The model shows the contributions of different parameters to the steady-state stress. The exponential contains the term $D/RL$, showing how larger diffusivity or lower growth rate will drive the stress to be more compressive. This is consistent with the observations of type I and II materials as well as measurements of the temperature and growth rate dependence. Both the exponential and the tensile stress depend on the grain size, which can explain the observed dependence of stress on the grain size. The rate at which the GB height changes is higher during the initial stage of coalescence than it is in the steady-state, which is consistent with the tensile to compressive transition in stress with thickness observed for type II materials. This effect was confirmed by studying the stress evolution in patterned films that grew as an array of hemispherical islands.\textsuperscript{142,143}

In order to compare the model predictions quantitatively with experiments, it is necessary to fully characterize the sample’s microstructure. This is important because in many films, the grain size changes as the film grows. As described by the structure zone model (SZM) of Thornton and Hoffman,\textsuperscript{25} the grain size can remain constant (zone I), change at the surface but not throughout the film (zone T), or change throughout the film (zone II). The consequence of the microstructural evolution for the stress-thickness can be seen in Fig. 7. The stress-thickness in Fig. 7(a) shows results from electrodeposited Ni which grows with a constant grain size (zone I). As predicted by Eq. (18), the constant growth rate and grain size lead to a constant incremental stress. This is consistent with the stress-thickness evolution reaching a
constant slope at larger thicknesses in the figure. The dependence of the steady-state stress on $R$, extracted from these data, is shown in Fig. 7(b). The red line represents the results of the model using parameters obtained from nonlinear least squares fitting of the data.

For comparison, evaporated Ni films deposited under ultra-high vacuum (UHV) conditions have very different microstructural evolution. Measurements by Yu et al. indicate that Ni grows with a zone II behavior. The film’s grain size is proportional to the thickness, and the grain growth proceeds through the thickness of the film so that the grains remain columnar in morphology. Therefore, we cannot assume that the grain size is constant with thickness. In addition, the extensive grain growth in the underlying layers can produce additional tensile stress due to densification of the film. As originally described by Chaudhari, increasing the grain size from its initial value creates stress by removing regions of excess volume at the GBs; this can be modified to consider the contribution of grain growth to the stress in growing films. The effect of the changing grain size and grain growth has significant consequences for the stress-thickness evolution measured in the evaporated Ni shown in Fig. 7(c). Instead of reaching a constant slope, the slope changes continuously with thickness, and a turnaround phenomenon from compressive to tensile incremental stress is observed (indicated by arrows).

Validation of the model is ongoing. New experimental studies are being performed in which all the necessary parameters are characterized so that the model can be compared directly with the data. Modifications of the model are also being made to include the effect of grain growth so that the results of zone II growth can be analyzed. A goal of all this work is to see if a series of kinetic parameters can be obtained for different materials and growth processes that will allow the stress to be predicted under different conditions.

2. Influence of energetic vapor flux

It is well known that energetic particle bombardment during growth can strongly modify the stress, affecting not only surface but also subsurface processes. Hoffman and Thornton were the first to study such effects, and they reported in a series of papers on various single metal films deposited by cylindrical magnetron sputtering over a large range of Ar working pressures a clear transition from tensile to compressive stresses when the working pressure was decreased. They additionally observed that the critical pressure for the compressive to tensile transition increased with the atomic weight of the metal target relative to that of the working gas, with increasing discharge voltage, and was dependent on the cathode geometry. They concluded that the atomic peening mechanism contributed to these effects, whereby the impact of incident energetic particles produced recoil implantation of the coating material surface atoms and entrapment of working gas atoms, at the origin of compressive stress build-up in the film. However, the influence of the deposition rate on the compressive stress magnitude could not be isolated from these experiments.

In an effort to provide a more complete picture of the stress-inducing effects during PVD film growth under energetic deposition conditions, the model presented above was extended to take into account the contribution of incident particle bombardment. Indeed, defect incorporation through ballistic-induced displacement sequences and/or collision cascades in the growing layer is ignored in the expression of the intrinsic stress given in Eq. (18). It is well known, however, that incident particles with energy in the range of several electron-volt up to 100 eV, as typically encountered in MS discharges or CAD, provide means to affect not only surface but also subsurface processes. Collisions between the impinging particles and surface atoms can result in reflection of incoming species from the surface, resputtering, or kinetically assisted surface diffusion. Arriving particles can also implant in the shallow surface through a series of knock-on mechanisms, providing, by forward momentum transfer, kinetic energy to atomic recoils. This will induce displacement of subsurface atoms located close to the impact site to more favorable sites, like vacant sites, but also will contribute to the creation of point defects (at interstitials and substitutional sites) above a certain energy threshold. If film densification occurs when energetic particle flux is involved, it is often accompanied by the development of compressive
stress, as a result of “atomic peening process.”

The volumetric distortion is proportional to the fractional number of atoms being displaced from their equilibrium sites, which based on the forward sputtering model of Sigmund implies a square root dependence of the compressive stress on the incoming particle energy.

This explains why low-mobility (type I) materials exhibiting tensile stress under nonenergetic deposition conditions are usually under compressive stress when subjected to energetic vapor fluxes. An illustration of this phenomenon is shown in Fig. 8(a) which compares the evolution of the film force per unit width of sputtered Ta films with the film thickness as a function of Ar working pressure. A clear transition from a compressive toward a tensile steady-state stress is observed with the increasing Ar pressure from 0.12 to 0.75 Pa. In particular, there exists a critical pressure, $P_c$, at which the net average stress would be zero ($P_c \approx 0.5$ Pa in the present example). The value of $P_c$ not only is material dependent (it increases with the target/gas mass ratio, $M_t/M_g$) but also is dependent on the deposition rate $R$, geometry of the deposition system (the target-to-substrate distance, $d_{TS}$), and nature of plasma discharge.

Conversely, the application of a negative bias voltage to the substrate, at a fixed working pressure, increases the compressive stress due to contribution of ion bombardment-induced defect formation in the growing layer [compare orange and green force curves in Fig. 8(a)]. The development of tensile stress at high working pressure is related to the formation of underdense columnar growth morphology, typical for refractory metals deposited under kinetically limited surface diffusion conditions.

Compared to thermal evaporation, the particle flux reaching the substrate for the case of sputter-deposition is characterized by a broad energy distribution of film-forming species. For conventional direct current magnetron sputtering (DCMS) discharges, the film-forming species are essentially neutrals, i.e., sputtered (target) atoms and gas atoms reflected back from the target, while for high-power impulse magnetron sputtering (HiPIMS), there is a large fraction of ionized species. The energy distribution of sputtered atoms and backscattered neutrals depends on the target material type and target voltage and is therefore intrinsically related to the plasma conditions. However, the energy of ions can be controlled by applying to the substrate a certain bias voltage, $U_b$. Also, particles leaving the target experience collisions with working gas atoms (if $d_{TS}$ is larger than the atom mean free path) during their transport in the gas phase such that the energy (and angular) distribution of the particle flux reaching the substrate may differ significantly from that of the nascent flux leaving the target. As an example, the energy distribution of Ta sputtered atoms and Ar backscattered atoms is shown in Fig. 9, for the case of DCMS of a Ta target with 280 eV Ar ions, at two different Ar working pressures. These data were computed using SRIM (Ref. 164) and SIMTRA (Ref. 159) codes, assuming a target-to-substrate distance of 18 cm. The particle energy range spans two orders of magnitudes, from a fraction of electron-volt to $>100$ eV. Increasing the Ar working pressure from 0.12 to 0.75 Pa leads to a modification of the energy distribution of the incoming particles: a large fraction of Ta and Ar atoms have energy lower than 2 eV, due to the collision-induced thermalization process during transport in the gas phase. Consequently, the average energy of sputtered Ta atoms, $E_{Ta}$, decreases from 31.1 to 12.3 eV, and that of backscattered Ar, $E_{gas}$, decreases from 52.9 to 25.9 eV.

Knowing the relative fraction of sputtered atoms and backscattered gas atoms and ions, the mean energy deposited per incoming Ta particle, $E_{dep}$, can be calculated from the following expression:

$$E_{dep}(eV) = E_{Ta} + \Phi_b E_{gas} + \Phi_i n_i(U_b - U_p),$$

where $U_p$ is the plasma potential (typically $\approx 10$ V in the standard magnetron configuration), $\Phi_b$ the ratio of backscattered particles to sputtered particles, $\Phi_i$ the ratio of ionized particles to sputter particles, and $n_i$ is the charge state of the

FIG. 8. (Color online) (a) Evolution of the stress-thickness during sputter-deposition of Ta films at different Ar working pressure (0.12–0.75 Pa range) and bias voltage values (ground, –60 and –190 V). Note that grounded (0 V bias) substrate conditions correspond to lines without mention of bias voltage and values given in parentheses correspond to the average energy per deposited atom, $E_{dep}$ (see text). (b) Evolution of the (compressive) steady-state stress with $E_{dep}$, determined from the data in (a). Data are taken from Ref. 158.
ionized species. As reported in Fig. 8(b), the compressive stress gradually increases with $E_{\text{dep}}$, suggesting a larger propensity of defect creation with increasing deposited energy. If the mean energy per deposited atom provides a simple way of characterizing the incoming particle flux in terms of growth energetics, this quantity does not reflect per se the important contribution of single collisional events related to the high energy tail ($>100$ eV) of the particle energy distribution, which are scarce events but at the origin of point defect creation. Indeed for Ta, the energy threshold for atomic displacement, $E_D$, is 80 eV.\textsuperscript{165} While for sputtered atoms, the fraction of species having energies higher than 80 eV is small (around 10\% at low pressure), this fraction reaches 31\% for the case of reflected Ar (see Fig. 9). The contribution of backscattered gas atoms should not be overlooked when addressing the issue of stress in sputtered films, especially for heavy-mass targets sputtered with lighter particles for which the backscattering yield $Y$ may reach $>20\%$ [$Y$ roughly scales with $(M_t/M_g)^{2/3}$].\textsuperscript{146} Relatively large compressive stress values, up to –4 GPa, have been reported for ion beam-sputtered Mo films, as a result of incorporation of interstitial defects which induce a volume expansion of the crystal lattice and an associated expanded stress-free lattice parameter.\textsuperscript{87,166} If “atomic peening” is a concept known for several decades, the kinetics of defect incorporation, their underlying mechanisms, and the interplay with film’s microstructure have not yet been systematically studied and elucidated. Fillon et al.\textsuperscript{167} have reported an increase in compressive stress for sputter-deposited Mo films at higher deposition rates, a behavior which is opposite to what is observed for materials grown under nonenergetic deposition conditions [see Fig. 7(a)]. For Mo films, Magnfält et al.\textsuperscript{168} have evidenced a densification process in which atoms are inserted at the GB, causing a source of compressive stress without significant lattice expansion. In this case, the stress field is biaxial, while the incorporation of point-defects in the grain interior would generate a triaxial stress state.\textsuperscript{166,169} Interstitials (and to some extent substitutional atoms) are, in general, unstable and can be annihilated by diffusion toward the nearest underdense regions if sufficient energy is provided. Based on these observations, we proposed to add to the thermally activated diffusion processes considered in the kinetic model presented in Sec. III A 1 two other processes related to subsurface point defects creation by atomic peening [see Fig. 10(a)], namely,

(1) Incorporation of excess atoms at the GB. The average in-plane grain size is denoted as $L$.

(2) Creation of defects in the grain, at a depth $l$ from the surface. The value of $l$ depends on the deposited energy (it is expected to decrease with increasing working pressure $P$) but is typically a few monolayers (up to $\sim 2$ nm).
FIG. 10. (Color online) (a) Schematic of the defect creation and annihilation processes due to energetic particle bombardment considered in the stress model. (b) Evolution of the steady-state stress of sputtered Ta films with different grain sizes $L$ vs deposition rates $R$. Symbols refer to experimental data obtained from real-time MOSS, and solid lines are fits to the model described in the text.

We assume that the first process is a diffusionless mechanism occurring by collision-induced displacement and relocation of atoms to more favorable sites in the vicinity of the GB. This region is marked by the shaded area in Fig. 10(a), and its width is approximated to be proportional to the distance $l$. The contribution of defect trapping at the GB to stress is expressed as

$$
\sigma_{\text{gb}}^\text{ener} = A_o \times \langle l/L \rangle, \tag{20}
$$

where $A_o$ is a parameter that depends on the deposition conditions ($E_{\text{dep}}, f$, and $R$) and that will be adjusted to compare the model with the data.

For the second mechanism, we consider the rate of defect creation to be equal to $f c_o$, where $f$ is the flux of energetic particles and $c_o$ is the number of defects generated per particle. We also take into account the possible annihilation of these defects at the free surface. The balance between defect creation and relaxation at the free surface depends on the characteristic time $\tau_s$ necessary for the defect to diffuse to the surface, and the growth rate $R$. Since the surface is moving upward due to propagation of the growth front at constant $R$, the expression of $\tau_s$ is not simply given by $\sqrt{D_s}$, $D_s$ is the diffusivity of the considered defect. For its complete expression and more details on the stress model, we refer the reader to Ref. 170. It states that the concentration of defects in the steady-state regime, $C_{ss}$, is given by

$$
C_{ss} = \frac{c_0 f}{R} \frac{1}{1 + \frac{l}{R \tau_s}}. \tag{21}
$$

Note that at high $R$ or low $D_s$, the rate of defect annihilation at the free surface goes to zero so that the value of $C_{ss}$ saturates at $c_0 f / R$. Conversely, when $R \to 0$ or $D_s \to \infty$, then $C_{ss} \to 0$. The contribution of defect trapping in the grain interior to stress can be written as

$$
\sigma_{\text{bulk}}^\text{ener} = \left(1 - \frac{l}{L}\right) \sigma_o C_{ss}, \tag{22}
$$

where the term $\sigma_o$ represents the stress per defect retained in the bulk. The factor $(1 - l/L)$ is present because we assume that the bulk stress effect is proportional to the fraction of energetic particles that are not within a distance $l$ of the GB.

The complete expression for the steady-state stress during deposition with energetic species is therefore given by

$$
\sigma_{ss}^{\text{tot}} = \sigma_{\text{growth}} + \sigma_{\text{gb}}^\text{ener} + \sigma_{\text{bulk}}^\text{ener}. \tag{23}
$$

where the different stress sources $\sigma_{\text{growth}}$, $\sigma_{\text{gb}}^\text{ener}$, and $\sigma_{\text{bulk}}^\text{ener}$ are given by Eqs. (18), (20), and (22), respectively.

The comparison of the predictions of the model given by Eq. (23) with experimental data acquired for a set of sputtered Ta films with different controlled in-plane grain sizes$^{171}$ is shown in Fig. 10(b), where solid lines correspond to best-fit values obtained from the nonlinear least-squares fitting procedure. One can observe that the trend of increased compressive stress at higher growth rates is captured by the extended kinetic model. This effect is more pronounced for the series of Ta films with the largest grain sizes ($L = 950\, \text{nm}$). Looking at the individual terms of Eq. (23), it is found that the dominant stress contribution is $\sigma_{\text{gb}}^\text{ener}$ which counterbalances and even outweighs the thermally activated surface diffusion term $\sigma_{\text{growth}}$. The same trend is also reported for sputtered Mo films and well reproduced by the model$^{170}$. With the decreasing grain size, the magnitude of the compressive stress is found to increase: an almost linear dependence of $\sigma_{ss}^{\text{tot}}$ with $1/L$ at a constant growth rate (not shown here) is observed, which can be ascribed primarily to the contribution of $\sigma_{\text{gb}}^\text{ener}$ which scales with $1/L$ [see Eq. (20)]. The other term, $\sigma_{\text{growth}}$, also contributes to a higher compressive stress at a lower grain size but to a lower extent in the present experimental conditions.

The agreement between the calculated stress and the data shows that the model is capable of reproducing the dependence on $R$, $P$, and $L$ seen in the experiments which incentivizes further pursuing this approach to modeling the stress.

An even more challenging area of study is to model the stress evolution in alloy thin films. These are especially important technologically. Experiments in sputtered metals$^{172}$ and nitrides$^{173}$ show that in some cases, the alloy behavior looks like a superposition of the stress behaviors of...
the individual constituents. This suggests that it may be possible to consider the stress in terms of a superposition of stress-induced effects in the single element systems. However, there exist situations, as presented below, where such simple considerations do not hold.

B. Impact of interface reactivity, alloying, and phase transformation on stress

The sequence of compressive-tensile-compressive (CTC) stress evolution, shown in Fig. 1(b), is typical for high-mobility metals deposited on weakly interacting and inert substrates, such as amorphous SiO₂, for which texture inheritance can be disregarded and chemical reactivity is insignificant. The resulting film microstructure is, in this case, typical for a polycrystalline material. On a crystalline template, adatoms will experience the interatomic potential of the substrate and will locate at favorable adsorption sites corresponding to minima of the potential energy landscape of the substrate. Under conditions of sufficient surface diffusivity, a pseudomorphic epitaxial growth is observed, up to a critical thickness above which the stored elastic strain energy due to lattice mismatch with the substrate is relieved by means of nucleation and glide of dislocations, cluster nucleation, surface undulations, etc. While the case of epitaxy, and how it affects growth stress, is not addressed in this article, we will highlight in what follows the influence of chemical interaction with the substrate on the stress evolution. The case of alloy thin films and dynamic (i.e., during growth) segregation will be addressed in Sec. III B 2.

1. Interface reactivity and surfactant effects

Figure 11 illustrates the influence of chemical reactivity between the substrate and the deposited metals on the stress response. In Fig. 11(a), the film force evolution during growth of evaporated silver on the amorphous germanium (a-Ge) layer is shown, compared to the deposition on a-SiO₂, as reported by Flötotto et al. A CTC behavior is observed for the Ag film on both substrates, but the position of the tensile peak maximum is drastically altered. This means that the onset of film continuity is reached at thicknesses of ~60 nm on a-SiO₂ and ~8 nm on a-Ge. TEM observations confirm that the Ag film deposited on a-Ge has much finer grain morphology, and the sequence of island nucleation, island coalescence, and formation of a continuous film occurs at a much smaller thickness compared to that on a-SiO₂. By performing angular-resolved x-ray photoelectron spectroscopy, the authors disclosed that Ge atoms are segregating during growth of Ag, forming a Ge-rich topmost surface layer. The Ge surface coverage is below 1 ML and decreases with the increasing Ag thickness (reaching 0.2 ML for the ~300 nm thick Ag film). In the precoalescence stage, the higher density of islands and the lower surface diffusivity of Ag adatoms on the parent a-Ge surface (the diffusion barriers are 0.45 and 0.32 eV for Ag on a-Ge and a-SiO₂, respectively) promote the formation of a continuous Ag film at earlier thickness, together with a smaller grain size. In the postcoalescence stage, the segregating Ge atoms also hinder the surface diffusion of Ag and lead to continued renucleation of new Ag grains during growth. Due to the much higher GB density of the Ag film deposited on a-Ge, the latter one develops larger compressive stress [see Fig. 11(a)], in agreement with the model presented in Sec. III A.

![Fig. 11. (Color online) (a) Real-time evolution of the film force per unit width during thermal evaporation of Ag films on a-SiO₂ and a-Ge surfaces at a growth rate of \( R = 0.035 \) nm/s; graph adapted from Ref. 175. (b) Real-time evolution of the film force per unit width during sputter-deposition of Cu films on a-SiO₂, a-Ge, a-Si, a-SiNₓ, and a-C surfaces at a growth rate of \( R = 0.065 \) nm/s.](image-url)
Figure 11(b) shows the evolution of the film force during sputter-deposition of Cu films on different types of amorphous substrates, namely, a-SiO\textsubscript{2}, a-Ge, a-Si, a-SiN\textsubscript{x}, and a-C layers. While Cu films deposited on a-SiO\textsubscript{2}, a-SiN\textsubscript{x}, and a-C layers share a similar CTC behavior (having a tensile peak maximum in the 7.5–9.0 nm range), the stress development on a-Si and a-Ge is markedly different: (1) in the earliest growth stage, for the Cu film thickness below 1 nm (\textasciitilde 5 ML), the instantaneous stress goes slightly tensile instead of being slightly compressive, (2) there is no characteristic shape of a tensile peak, and (3) at later growth stages, above 15 nm, less compressive stress is being developed. Noticeably, for a-Ge, the stress shows a complex evolution, characterized by consecutive compressive and tensile stress transients below \textasciitilde 12 nm. SDRS experiments reveal the absence of isolated Cu islands on a-Si and a-Ge surfaces for Cu film thicknesses below 2 nm, which suggests a strong chemical reactivity at the interface. The formation of an interfacial Cu\textsubscript{2}Ge alloy is evidenced in the Cu/Ge multilayer from analysis of XRD patterns.\textsuperscript{176} The driving force for this compound formation is yet to be deciphered, especially to identify the fast-diffusing species and any surfactant effect. It is interesting to note the significant improvement of the [111] out-of-plane texture of Cu films deposited on a-Ge compared to what is observed on a-SiO\textsubscript{2}, while opposite results were reported by Flötotto et al.\textsuperscript{175} for Ag on a-Ge. In the case of Ag/a-Ge, the continuous presence of segregating Ge adatoms throughout growth reduces the diffusion length of Ag adatoms, favoring copious renucleation events, which can explain the development of randomly oriented grains. The stronger reactivity of Cu, compared to that of Ag, together with dynamic intermixing favored by higher energy particle flux hitting the film surface under sputter-deposition conditions compared to thermal evaporation likely explains the distinctive stress and texture evolutions of these metals on a-Ge. Finally, it can also be concluded from the stress data of Fig. 11(b) that a-SiN\textsubscript{x} and a-C behave as weakly interacting surfaces, similar to native silicon dioxide, and can be used as suitable buffer layers for depositing Volmer–Weber metal films for plasmonics or catalysis applications.

2. Alloying effect

Secs. III.A.1 and III.A.2 have reviewed and established the importance of GB in regulating the stress evolution in polycrystalline thin films. These examples concerned the case of single species films. For alloys, with atoms having different surface mobilities, segregation energies, and atomic radii, one may expect synergistic or antagonist effects to come into play, resulting in different growth stresses. Changing the alloy composition affects the enthalpy of mixing, which likely also influences the stress state. Fu and Thompson\textsuperscript{172} have studied the compositional dependence of intrinsic stress of Fe\textsubscript{x}Pt\textsubscript{1-x} alloys obtained by cosputtering from elemental targets at room temperature. This system is interesting because, when deposited as pure elements, Fe and Pt behave as low-mobility and high-mobility species, developing tensile and compressive steady-state stress, respectively. In the bulk, Fe and Pt are rather miscible elements: the Fe-Pt phase diagram exhibits a high-temperature fcc solid solution domain extending over the whole compositional range, while ordered Fe\textsubscript{3}Pt, FePt, and FePt\textsubscript{3} compounds crystallize at lower temperatures. With the increasing Pt content in the alloy film, the intrinsic stress changed from tensile to compressive. However, the stress variation did not obey a linear mixing trend; in particular, the compressive stresses of Fe\textsubscript{x}Pt\textsubscript{1-x} alloys with x ranging from 0.54 to 0.79 were larger than that of pure Pt. The authors explained this behavior by the propensity of the more mobile Pt adatoms to segregate at the GB, as confirmed from atom probe tomography (APT) analysis. The presence of Pt atoms in the GB modifies its chemical potential so that more compressive stress is built up if the chemical potential difference between the surface and the GB is enhanced with the increasing Pt content in the film. The authors also pointed out that the driving force for the chemical enrichment at the GB is likely attributed to enthalpy minimization. They also noticed that the alloy with x = 0.35 exhibited a “zero-stress” state together with negligible compositional segregation of one species at the GB compared to the bulk of the grains. It is noteworthy to disclose, for this specific system, that the alloy composition had little effect on the resulting grain size and that the intrinsic stress was not affected by a change in the deposition rate at fixed alloy composition.

There are other recent reports which highlight the influence of solute segregation on the microstructure and peculiarities in the stress behavior of binary alloys. It was recognized that the addition of a few at. % of Al (Ref. 177) or Ni (Ref. 178) in Cu alloys formed by cosputtering leads to a higher amount of compressive stress in the continuous (postcoalescence) film regime, compared to the pure Cu case. Similar findings have been obtained in the case of the Cu-Ge system upon increasing the Ge content.\textsuperscript{179} This effect is consistent with a tensile peak maximum being shifted to lower film thicknesses, indicating a reduced island size during nucleation and at percolation in the presence of Ge atoms. While a clear refinement (by a factor of 5) in the grain size was revealed for Cu-Al alloys upon alloying Cu with 1 or 2 at. % Al, pointing to the reduction of adatom surface mobility, similar to the incorporation of impurity elements like oxygen,\textsuperscript{180} the decrease in the grain size was more moderate (a factor 2–3) in the case of alloying with Ni. Kaub\textit{et al.} proposed that the highest compressive stress for the Cu-Ni alloy with the least Ni content (5 at. %) was due to Ni-enrichment at the GB, with a subsequent increase in the Ni content resulting in a decrease in the compressive stress magnitude once the GB saturation is achieved.\textsuperscript{178}

For the case of low-mobility metals, as exemplified in the Fe-Cr system, addition of Cr in solute concentration into Fe resulted in a decrease in the tensile stress in the postcoalescence regime.\textsuperscript{181} An overall good correlation between the change in grain size L and the magnitude of tensile stress σ was observed (with σ scaling with 1/L) although variations in the degree of film texture, GB character (high-angle
versus low-angle), and GB chemistry$^{182}$ may also control the stress state in such systems.

Fu and Thompson also addressed the case of immiscible binary alloys by investigating the Fe-Cu system.$^{183}$ The stress evolution was monitored during sputter-deposition of a Fe$_{0.51}$Cu$_{0.49}$ alloy at different \textit{in situ} deposition temperatures, from 25 to 325°C, using a 10 nm thick Si$_3$N$_4$ buffer layer acting as a diffusion barrier layer to prevent any deleterious silicide formation at the film/substrate interface. Increasing the deposition temperature resulted in the phase separation of the FeCu alloy into body centered cubic (bcc) Fe-rich and fcc Cu-rich phases. While similar stress evolutions were recorded in the initial stages of island nucleation and coalescence, the magnitude of the compressive stress was found to increase notably during the continuous growth regime with deposition temperature. Note that the grain size and surface roughness significantly increased with deposition temperature. From cross-sectional and plan-view STEM observations, it was concluded that the Fe islands nucleate off of the Cu islands. The origin of increased compression at higher deposition temperature could not be associated with a change in the atomic volume during phase separation, as Fe and Cu share a similar atomic volume. Rather, trapping of excess atoms at surface defects and/or at GB was invoked. The formation of a rougher surface with increasing deposition temperature would, however, reduce the diffusion of adatoms into GB due to the self-shadowing effect and the Ehrlich-Schwoebel barrier at the step edge.

3. Stress evolution in phase transforming thin films

As described in this review, the growth of a film—from an embryonic island to a coalescence state—will result in a dynamic evolution of the thin film stress states. This stress evolution has then been used to infer the adatom mobility of the deposited species in relationship to the film’s microstructure progression. In a similar manner, the same \textit{in situ} stresses can be used to help deciphering phase state changes within the film, with such changes being crystalline-amorphous transformations, polymorphic crystalline transitions, and precipitation. Thus, stress monitoring can provide unique diagnostic insights into both the adatom mobility between phase states and the microstructure evolution between phase states. The stress generators created by the phase changes are nominally linked to the volumetric change created by the new phase and/or the interfacial strains resulting from the new phase in contact with either other phase(s) or the substrate. In this section, a few highlights of stress dependent responses in phase changing films are given to illustrate how film stress has been employed as an investigative response into understanding mechanisms of thin film stress evolution.

In the recent works by Li \textit{et al.}$^{184,185}$ the growth stresses for the hexagonal close packed (hcp) to bcc Ti phase transformation were monitored as the Ti film grew. Ti was deposited in a multilayer stack using two different bcc template layers—Nb or W—to determine the effect of the template bcc lattice matching on the stabilization. In each report, the Ti layer thickness equaled that of the bcc template layer. For the Ti/Nb multilayers,$^{184}$ the bcc Ti was stabilized up to 2 nm after which it reverted to its bulk hcp phase with a further increase in thickness. This bcc Ti phase is equivalent to its bulk high temperature $\beta$-Ti phase. By extrapolating the $\beta$-Ti lattice parameter to room temperature, it was found to be 0.327 nm, which was very similar to that of the bcc room temperature Nb parameter of 0.330 nm. The close matching resulted in a coherent interface. As the Ti layer initially grew on the Nb layer, the stress-thickness revealed a slight tensile (or positive slope) condition (Fig. 12). This stress response was associated with the tensile strain that occurred as the bcc Ti lattice matched the slightly larger lattice spacing of the Nb layer. At approximately 2 nm of Ti growth, the stress-thickness slope transitioned and became negative or compressive, also shown in Fig. 12. Using postmortem electron diffraction of the multilayers, this thickness corresponded to the change from bcc to hcp Ti.

In comparison, when Ti was deposited onto the bcc W template layers in the Ti/W multilayer, a different stress response was observed.$^{185}$ Unlike Nb, W’s lattice parameter of 0.316 nm provided a significant strain that consequently reduced the Ti layer’s stabilization to $\sim$1 nm. During Ti growth, the stress thickness was found to exhibit a very steep (negative) slope gradient over the first $\sim$1 nm. This was associated with the bcc Ti being under compression in an attempt to match the W template. Ti transformation to hcp Ti after 1 nm retained the compressive stress state, but the stress in the thicker regime was less compressive, indicating some stress relief with the transformation. In both examples, the \textit{in situ} stress measurements provided real-time diagnostics of the phase changing behavior of the growing film.

Beyond polymorphic crystalline phase transformations in multilayers, amorphous to crystalline transformations (or crystallization) have also been observed. For example, Mo/ Si multilayers used for x-ray optic mirrors can stabilize an amorphous Mo layer up to a thickness of $\sim$2 nm.$^{186}$ This

![Fig. 12. (Color online) In situ growth response of Ti on Nb. Note the slight positive or tensile stress response of Ti up to 2 nm where upon it transitions to a negative or compressive stress for larger layer thicknesses. This thickness represented the change from bcc to hcp Ti. Figure adapted from Ref. 184.](image)
Amorphous stabilization can be extended by alloying Mo with Si. Fillon et al.\textsuperscript{41} provided a nice summary of this stabilization in relationship to the stress dependence for a series of Mo$_{1-x}$Si$_x$ films. For $x < 0.15$, a tensile stress was found to develop after $\sim 3-5$ nm of growth with this tensile stress gradually leveling out; for $0.16 \leq x < 0.19$, a large tensile stress developed after $\sim 13-20$ nm of growth which was subsequently retained; and for $x \geq 0.19$, the stress state rapidly became compressive after $\sim 3$ nm. Each of these stress states were found to correspond to the phase state of the film. For $x \geq 0.19$, the films were amorphous, whereas compositions below this value were related to a critical thickness-dependence for an amorphous-to-bcc solid solution transition. The authors explained the critical thickness dependence in terms of volumetric and interfacial energy descriptions for stabilization, which has also been developed for polymorphic transformations.\textsuperscript{187} These stresses would then be related to the volumetric strains that would accompany the crystallization of the film. What was intriguing in their findings was that the largest tensile stresses corresponded to the largest grain sizes. In polycrystalline films, the attractive forces at the GBs are nominally inversely proportional to the grain size. To explain this potential conflicting issue, the authors correctly pointed out that these films do not initially grow polycrystalline but are amorphous, and the stresses dominated by volumetric changes from the transformation would propagate as elastic strains as the film thickens. Thus, their findings emphasize previous sections of this review concerning microstructure-stress relationships as well as expanding those in recognition of the influence of a phase transformation on the stress response. Recently, this scenario of a spontaneous, compositional- and thickness-dependent crystallization process in Mo$_{1-x}$Si$_x$ alloys was confirmed by combining in situ and real-time wafer curvature and XRD and x-ray reflectivity experiments simultaneously during thin film deposition.\textsuperscript{188} Data reported in Fig. 13(a) show that the sudden tensile rise that manifests above a certain deposition time is concomitant with the rapid increase in the integrated XRD intensity [Fig. 13(b)] of the bcc (110) peak for Mo$_{1-x}$Si$_x$ alloys with Si content $x \leq 0.20$. This crystallization is also accompanied by a sudden increase in surface roughness [Fig. 13(c)], which continues during later growth. A significant delay in the amorphous-to-bcc solid solution transition is observed when the Si content approaches a critical composition, $x \sim 0.20$. For $x \geq 0.20$, this critical thickness is found to be $\sim 10$ nm. The film with $x = 0.25$ does not exhibit any phase transition with deposition time and grows persistently in an amorphous state, with a slight compressive stress and smoother surface. From a simulation of the temporal evolution of the XRD integrated intensity [Fig. 13(b)], the propagation velocity of the crystallization front was found to be $\sim 13$ nm/s, highlighting for a relatively fast process. This is, however, several orders of magnitude lower than what usually takes place during laser-induced or metal-induced explosive crystallization (EC) of $a$-Si and $a$-Ge (velocity on the order of cm/s to m/s).\textsuperscript{189–191} Noteworthily, the initial stress state of precursor $a$-Si films is suggestive to influence the EC mechanisms and resulting grain sizes of the formed poly-Si films.\textsuperscript{191}

Similar crystallization stress state studies have been investigated in amorphous Te-based alloys, which are potential phase changing materials for optical data storage. In the work by Leevard Pedersen et al.,\textsuperscript{192} a series of ex situ heats were conducted where the stress associated with the onset of
crystallization was monitored during the annealing. The authors noticed that the observed stress accounted for only \(~9\)% of what would be computed from a purely elastic crystallization reaction. Further stress relief noted in the films was then associated with the plastic flow in the amorphous phase. Moreover, the stress relaxation rate after crystallization was found to be minimal, suggesting to the authors that inelastic stress changes must be accommodated by viscous flow in the amorphous phase. Thus, from these phase changing stress studies, the authors concluded that a small volume change and a small elastic modulus would be preferred in a phase changing optical device film if one is to avoid extensive stress accumulation from the accompanying viscous flow that occurs upon crystallization.

Beyond internally induced stresses from a phase transformation, film stress dependent phase changes can also be influenced by external processing variables such as temperature and deposition method. In the work by Clevenger et al.,\textsuperscript{193} the tetragonal \(\beta\) to cubic \(\alpha\) phase change in Ta films was investigated. In their experiments, 100 nm thick Ta films were grown in both tensile and compressive stress states dependent on the deposition method, evaporation, or DCMS, with the sputtered films further modified by varying pressures and RF substrate bias. With the variety of Ta films deposited, they were heated up to 850 °C at 10 °C/min under purified He with the stress monitored during the annealing by a wafer curvature method. The \(\beta\)-to-\(\alpha\) Ta phase transition occurred over a range of temperatures between 600 and 800 °C, with the onset of the transformation strongly dependent on the initial film’s deposition process conditions. This \(\beta\)-to-\(\alpha\) phase change was noted to be the main stress relief mechanism for the Ta films and was most pronounced in the intrinsically compressive stress states. Incomplete compressive stress relaxation was observed if the film was initially deposited in the \(\alpha\)-phase state or if the \(\beta\)-phase only partially transformed.

Thin films of NiTi, commonly referred to as the Nitinol shape memory alloys, provide a unique study of a phase changing material that can be cyclically iterated between two distinct phases. Such materials are candidates for MEMS applications. Fu et al.\textsuperscript{194} have reported how variations in processing conditions and Ti contents can alter this phase dependent stress response for a series of Nitinol based films. In their paper, they varied the sputtering deposition pressure and Ti composition and linked these variables with the phase changing stress response via \textit{ex situ} annealing. At the lowest sputtering pressure (0.8 mTorr), the films exhibited compressive stress states which were associated with the sputtered atoms having lost minimal kinetic energy from a reduced number of collisions with gases in the vacuum chamber. These conditions contribute to enhancing atomic peening and adatom mobility during growth, as discussed previously in Sec. III A. With the increasing sputtering pressure (3.2 mTorr), the films became tensile which then decreased with an ever increasing film thickness >250 nm or films deposited at even higher pressures (10 mTorr). These reduced tensile stress states were rationalized to various relaxation mechanisms within the microstructure of the evolving film.

In each film, the as-deposited TiNi film state was noted to be amorphous. Upon annealing the films up to \(~130\) °C (400 K), each exhibited an evolution toward an ever increasing compressive stress condition where upon the stress then reverted toward either tensile direction (for the 0.8 mTorr film) or relatively invariant stress evolution (for the 3.2 mTorr film) during densification of the films’ microstructure [Figs. 14(a) and 14(b)]. Upon crystallization, a notable slope change was observed in the film stress state, which then readily became compressive (negative slope) with increasing temperature. Cooling the films from \(~630\) °C (900 K), the stress reverted toward a tensile stress state and the shape memory austenite-martensitic transformation became readily apparent as an abrupt reduction in stress below 400 K, also shown in Figs. 14(a) and 14(b) with appropriate phase transformation identifiers labeled within the figure. These changes in stress relief provide clear indications of a phase-dependent stress behavior for alloys whose composition does not change but undergo either a crystallization or a diffusion-less shape memory transformation with temperature.

The stress state associated with this shape memory transformation was also found to be composition-dependent [see Fig. 14(c)]. For films that were 47.5 at. % Ti, a small change in the stress evolution curve was observed, suggesting that the martensitic transformation was not significant. When the films were 48.5 at. % Ti, the shape memory transformation did increase, which resulted in an increase in the recovery stress upon the martensitic-austenitic transformation. But the heating and cooling stress curves did not converge to the same temperature-stress levels at ambient temperatures, leading to a partially closed hysteresis loop. This separation was suspected to be associated with an additional R-phase transformation in the shape memory effect; the R-phase is an intermediate structure between the high temperature, high symmetry austenite structure and the low temperature, lower symmetry martensitic structure. Upon increasing the Ti content to 49.1%, a two-step transformation was clearly observed upon heating and cooling and is considered clear evidence of this R-phase transformation. Once the film composition became Ti-rich, i.e., 50.2 at. %, the two-step transformation was not present, the stress versus temperature heating and cooling loops were closed, and the films exhibited a perfect shape memory transformation. A further increase in the Ti content decreased the martensitic transformation temperature but with only a partial relaxation of the stress states that were observed at the previous Ti-rich compositions.

The study by Shen et al.\textsuperscript{195} compiled many of the characteristics described above but in a single system which undergoes a composition-dependent phase state-stress dependent relationship. In their work, reactive sputtering of W in an \(\text{Ar} + \text{N}_2\) atmosphere was undertaken to produce \(\text{WN}_x\) films. It was found that N concentrations less than 8 at. % retained a bcc solid solution phase, whereas increasing the N content between \(~12\) and 28 at. % vitrified the W film. If the N content reached \(~32\) at. %, the film again crystallized as the
were rationalized to be pressure-dependent variations in the concentration of $/C_24$ increased, with the lowest value being an amorphous film at however, the stress decreased as the N concentration increased, with the lowest value being an amorphous film at a concentration of $\sim 20$ at. % N. These structural changes were rationalized to be pressure-dependent variations in the mixed Ar:N$_2$ gas ratio used during sputtering. The gas mixture, at a fixed sputtering pressure of 6 mTorr, appeared to influence the average energy of the speciesbombarding the film during deposition. This impacted the microstructure and corresponding residual stress response. As the N content increased, the columnar microstructure observed in the solid solution W films transitioned to a disordered amorphous network that promoted a loss in intercolumnar coupling and increased porosity in the GBs. These phase-changed microstructure alternations facilitated the observed relaxation of the compressive stresses.

In some cases, the phase transforming stress is a result of intrinsic impurities that cause a phase transformation. In the very early stages of W film growth, the nominally bcc phase can adopt a A15 structure, called $\beta$-W in the literature. Kaub and Thompson$^{200}$ have shown that small amounts of Ti solute coupled with in situ annealing during deposition will change the thickness stability and the magnitude of the initial stress state for A15 W. They suggested that Ti reacts with the residual oxygen in the film, which has been suggested to facilitate the stability of bcc W. APT confirmed a higher presence of titanium oxide complexes in the mass spectrum, providing evidence to support the notation that oxygen is likely regulating the phase stability of A15 W, and by controlling its reactivity with W, the phase transformation is manipulated. Thus, alloying, in this case, is not necessarily controlling the stress state from a reaction between the solvent and solute species but rather a means to control impurities in the system responsible for phase transformations.

A concluding comment is made concerning intermetallic ordering in phase changing films and their phase-stress responses. Very much like the martensitic transformation of NiTi, where the composition does not change, an ordered intermetallic film (where the atoms reside in specific lattice sites) can have an equally pronounced effect on film stress. In the work by Fu et al.,$^{40}$ a series of FePt thin films were sputter deposited at various in situ deposition temperatures to promote the solid solution A1 to an ordered intermetallic $L_{10}$ phase change. It was found that upon increasing the degree of ordering, the film’s compressive stress was reduced. Nominally, an increase in deposition temperature would promote higher mobility and potential microstructures that would increase the compressive stress. This deviation in behavior for the chemically ordering film was contributed to the adatoms becoming less mobile as they adopted their thermodynamically preferred lattice registry sites in the growing film. Even more interesting was the stress relaxation response that occurred with growth interruption (or, in other words, how the stress evolves with time after deposition ceases). As the film showed an increase in the degree of chemical order, the rate of compressive stress relaxation dramatically increased. One could assume that a film with a higher compressive stress would be biased for greater stress relaxation. In contrast, the opposite was noted as the least ordered films were the most compressive. The rapid stress relaxation noted for the ordered films contributed to these ordered films having an additional chemical potential contribution to bias the stress relaxation according to the model proposed by Chason et al.$^{65,140}$

Phase transformation, whether polymorphic, crystallization, order-disorder reactions, etc., has a dramatic effect on the stress evolution of thin films. These transformations result in varying stress relaxation mechanisms and provide insights into the mobility responses of the adatoms involved in these changes. The phase transformations offer stress relieving mechanisms that will change the mechanical attributes of the films that contain them. Through this brief section’s overview, a few highlights have been given to illustrate how thin film stress evolves under a variety of phase transformation types and how residual stress measurements offer a very valuable diagnostic tool for investigating such phase transformations and their influence on mechanical responses.

### C. Stress gradients and texture development

The structure-stress dependences in nanocrystalline thin films were discussed in Secs. III A and III B mainly with respect to the variation of the grain size, fraction of GB, and

---

**Fig. 14. Stress vs annealing behavior under various transformations:** (a) NiTi film grown at 0.8 mTorr, (b) NiTi film grown at 3.2 mTorr, and (c) different compositions of NiTi films (Ref. 194). Reprinted with permission from Fu et al., Surf. Coat. Technol. 167, 120 (2003). Copyright 2003, Elsevier.
growth kinetics/energetics. The film microstructure can also evolve with respect to the orientation of the grains so that the influence of texture development on the stress gradient has to be considered.

In Figs. 15(a)–15(e), the results from cross-sectional x-ray nanodiffraction analysis on a 15 μm thick CrN coating deposited on the Si(100) substrate using MS in an Ar + N2 gas mixture at a temperature of 350 °C using a sequence of three bias voltages of −40, −120 and −40 V are presented.96 Three consecutive CrN sublayers, each 5 μm in thickness, were deposited by varying the energy of incident ions and without interrupting the film growth. In Fig. 15(a), a SEM micrograph of the film cross-section demonstrates the film morphology with three sublayers and needlelike crystallites in the top and bottom sublayers. In Fig. 15(b), diffraction intensity as a function of Bragg’s angle 2θ is plotted. The data were obtained by integrating the Debye–Scherrer rings in the δ range of 80°–100° (Fig. 2). The results indicate that the middle region of the monolithic CrN coating possesses different lattice parameters as other two sublayers. The variation of the intensity documents that the texture of the film changes as a function of the thickness. In Fig. 15(c), a variation of the diffraction intensity along the Debye–Scherrer ring of CrN 200 reflections documents that the coating grows in the near-substrate regions with CrN (100) planes oriented parallel to the coating–substrate interface and that this preferred orientation changes with the distance from the substrate. In Fig. 15(d), a variation of full width of half-maximum (FWHM) of CrN 111 reflection is presented. In the interface region, large FWHM values indicate the presence of small crystallites and/or the presence of strains of second and third order. At a distance of ~3 μm and more from the interface, the small and the large FWHMs for δ ≥ 0 and δ ≥ 90 indicate the presence of anisotropic crystallites elongated along the surface normal, respectively. At depths of ~5 and ~10 μm, when the bias changed, one can observe an increase in FWHMs, indicating the presence of small crystallites and/or a large amount of crystallographic defects, which can be interpreted by the nucleation process.

Finally, in Fig. 15(e), the residual stress dependence in the coating evaluated using Eq. (10) is presented. At the interface, a relatively large compressive stress of ~1.5 GPa in the coating nucleation region can be explained by the presence of small crystallites and large density of GBs, which dominate the microstructure and allow for the generation of structural defects during the film growth. When the film grows, the magnitude of the compressive stress decreases and then saturates at a depth of ~10–12 μm. The decrease and the saturation can be explained by the gradual evolution of V shape grains and the decrease in GB density. When the bias changed from −40 to −120 V, at a depth of ~10 μm, ion bombardment with more energetic particles leaves a higher fraction of residual radiation-type damage in the form of clusters of point defects (mainly interstitials) and/or trapping of atoms at GB sites, as discussed in Sec. III A 2, resulting in compressive stress of a higher level (~2.5 GPa). After the bias was changed again to −40 V, the stress decreased and saturated because of microstructure saturation.

The results in Fig. 15 indicate a very complex nature of the microstructure and stress evolution in the coating, which is due to self-organization phenomena as well as process parameters. The results in Fig. 15 were obtained from one scan lasting less than 2 min and document the possibility of cross-sectional x-ray nanodiffraction to correlate residual stress gradients with microstructural evolution and process parameters.

Fig. 15. (Color online) Experimental results collected using the cross-sectional x-ray nanodiffraction approach from a 15 μm thick CrN nanocrystalline coating on a Si(100) substrate (Ref. 96): (a) SEM image of the film deposited at bias voltages of −40 and −120 V. (b) Diffraction scans collected at different depths reveal the presence of three sublayers with different lattice parameters and crystallographic texture gradients. A map (c) of diffraction intensities for CrN 200 reflection demonstrates smooth transitions of fiber textures across the coating. FWHMs of CrN 111 reflection (d) demonstrate the complex microstructural development with three nucleation zones. Coating depth dependent in-plane residual stresses (e) correlate with the microstructural evolution [(a)–(d)] and process conditions.

JVST A - Vacuum, Surfaces, and Films
IV. STRESS IN ENGINEERING COATINGS AND COATING SYSTEMS

Fundamental studies on structure-stress interrelation and advanced measurement methods have become fundamental tools for the understanding of material performance in real-life industrial environments. In fact, the accurate knowledge of residual stress distributions at multiple-length scales (from macro- to nano-) can be extremely relevant to better understand the mechanical behavior and reliability of microsystems and thin films in practical applications.

In Secs. IV A to IV D, a series of industrial applications is presented, where the importance of residual stress control and monitoring and the correlation between residual stress and in-service performance are highlighted and discussed with specific reference to Secs. II and III. In particular, we present selected examples from the areas of microelectronics, optical coatings, coatings on plastics, and tribological coatings for aerospace. Although more examples from other fields, such as health and energy, could also be shown, the following considerations and case studies are highly relevant in order to highlight the main and effective strategies to tailor intrinsic stresses in coatings, with specific reference to different coating/substrate combinations as presented during the 2016 Stress Workshop.

We also provide in Sec. IV E a brief synthesis of the most common stress generation and relaxation mechanisms in relevance to the case studies presented in this paper and outline novel approaches to manage stress. Finally, Sec. IV F presents our current understanding on crack initiation in stressed layers.

A. Stress in microfabricated systems

The promise that nanotechnology holds with respect to increased performance and reliability lies in the ability to utilize effects that naturally arise in structures with greater surface-to-volume ratios. For example, the increased mechanical flexibility imparted by MEMS features by incorporating lithography with novel deposition and etching techniques has allowed for intricate and unique arrangements of materials. Lithography with novel deposition and etching techniques has allowed for intricate and unique arrangements of materials. Because this substrate orientation has been shown to possess a larger lattice parameter than Si, such as Si$_{1-x}$Ge$_x$, or tensile strain by incorporating materials with a smaller lattice parameter, such as Si$_{1-x}$C$_x$. Liner films possessing large residual stresses can be lithographically patterned to generate stress concentrations near CMOS features, representing another scheme that has demonstrated increased carrier mobility in the underlying semiconductor.

1. Characterization of residual stress through the wafer curvature

Assessing the effectiveness of these approaches first requires a determination of the intrinsic stress that these stressor materials contain both before and after device manufacture. Standard characterization involves wafer curvature measurements performed on blanket stressor films deposited on Si substrates. Quantifying the corresponding residual blanket film stress ($\sigma_B$) relies on using the appropriate mechanical model that describes the elastic response of the layered, composite system. If the film and the substrate are perfectly bonded and uniform in thickness, then the Timoshenko model or the Stoney equation in the case of film thicknesses much less than that of the substrate, can be used to extract film stress from the differences in the curvature before and after film deposition. For a single-crystal substrate, its elastic anisotropy and crystallographic orientation affect the observed curvature due to elastic stress. Neumann’s principle dictates that the symmetry of the crystal system must be present in the symmetry of the property, which is described by second-rank tensors for stress and strain. For example, strain induced in Si (111) and Si (001) substrates, which possess threelfold and fourfold symmetry, respectively, generates radially symmetric wafer curvature profiles. However, Si (110) substrates exhibit twofold symmetry and can support two independent radii of curvature. Because this substrate orientation has been shown to produce improved p-type device mobility in planar FET (Ref. 211) and FinFET devices, it is important to properly analyze the anisotropic strain distributions and the corresponding curvature in such substrates.

The intrinsic elastic anisotropy present in Si can result in a distribution in the curvature as a function of in-plane angle even if the overlying film possesses isotropic stress. This effect is demonstrated in Fig. 16, which depicts the difference in the curvature of Si substrates with (001) and (011) orientations due to an overlying, 0.5 $\mu$m thick Si$_3$N$_4$ film. As can be seen, the residual stress of the amorphous Si$_3$N$_4$ film induces a curvature displaying two-fold symmetry in the Si...
The greater elastic stiffness along \{011\} produces minima in the curvature (~23.9 km\(^{-1}\)) relative to that (~34.4 km\(^{-1}\)) along the more compliant \{100\} in Si (011), resulting in a difference of approximately 44% along these two principal directions. In contrast, a similar Si\(_3\)N\(_4\) film deposited on a Si (001) wafer induces a radially symmetric curvature of approximately ~36.9 km\(^{-1}\) even though the in-plane elastic stiffness within (001) varies by approximately 30%; it is the symmetry of the wafer’s crystallographic orientation that dictates distributions in the curvature. Comparisons to an analytical model developed for arbitrary orientations within anisotropic thin films and substrates\(^{213}\) allow for a quantification of Si\(_3\)N\(_4\) film stress of 1.19 and 1.14 GPa on the Si (001) and (011) substrates, respectively.

2. Stress distributions in CMOS features

While wafer curvature measurements provide data on the stress state within blanket stressor films that provide the driving force for elastic deformation, it is necessary to employ techniques with a greater spatial resolution to assess strain distributions that arise within and near CMOS devices. Several approaches have been applied to this task, including micro-Raman microscopy\(^{214,215}\), TEM\(^{207,216}\), and synchrotron-based microdiffraction.\(^{217,218}\) Micro-Raman spectroscopy measures stress indirectly through a correlation of phonon frequency shifts with those for samples with a known stress state (uniaxial, biaxial, etc.). However, laser-induced heating of the sample can artificially induce such frequency shifts,\(^{215}\) particularly in silicon-on-insulator (SOI) layers. TEM-based techniques clearly require significant sample preparation to produce electron-transparent specimens that can also modify the original stress state in the features under investigation. Among the techniques that allow for the in situ measurement of strain at a submicron scale, synchrotron-based x-ray microbeam diffraction is best suited to determining the elastic strain tensor components.

Strain fields in SOI regions underneath Si\(_3\)N\(_4\) stressor features have been shown to extend up to 40 times the thickness of the stressor.\(^{219}\) Figure 17(a) depicts a 0.84 \(\mu\)m wide, 105 nm thick Si\(_3\)N\(_4\) feature deposited on a SOI substrate, in which the free edges of the feature induce stress in the underlying layers. Out-of-plane SOI strain was determined by measuring the change in the (008) reflection from the SOI region using an x-ray microbeam width of approximately 0.25 \(\mu\)m and a photon energy of 11.2 keV. Depth-averaged strain distributions within the SOI layer were compared with elastic simulations based on an anisotropic edge force model\(^{220}\) and the boundary element method (BEM\(^{219}\)) in Fig. 17(b). Fitting of the models to the observed strain profiles yielded a compressive stress value (\(\sigma_3\)) of ~2.5 GPa in the unrelaxed Si\(_3\)N\(_4\), equivalent to that found by wafer curvature measurements performed on blanket wafers. As shown in Fig. 17(b), the BEM model produces a better representation of the observed strain distributions and confirms the fully elastic strain transfer between the Si\(_3\)N\(_4\) feature and SOI although the sharp strain gradients near the feature corners are not captured by the measurements due to the finite size of the x-ray beam. The out-of-plane compressive strain underneath the Si\(_3\)N\(_4\) stressor feature corresponds to in-plane tensile strain due to Poisson contraction, whereas the portions of the SOI outside of the feature possess in-plane compressive strain, which would correspond to the device channel regions. However, similar strain mapping across wider Si\(_3\)N\(_4\) features revealed that the magnitude of strain induced in the SOI decreased with the decreasing stressor width,\(^{221}\) signaling less efficacy in strain transfer for smaller device sizes. Microdiffraction measurements performed on SOI device channels and adjacent, embedded Si\(_{1-x}\)C\(_x\) stressor features show that the elastic strain is also fully transferred\(^{222}\) and suggest a more effective strategy as device dimensions decrease.

3. Stress gradients within microelectronic metallization

A final example that illustrates the effects of stress in microelectronic technology is the presence of stress gradients within the metallization that provides power to CMOS devices. It has been known that large current densities within metallic lines can be sufficient to generate atomic mass flow, or electromigration, and ultimately create voids leading to open circuit failures.\(^{223}\) Strategies developed over the past five decades to reduce diffusion paths within the interconnect microstructure involve incorporating selected solute atoms that hinder flux of the metallization and minimizing the number of GBs that are aligned parallel to current flow. However, such approaches have not been effective at limiting the decrease in the median lifetime of copper-based interconnects, which is halved in every new generation of CMOS technology, where device density is also doubled.\(^{224}\) Novel capping layers have been explored to mitigate...
diffusion along the top surface of the metallization, a location that is highly susceptible to electromigration. Capping films are required to limit oxidation and impurity diffusion into interconnects during manufacture. However, the fabrication of copper metallization contains numerous processing steps that expose the interconnects to high temperatures. The mismatch in coefficients of thermal expansion between copper and the other materials that comprise the back-end-of-line levels can induce significant tensile stress in the metallization due to thermal cycling. In particular, plastic deformation generated within Cu can result in tensile stress gradients that can double the total stress near the cap/metallic film interface relative to the bulk stress in the film. Because greater tensile stress can exacerbate voiding in interconnect structures, the accurate measurement of stress distributions within these features is critical to understanding their vulnerability to electromigration.

Glancing-incidence x-ray diffraction (GIXRD) can be used to nondestructively probe strain gradients within metallic films. By employing angles of the incident and diffracted x-ray beam with respect to the top sample surface close to the critical angle for total external reflection, we can control the depth to which diffraction information is collected, as already mentioned in Sec. II B. Using the lattice spacings measured by GIXRD in combination with conventional $d_{hkl}$ vs $\sin^2(\psi)$ stress techniques for a particular $hkl$ reflection, stress gradients that develop near the top surface of crystalline films can be quantified relative to the bulk, in-plane stress. Figure 18(b) depicts the stress gradient extracted from GIXRD measurements using the (220) reflection from a 2.2 $\mu$m thick Cu film capped with a 25 nm SiC$_3$N$_4$H$_2$ layer, deposited at a temperature of approximately 350 $^\circ$C. The constraint imposed by the overlying cap during such thermal excursions is responsible for a greater tensile stress near the cap/Cu interface relative to that deeper in the Cu film where plastic relaxation is more prevalent. Although the stress gradients, $\Delta \sigma$, are collected as a function of incidence angle, it is important to transform these distributions to ones as a function of depth within the film.

In order to accomplish this, the distribution in x-ray beam intensity, which decreases exponentially with the depth due to photoelectric absorption, must be deconvoluted from the measured shift in lattice spacing. However, difficulties in solving this inverse problem can result in significant error in obtaining accurate stress profiles. To better illustrate this issue, we superimpose on the measurements in Fig. 18(b) the predicted curves corresponding to two hypothetical stress gradients, as depicted in Fig. 18(a). One distribution assumes a constant value of $\Delta \sigma$ to a depth of 70 nm followed by an exponential decrease in the stress gradient (decay length of 39 nm) to the bulk stress value, while the other possesses a constant $\Delta \sigma$ that extends to a depth of 105 nm. Least-squares fitting was used to identify these dimensions as well as the maxima in the stress gradients, $\sigma_0$, of 125 MPa. Although the stress gradient distribution with an exponential decay exhibited a lower fitting error, both profiles indicate a depth in the range of 100–200 nm below the cap/Cu interface over which plastic deformation is affected by the overlying cap and essentially overlap when plotted as a function of incidence angle [Fig. 18(b)], highlighting the challenges in finding a unique stress gradient.

4. Summary—Stress in microfabricated systems

Strain distributions within microelectronic features influence many aspects of device performance, from increased carrier mobility to greater susceptibility to electromigration-based failure. Extracting stress distributions within these features through complementary techniques affords us a better understanding of how such distributions impact overall behavior, as well as how adjacent strain fields interact with each other as device density increases. With the introduction of more complex device geometries and smaller feature sizes, future characterization methods must be able to investigate all components of the stress tensor in three dimensions within key regions of these features to ensure that strain in the entire composite can be properly tailored to enhance the key properties essential for these future devices.
In order to qualify for optical applications, the following criteria should be respected when choosing the appropriate film material and film deposition process:

1. Optical coating requirements

In order to fully describe the behavior of OC in relation to the fabrication process as well as to the effect of environmental conditions and the conditions of use, \( \sigma_{\text{tot}} \) is usually expressed as (e.g., Ref. 56)

\[
\sigma_{\text{tot}} = \sigma_{\text{int}} + \sigma_{\text{therm}} + \sigma_{\text{env}},
\]

in relation to the intrinsic stress (\( \sigma_{\text{int}} \)), thermal stress (\( \sigma_{\text{therm}} \)), and the stress due to the sorption of gases and vapor from the surrounding environment within the pores and internal cavities (\( \sigma_{\text{env}} \)).

A classical example of the evolution of the total stress in the most typical optical film materials, namely, SiO\(_2\), during the fabrication cycle using evaporation is shown in Fig. 19. During deposition, the total compressive stress increases with the increasing thickness. At the end of the deposition cycle, the Meissner trap is degassed, leading to partial filling of the film’s pores, accompanied by a slight increase in the compressive stress. This stress further significantly rises due to the sorption of water vapor during venting, followed by a slow stress relief due to the microstructural relaxation.

The effect of water sorption has been studied in detail by Hirsch, who proposed a model in which the propensity to the accommodation of water molecules is related to the size of cylindrical pores and the adsorbent dipole moment. In relation to the water vapor sorption phenomenon, much of

---

**B. Stress in optical coatings**

Advances in the physics and technology of thin films have significantly been stimulated by their use in optical systems for numerous conventional and high-tech applications. This most frequently includes transparent dielectric OIFs, attractive forces within pores lead to tensile stress, while gas or vapor entrapment and its physio-chemical isomerization in inner cavities or at GBs lead to compressive stress. In practical situations, the total internal stress, \( \sigma_{\text{tot}} \), in OC is determined from bending curvature measurements: The curvature \( \kappa \) is obtained by capacitance, electromechanical, interferometric, and other measurements, usually using a circular plate or a cantilever beam, and \( \sigma_{\text{tot}} \) is then calculated from the Stoney formula, as described and discussed in detail in Sec. II A.

In order to fully describe the behavior of OC in relation to the fabrication process as well as to the effect of environmental conditions and the conditions of use, \( \sigma_{\text{tot}} \) is usually expressed as (e.g., Ref. 56)

\[
\sigma_{\text{tot}} = \sigma_{\text{int}} + \sigma_{\text{therm}} + \sigma_{\text{env}},
\]

in relation to the intrinsic stress (\( \sigma_{\text{int}} \)), thermal stress (\( \sigma_{\text{therm}} \)), and the stress due to the sorption of gases and vapor from the surrounding environment within the pores and internal cavities (\( \sigma_{\text{env}} \)).

A classical example of the evolution of the total stress in the most typical optical film materials, namely, SiO\(_2\), during the fabrication cycle using evaporation is shown in Fig. 19. During deposition, the total compressive stress increases with the increasing thickness. At the end of the deposition cycle, the Meissner trap is degassed, leading to partial filling of the film’s pores, accompanied by a slight increase in the compressive stress. This stress further significantly rises due to the sorption of water vapor during venting, followed by a slow stress relief due to the microstructural relaxation.

The effect of water sorption has been studied in detail by Hirsch, who proposed a model in which the propensity to the accommodation of water molecules is related to the size of cylindrical pores and the adsorbent dipole moment. In relation to the water vapor sorption phenomenon, much of
the effort in the field of OC has been devoted to the suppression of porosity by applying appropriate energetic conditions, leading to high film packing density (see, e.g., Refs. 225, 232, and 236).

Since most of the OCs appear in the form of discrete multilayer stacks of \( n_{\text{Ht}}, n_{\text{Ls}}, \) and \( n_{\text{Ms}} \) layers, it is generally accepted that the final stress, represented by the coating/substrate curvature \( \kappa \), scales with the stress of individual layers \( \sigma_i \) weighted by the corresponding film thickness \( h_i \) (e.g., Refs. 237 and 238)

\[
\kappa = \frac{6(1-\nu_f)}{E_f h_f^2} \left[ \sigma_1 h_1 + \sigma_2 h_2 + \cdots + \sigma_n h_n \right],
\]

where \( \nu_f \) is the Poisson ratio, \( E_f \) the Young’s modulus, and \( h_i \) substrate thickness.

It has been demonstrated that the curvature method generally used for the assessment of the stress can also be effectively applied for the determination of the films’ coefficient of thermal expansion (CTE or \( \alpha_f \)) as well as of the Poisson ratio (\( \nu_f \)) using the two-substrate method.\(^{239}\) In this approach, the film is simultaneously deposited on two different substrates (1 and 2) with their known (bulk) CTE (\( \alpha_{s1} \) and \( \alpha_{s2} \)), while assuming that the microstructures are the same. Taking into account the thermal stress for the two individual coating/substrate combinations\(^{240}\)

\[
\sigma(T) = (\alpha_f - \alpha_i) \left( \frac{E_f}{1-\nu_f} \right) (T - T_i),
\]

where \( T_i \) is the initial temperature, measuring the variation of stress with temperature \( (d\sigma/dT) \) for individual substrates \( s_1 \) and \( s_2 \) can then be used to derive \( \alpha_f \) and \( \nu_f \) from the two following equations:

\[
\alpha_f = \frac{\alpha_{s1} \left( \frac{d\sigma}{dT} \right)_{s1} - \alpha_{s2} \left( \frac{d\sigma}{dT} \right)_{s2}}{(\frac{d\sigma}{dT})_{s1} - (\frac{d\sigma}{dT})_{s2}},
\]

\[
\nu_f = \frac{1}{E_f} \left( 1 - \nu_i^2 \right) \left( \frac{d\sigma}{dT} \right)_{s1} - \frac{1}{E_i} \left( \frac{d\sigma}{dT} \right)_{s2} - 1.
\]

The only unknown in the latter equation is the reduced Young’s modulus of the film (\( E_f \)) that can be obtained from depth-sensing indentation measurements, while \( \nu_i \) and \( E_i \) are the Poisson ratio and the Young’s modulus of the diamond indenter. As an example, using c-Si and GaAs wafers with their known CTE values of \( \alpha_{s1} = 2.6 \times 10^{-6}\)°C\(^{-1} \) and \( \alpha_{s2} = 5.12 \times 10^{-6}\)°C\(^{-1} \), this approach yielded \( \alpha_f = 4.9 \times 10^{-6} \)°C\(^{-1} \) and \( \nu_f = 0.22 \) for Nb\(_2\)O\(_5\), \( \alpha_f = 4.4 \times 10^{-6} \)°C\(^{-1} \) and \( \nu_f = 0.27 \) for Ta\(_2\)O\(_5\), and \( \alpha_f = 2.1 \times 10^{-6} \)°C\(^{-1} \) and \( \nu_f = 0.11 \) for SiO\(_2\) obtained by dual ion beam sputtering (DIBS) (for more details and references, see Ref. 240). More recently, this technique has also been applied for the assessment of the CTE and \( \nu \) values of the novel hybrid (organic–inorganic) optical coatings prepared by ion beam assisted chemical vapor deposition (IBA-CVD).\(^{241}\)

A new in situ real-time approach to perform and analyze scratch tests in relation to stress in transparent coating/substrate systems has recently been introduced.\(^{242}\) This method allows one to observe, in real-time, the contact region during the scratching process. As an example, thin TiO\(_2\) layers exhibiting stress levels ranging from tensile to compressive were deposited by ion beam assisted deposition (IBAD) onto plastic substrates. Failure processes obtained using increasing (and novel) decreasing load scratch sequences were linked to the internal stress in the coatings, allowing one to derive a stress management diagram and to evaluate the yield stress of the layers. Such an approach helped to enhance understanding of the OC failure mechanisms, but it also outlined a new pathway to increase scratch measurement reproducibility.

In Sec. IV B 2, we will discuss the relation between the fabrication techniques and the mechanical stress and stress evolution in OC.

2. Effect of the fabrication technique on stress in optical coatings

OC has traditionally been fabricated by PVD techniques from a solid source; this includes thermal and EBE, MS, ion beam sputtering and DIBS, and filtered CAD. Advances in low pressure plasma processing, and in plasma enhanced chemical vapor deposition (PECVD) in particular, have greatly increased the interest in PECVD for the fabrication of OC, while novel techniques such as atomic layer deposition (ALD) have also been considered.

As already mentioned above, considerable effort has been devoted to ensuring film densification in order to obtain high packing density. This can be accomplished by increasing substrate temperature, \( T_s \), or by an appropriate control of ion bombardment energy, \( E_b \), typically below 1 keV. The film microstructure obtained by such processes can be well described by the SZM, first proposed for evaporated metals\(^{243}\) and for MS.\(^{225,244}\) The SZM has then been revised by Messier et al.\(^{245}\) who showed that \( T_s \) required for high packing density can be decreased by superimposed ion bombardment due to growth-related effects such as interfacial atom
mixing, high surface mobility (diffusion) of deposited species, resputtering of loosely bound species, and deeper penetration of ions below the surface, leading to the displacement of atoms (forward sputtering or knock-in effects). This has since been clearly demonstrated for OC and other coatings prepared by MS as well as by PECVD. Different approaches have been used to quantitatively describe ion bombardment and its effect on the film microstructure and properties (also see Sec. III B). It appears that a key parameter for describing such effects is the energy, $E_{\text{dep}}$, delivered to the growing film per deposited particle that, based on Eq. (19), can be expressed as $E_{\text{dep}} \sim E_i \Phi_i/\Phi_m$, where $E$ denotes the energy, $\Phi$ denotes the particle flux, and indices $i$ and $m$ refer to ions and neutrals, with the latter one including condensing precursor species. Generally, $\Phi_m$ can be experimentally determined as $\Phi_m = R \rho N_A/m_A$, where $R$ is the deposition rate, $\rho$ is the density, $N_A$ is Avogadro’s number, and $m_A$ is the atomic mass. These relationships are also reflected in the recently updated structure-zone diagram that includes the flux of energetic particles as one of the key parameters.

Typical values of stress for the most frequently used OC prepared by PVD (Ref. 56) and PECVD have been summarized; they typically range from −500 to −100 MPa in compression for SiO$_2$, Ta$_2$O$_5$, and TiO$_2$, while they are around +400 MPa for MgF$_2$. Stress in PECVD optical films is generally lower compared to PVD. For example, using a dual-mode microwave/radiofrequency plasma for the deposition of SiO$_2$, SiN$_{1.3}$, SiON, and TiO$_2$ films, the stress could be adjusted from tensile (around +200 MPa) to compressive (between −100 and −500 MPa), with zero stress values obtained for a substrate bias voltage of about −150 V. Such a transition has been associated with the conversion of a highly porous structure to a dense structure related to the energies required for the displacement of individual atoms constituting the film. This phenomenon has been explored for the fabrication of porous/dense single-material discrete layer and inhomogeneous layer (rugate) filters that have been proposed as all-optical gas and vapor sensors.

The performance of hard DLC ($\alpha$-C:H) films obtained from hydrocarbon gases or vapors under high energy ion bombardment has frequently been related to (and limited by) high residual compressive stress. Such a high stress has been associated with the incorporation of hydrogen that is not chemically bonded, but it is physi- or chemisorbed on inner surfaces or trapped in the voids. It has been shown that significantly lower stress values are obtained for high $\Phi_i/\Phi_m$ ion flux ratios but lower ion energy, mainly due to an enhanced surface mobility and reduced subplantation, as also supported by MD simulations.

Over the years, there has always been a debate about the effect of the deposition technique on the performance of OCs, especially their optical and mechanical properties. In response to that, Klemberg-Sapieha et al. performed a systematic round robin study on the comparison of the most typical $n_H$ and $n_L$ OC, namely, SiO$_2$, Ta$_2$O$_5$, and TiO$_2$ using PVD (IBAD, DCMS, HiPIMS, and FCAD) and PECVD. It has been found that following optimization for the best performance (highest packing density), the optical properties were very similar (or almost identical), while most of the differences were in terms of the mechanical properties, such as hardness, Young’s modulus, and particularly stress (see Fig. 20). It has been concluded that lower $\sigma_{\text{tot}}$ values are obtained at higher $\Phi_i/\Phi_m$ values. This is also further supported by a recent study using reactive HiPIMS where the lowest stress for SiO$_2$ and Ta$_2$O$_5$ films compared to other techniques has clearly been demonstrated (see Fig. 20). In addition, multilayer OIFs, fully fabricated by HiPIMS, were also found to exhibit significantly lower residual stress levels compared to single layers as well as to other techniques, an effect accompanied by a substantially increased scratch resistance, and high environmental and long-term stability.

This is also in agreement with a study of the effect of $E_i$ on Nb$_2$O$_5$ and Ta$_2$O$_5$ $n_H$ films prepared by DIBS in which $\sigma$ has systematically been compressive, but partial stress relaxation has been observed for $E_i$ above about 250 eV. This is in line with a model according to which the process of stress reduction consists of two stages (e.g., Ref. 156): (1) the atoms initially implanted in the film are in metastable positions, while (2) some fraction of the incoming ion energy is transferred to those atoms, releasing them by a thermal spike from their metastable positions, moving them to a stable location.

3. Strategies to manage stress in optical coatings

The performance of OC systems and devices may be negatively affected by the deleterious effects of mechanical stress. In this section, we discuss several strategies that allow one to decrease or compensate stress in the OIF applications, namely, (1) use of new or novel film materials, (2) implementing minimum stress as a target in optical filter design, and (3) multilayer stack release to obtain free-standing optical filters.

![Fig. 20. (Color online) Comparison of (optimized) residual stress in low index (SiO$_2$) and high index (Ta$_2$O$_5$) films prepared by different complementary methods including IBAD, pulsed DCMS, DIBS, and PECVD (according to Ref. 253). The values for films prepared by HiPIMS are taken from Ref. 254.](image)
Application of inorganic OIF on plastic substrates such as in ophthalmic applications can be limited by the incompatibility of the coating and the substrate material due to a large difference in CTE, possibly leading to cracking, delamination, and reduced durability. In this context, hybrid films have recently been proposed and explored, fabricated by IBA-CVD, in which an organic precursor is introduced into the path of an ion beam. Such hybrid films have been found to possess optical properties suitable for OIF applications, while providing mechanical performance enhancement due to high mechanical flexibility, high hardness/modulus ratio, and high durability of the plastic optical components, leading, for example, to decreased “crazing” (visible fracture patterns). In addition, such hybrid films are substantially less prone to water vapor sorption compared to their inorganic counterparts. This effect has important consequences on the CTE differences therein), each individual optical layer is represented by its CTE contribution to the total stress as illustrated in Fig. 21 demonstrated by a significantly reduced difference between the stress before and after venting, $\Delta\sigma$.

Excessive total stress in OIF can exhibit a negative effect on the performance of optical components such as astronomical mirrors due to substrate bending. This can be compensated by depositing equivalent film thicknesses on both sides of the substrate by taking into account the expression for the total stress of the OIF and for the individual layers.\(^{257}\)

Recently, a new approach has been demonstrated that simultaneously considers both optical and mechanical targets in the OIF design.\(^{258}\) Using a supplemental module in the OpenFilters open source design software (Ref. 258 and references therein), each individual optical layer is represented by its refractive index $(n_{\text{hi}}, n_{\text{li}})$ and by its stress. This combined optical/mechanical optimization and design technique has been experimentally validated by fabricating multilayer stacks using EBE, in combination with their mechanical stress assessment performed as a function of temperature.

Two different stress-compensation strategies have been compared including (1) a design of two complementary OIFs on either side of the substrate and (2) implementing the mechanical properties of the individual materials in the design of the OIF on one side only. This approach has been tested by the manufacture of a Fabry–Perot etalon used in astronomy; using evaporated SiO$_2$ and TiO$_2$ films, the substrate curvature could be decreased by 85% and 49% for the first and second strategies, respectively.

OIFs have found widespread use in the fiber optic telecommunication industry. For such applications, they must be thermally stable, exhibit excellent spectral performance, have a low polarization dependence, and must be readily produced in high volume.\(^{260}\) One of the frequent ways to compensate stress is OIF annealing for stress relaxation.\(^{261}\) However, when pushed toward an increased complexity of the fiber optic network architecture, more efficient components require more complex filters, which imply thicker coating designs and smaller final filter size. In such cases, coating stress at the edges of the filter is relieved through bending or distortion of the substrate, and the coating becomes thinner. The coating at the center of the filter retains a high stress condition and remains thicker. The thickness difference leads to the central wavelength variation and the resulting passband width performance degradation.

As a remedy to the above problem, it has been proposed to release the coating after deposition and to reattach it to a suitable substrate, giving rise to an OIF with an ultralow stress (see Fig. 22). Specifically, the OIF is first fabricated on a sacrificial substrate provided with a moisture sensitive surface. The coating is then released by scratching a small part of the disk with a diamond tip: this causes moisture to penetrate to the NaCl layer and spontaneously (\(<100\text{ ms}\)) release the dielectric filter. The resulting coating flakes are flat and have sizes of up to 60 cm$^2$. Since these freestanding filters have a large thermal central wavelength shift (typically 9 pm$^\circ\text{C}$), too high for telecommunication applications,\(^{260}\) the freestanding filter is then attached onto either a high-CTE glass or a stainless-steel washer with the appropriate CTE (see Fig. 22). This process made it possible to produce high performance filters (200, 100, 50 GHz, etc.) with smaller physical dimensions, while maintaining or improving optical performance.

### C. Stress in the coatings on plastics

The topic of stress in thin film-coated structures would not be complete without an examination of the specific challenges and opportunities of coating polymer components, many of which enable the light-weight, robust, mass-producible products that surround us. While the underlying physics is not different from stress-related issues at other interfaces and substrates, this section looks at particular systemic influences on stress for two popular polymer substrate configurations: roll-coated polymer webs and injection-molded polymers. The particular origin of stress-related issues is related to the inherent difference between the CTE of polymers compared to inorganic coatings by about 2
orders of magnitude and the phenomena related to the complexity of the interfacial region (or “interphase”) between the two materials.55

1. Stress on roll-to-roll–coated polymer webs

The mechanics of a coating on a thin polymer web (e.g., for a window film, as thin as a few micrometers) introduces significant application effects at stress levels that would be considered benign when compared to coatings on glass or other rigid substrates. As an example, the Stoney formula indicates a significant deflection of a 25 μm-thick substrate (with a Young’s modulus $E = 4$ GPa) with a 200 nm thick coating with a residual compressive coating stress of about 60 MPa, giving rise to a 50 mm radius of curvature. It is easy to see how such a deflection may affect the installability of a window film, for example, if there were not effective ways to mitigate it for a successful product. Similarly, it explains why brittle vacuum-deposited coatings on thin substrates can be quite prone to cracking, resulting in delamination or crazing defects if not designed properly.

The resulting effect on a product’s utility can be quite severe, as the following examples demonstrate: Structurally perfect barrier coatings (such as those produced by ALD) based on dielectrics can crack under stress and create diffusion paths that diminish the barrier function; window films with brittle dielectric coatings can incur “crazing”; electrically conductive coatings may exhibit much higher sheet resistances due to microfractures; the permissible bending radius of films with large layer counts may be so large that it inhibits handling, installation, or further downstream processing; high-k insulating films may incur reduced electric breakdown strength due to stress-induced microfractures. Therefore, understanding of the sources of stress and availability of the fundamental mitigation strategies for industrial applications are very important.

2. Sources of stress in coated polymer webs

While some stress-inducing mechanisms are the same as for any other substrates (nucleation/film growth effects, differences in CTE), polymer webs incur stresses related to web handling and due to the thermal (i.e., primary physical variable) conditions during the coating process.262 A very common substrate for window film products is biaxially oriented polyethylene terephthalate (BOPET) which requires bulk processing temperatures below the “glass transition temperature” ($T_g$) of approximately 70 °C to retain its dimensional and mechanical characteristics. Coating at higher temperatures will generally (especially when unconstrained) induce shrinkage that was designed into the material during the stretching process and thus impose compressive stress on any coatings. Predicting the behavior of releasing the shrinkage of a BOPET during a constrained (e.g., web coating) operation requires very thorough knowledge of the material’s processing history and web coating conditions and will complicate the resulting substrate/coating stress regime. While the CTE difference between polymers and typical vacuum-deposited coatings can be quite large, the stress influence is frequently relatively benign due to the small permissible temperature excursion.

BOPET is an example for a web substrate that is produced by extrusion of a molten resin through a die and (if desired) subsequent stretching and heat-setting,263 thus
impacting heat-shrinkability functionality that is advantageous for many applications that require three-dimensional shape conformity. Some polymers (such as polyimide) require a different film manufacturing method utilizing casting from a solvent solution, while the casting process can significantly influence the structural and mechanical properties of the substrate.\textsuperscript{264} Such films are typically not heat-shrinkable.

Another significant influence on coating stress during the deposition process is web handling. Polymer webs are elastic materials, and the high anisotropy of the web geometry and often of the web manufacturing process has significant implications on stress profiles. A web coater needs to maintain a certain tension (primarily for sufficient process cooling over a chilled coating drum and also for other effects such as scratch control and roll formation), and the resulting elastic elongation in the machine direction induces a constriction and thus compressive stress in the transverse direction. This nonisotropic stress/strain profile in the uncoated substrate will be superimposed with the coating stress induced by the sequential stages of coating formation and can significantly affect the subsequent processability and characteristics of the use of the film (Fig. 23). Processing of very thick films over small-diameter rollers may impart additional stresses onto the substrate and possibly the coating and create complex load scenarios especially when real-life “noise” factors such as roller misalignment and web distortion are considered.\textsuperscript{265}

Stress mitigation efforts in web coating often focus on balancing the deposition process conditions with the web handling constraints, and more complex filter stacks are often designed as dual-sided coatings (on both sides of the substrate) for sufficient stress compensation. Very elastic films that would incur high stress anisotropy during web handling may be coated with the assistance of a more rigid carrier film. In view of the limited thermal budget due to not only the melting temperature but also typically more importantly the $T_d$ value of many polymer web substrates, this obviously eliminates classical high-temperature thermal annealing of a coating for effective recrystallization for stress relief, so a post-treatment is generally not an option, and the stress mitigation must be accomplished during the deposition process itself. For some processes, it may be possible to decouple the thermal load of a coating process from the substrate by special annealing methods such as flash lamp annealing processes.\textsuperscript{266} In multiple coating systems where the product is made from multiple layers of coated polymer films (e.g., high-end window film), the aforementioned effects of web tensioning and additional coating processes (particularly wet-coatings using solvents, as for adhesives and hardcoats) and an appropriate product design and coating sequence can be successfully employed to produce stress-neutral products from quite stressy subcomponents.\textsuperscript{262} Nevertheless, the demands on production consistency and precise process control and process monitoring are very high to maintain high production yields.

A stress mitigation strategy may also be found in substituting brittle ceramic (dielectric) coatings with polymer or polymer composite coatings that typically possess a much lower modulus and can conform better to distortions.\textsuperscript{267,268} It has been shown that the crack onset value on many organic layers is above 12% as compared to 1% for many ceramic materials, which translates to a much higher tolerance for tight bending radii. The improved stress tolerance of such polymer-based thin films and coatings may also be an enabling characteristic for successful high-performance thin film barriers for flexible polymer devices such as organic light-emitting devices.\textsuperscript{269} The approach of utilizing hybrid (organic–inorganic) layers is also a powerful stress mitigation tool for coatings on injection-molded polymer products, as is outlined in an example in Sec. IV C 3.

3. Stress in coatings on injection-molded polymers

The proliferation of high-end polymer optics relies on producing high-end coatings, especially for dielectric filters for antireflection, color correction, or other functional requirements. Coatings on polymers particularly benefited from the success of adding an antiscratch coating onto lightweight polycarbonate (PC) ophthalmic lenses, and coatings have expanded from high-performance optical filters to low-cost/mass-produced antireflection and reflector coatings. Coating adhesion and durability are frequent requirements, especially in applications when exposed to harsh environments.

Compared to web coatings, the range of substrate polymers is much broader, but the most prominent material groups are PC (including the CR39 derivative, widely used for ophthalmics), polymethyl methacrylate (PMMA/Acrylic), poly styrene, and copolymers (such as blends of acrylic and styrene, e.g., styrene acrylonitrile).\textsuperscript{270} Other common specialty polymers for coated optics are polyetherimide (“Ultem®”) and cyclo-olefin polymers (“Zeonex®”). Polymers are selected due to their properties such as refractive index, spectral transmission, UV sensitivities, mechanical properties, thermal tolerance, molding process properties, and cost just to name a few.

High-end optical engineering polymers exhibit high compositional stability that is primarily mandated by the requirement
of stable and predictable optical performance. In some low-cost polymers, oligomer migration over time can affect long-term interface properties. Oligomers are chain fragments or short monomer structures that may have considerable mobility and different physical properties compared to the bulk. They may occur as a side effect of injection molding or can be intentionally added to modify bulk properties; for example, thermoplastics often contain ester oligomers as plasticizers to make them more impact-resistant. The issue with oligomer migration with respect to stress is that the oligomers accumulating at the component surface may act as a (largely unintended and uncontrolled) release layer that weakens the substrate–coating interface, and delamination can occur over time even though initial adhesion was quite sufficient. Similarly, additives such as flow enhancers or mold release agents may also be sources of interface “skins” that typically degrade the stress tolerance at the substrate/coating interface.

Most injection-molded polymers are compatible with a wide range of coating processes as long as thermal and chemical limitations are observed and exhibit good adhesion properties that tolerate quite high stress levels. PMMA, however, has been shown to be degraded by deposition process-related plasma radiation, resulting in poor adhesion and requiring special coating methods. Satisfactory coating adhesion on PMMA requires either a plasma-free vacuum coating process (such as thermal evaporation or IBAD) or the use of a radiation protection coating.

The molding conditions of a component have a major influence on the surface characteristics presented to a coating. During the molding process, the polymer is subjected to very high pressures, temperatures, and often high shear rates that can significantly alter the component’s chemical and mechanical properties from the bulk polymer resin, and especially, the cooling profile (the outer skin solidifies first, while the core is still in the melt stage) and bulk geometry can introduce significant compositional material gradients that may vary on different areas even within the same injection-molded part. High-precision optical polymer components may be produced with injection compression molding, which not only greatly improves the fidelity of the optical contour to the mold but also imparts higher stresses and relaxation phenomena. Advanced mold design, flow modeling, and molding process control make it possible to control material variations, and the same conditions that support good optical properties of an injection-molded lens typically tend to provide uniform coating interface conditions for the functional surfaces, creating more predictable interface conditions for managing coating stress.

Another notable influence on coating stress in injection-molded parts is the geometry which may induce large differences in the angle-of-incidence in directional coating processes. Different ones, such as MS, impart film thickness and sometimes film density and nonuniformity on optics with high-aspect ratio contours, which affects the optical performance as well as the interfacial stresses at the film–substrate interface. The assumption here would be that the highest compressive stress levels (due to thickness and bulk density) would be at surfaces normal to the incident deposition flux, whereas off-angle geometry features exhibit lower density and lower thickness and thus lower interfacial stresses. This can make the design of stress-compensated precision optics more complicated. It is recognized that conformal deposition processes, such as PECVD (Refs. 11 and 54) and lately particularly ALD (Ref. 274), present an attractive option to address this issue although it remains to be seen whether the elevated deposition temperatures required for many of the attractive coating materials can be lowered to open up the application space of injection-molded polymer optics.

When it comes to the choice of coating material systems, particular improvement of the performance of antireflective stacks on plastic lenses has been demonstrated by implementing hybrid (organic–inorganic) layers in the optical stack. It was demonstrated that such hybrid systems provide a significantly improved resistance to temperature- and humidity-induced stress variation due to a higher elastic recovery of hybrid (SiO2-Chx, ZrO2-Chx) coatings compared to their inorganic counterparts. Such coatings prepared by IBA-CVD show a higher thermal expansion close to that of polymer substrates and a relatively high H/E ratio (up to 0.16), as well as high durability following accelerated environmental tests including exposure to high temperature/high humidity, UV and solar radiation, as well as a saline solution.

Important improvement in the performance of inorganic films on polymer substrates has been demonstrated by specifically tuning the surface reactions when applying plasma-based deposition techniques. Surface interactions involving energetic photons, ions, and reactive species (free radicals) have been found to lead to the formation of a physically thick structured interfacial region (interphase) between the inorganic coating and the plastic substrate (see Fig. 24). Similar depth profiles have been observed for different combinations of materials, including SiO2 and SiN1.3 on polyethylene terephthalate, PC, PMMA, and other polymers. The interphase has been found to be up to several tens of nanometers thick; it consists of a cross-linked layer [region (i) in Fig. 24], followed by a transition layer [region (ii)] formed by intermixing the growing film with the substrate materials and possibly by voids. In the case of SiN1.3 shown as an example in Fig. 24 and aimed for antireflective or barrier applications, n increases from 1.59 for PC to 1.80, while hardness increases by 2 orders of magnitude (from 0.2 GPa for bulk PC to about 2 GPa for the cross-linked surface layer and up to 18 GPa for SiN1.3). This inhomogeneity generally leads to a more uniform stress distribution at the interface, giving rise to better adhesion, tribological properties, flexibility, stretchability, and other functional characteristics suitable for coated plastics.

4. Summary—Stress in the coatings on plastic

Stress issues for coatings on polymers can lead to significant functional failures at relatively low stress levels compared to other more rigid substrates. The manufacturing
conditions for polymer web or molded polymer components can induce complex mechanical and chemical conditions at the coating/substrate interface, and for thin polymer webs, the coating may become a significant structural/mechanical component of the composite. Successful stress management of coated polymer products depends on a thorough understanding of the interplay between the material properties of the coating and the substrate as well as the processing conditions, both during coating deposition as well as postprocessing.

D. Effect of stress on the performance of tribological coatings for aerospace applications

Tribological coatings are employed to mitigate wear-related issues in a variety of aerospace components. The ability of coatings to meet the demands of specific applications depends many times upon the relationship between the internal coating stress and the applied contact stress. However, the relationship between stress and the tribological performance is not well understood. Yet, coatings are expected to achieve sufficient tribological performances, especially the components that experience high contact stresses. In order to illustrate this situation, in the following, we describe three examples of critical applications where stress must be managed to obtain sufficient tribological characteristics, namely, landing wheel bearings, rotorcraft gear-box components, and rotating elements in satellite guidance systems.

1. Landing gear bearings

Landing gears on aircraft are equipped with grease-lubricated, tapered roller bearings. Tapered roller bearings are designed to simultaneously accommodate high radial and axial loads. Because of the differences between the inner and outer raceway angles, a force component is generated during operation that drives the tapered rollers against the rib face in a sliding contact. At touchdown, the rotational speed of the bearings accelerates instantaneously from rest to accommodate the landing speed of the aircraft. Since the grease is usually very cold at the time of landing, the amount of lubrication from the bleed of oil out of the grease is initially very small. When extremely large axial forces are involved in the landing, such as aircraft landing on the decks of aircraft carriers, very large loadings can be generated at the rib face/roller end sliding contact. Large contact stresses applied to a poorly lubricated, rapidly and highly loaded sliding contact bring asperities of mating surfaces (in this case, the roller ends and rib face) into intimate contact, and scuffing or galling type wear can ensue.

Currently, a metal-doped amorphous carbon coating is being used in some of these applications to reduce the risk of scuffing at the roller/rib contact of landing gear bearings. Specifically, the coating is a nanolaminate material consisting of nanocrystalline β-TiC precipitates in an amorphous carbon matrix, i.e., TiC/a-C.

Several application-specific requirements were considered in the selection of this coating. Because of the large shear stresses applied at the rib/roller interface during landing, the intrinsic (compressive) stress in the coating needed to be <1 GPa and therefore managed during deposition. The deposition process chosen to apply TiC/a-C was closed-field unbalanced MS, which is a process that can generate large argon ion fluxes (Φ) to the substrate during coating growth. The maximum temperature during deposition and the indentation modulus (Y) of the coating had to be compatible with the tempering point and elastic modulus of the steel, ~180°C and 210 GPa, respectively. To accommodate the shear stresses during the initial contact, the dry sliding friction coefficient of the coating was required to be small (μ ~ 0.06) and the interfacial shear strength required to be large (τ ~ 45 MPa). Finally, it was desirable that the coating wear rate should be as low as possible. An Archard-type wear model would suggest that the coating should be as hard as possible. However, hard coatings have large elastic moduli, which are undesirable for this type of application since (1) large indentation modulus coatings can have large intrinsic stresses and (2) large modulus coatings tend to wear by fracture due to columnar morphologies. An alternative method to achieving low wear rates is to avoid wear by columnar fracture of the coating during sliding contact, engineer the coating to be slightly harder than its counterpart, and use a coating that establishes a stable velocity accommodation mode (VAM) in the tribological contact.

Although TiC/a-C has a low indentation modulus (110 GPa), it can develop large intrinsic stresses during deposition since it does not have a columnar morphology, and as with sputter-deposited coatings, the magnitude of the intrinsic stress increases with the coating thickness. It was found that a ~1 μm TiC/a-C coating had a compressive stress of ~1.2 GPa, a hardness of 8 GPa slightly larger than that of the steel rib face (~7 GPa), and a TiC/a-C/steel interface that displays an excellent VAM through the creation of a relatively thick and durable graphitic transfer layer in the dry and partially lubricated sliding contact.
In a demanding field test where standard tapered roller bearings were able to achieve at most 1 or 2 landings, bearings with the TiC/a-C coating applied to the roller ends were able to exceed the target of 50 landings. The coating was able to deliver the desirable tribological performance because of its ability to inhibit scuffing through the formation of a thick and durable graphitic velocity accommodation layer, its low friction coefficient against steel, the absence of columnar fracture-type wear, and its ability to withstand large shear stresses at the roller end/rib face.

2. Oil-out protection for rotorcraft bearings and gears

A sudden loss of lubrication (oil-out) in rotorcraft gearboxes produces an increase in frictional forces and a rapid temperature rise of meshing components, leading to closures of radial clearances and mechanical seizures of bearings and gears.283 One of the strategies that are being pursued to retard temperature rise in gearboxes after a sudden loss of lubrication is the use of tribological coatings. Unlike the sliding contact described in the previous example, bearings and gears in rotorcraft gearboxes function under the rolling and mixed mode (i.e., rolling/sliding) contact with high Hertzian contact stresses. The desirable functionality of a coating to provide oil-out protection to rotorcraft bearings and gears is to reduce the frictional heating of contacting surfaces after loss of lubrication, while remaining intact during normal operation. Meshing of gear teeth, roller/cage, and roller end/rib face contacts is the source of large heat generation after an oil-out occurrence. Hence, a tribological coating that is mechanically compatible with the steel components (\(Y \leq 210\) GPa) with greater hardness (\(H > 7\) GPa), a low deposition temperature (\(T < 180^\circ\)C), and a small friction coefficient is a desirable candidate for bearings and gears in these applications. A survey of the literature indicates that the family of DLC coatings can satisfy these requirements.284 Specifically, a tungsten-containing, amorphous hydrocarbon (WC/a-C:H) coating appears to be well-suited to the above coating requirements.

Similar to TiC/a-C, WC/a-C:H is a nanolaminate consisting of nanocrystalline WC precipitates in an amorphous hydrocarbon matrix and is commonly deposited by closed field unbalanced MS. First developed by Dimigen et al.,285 typical WC/a-C:H coatings utilize a thin Cr layer to provide a metallurgical bond to steel, can have thicknesses of \(~5\) \(\mu m\), are hard (\(H \sim 12\) GPa), and have an indentation modulus of \(Y \sim 156\) GPa. A \(~2.8\) \(\mu m\) thick WC/a-C:H coating was reported to have a compressive intrinsic stress of \(\sigma \sim 1.5\) GPa and a relatively low dry friction coefficient against steel after run-in (\(\mu \sim 0.2\)).281 Although WC/a-C:H functions very well as a wear-resistant coating in many applications, it sometimes cannot withstand the rigors of the rolling and mixed mode contact under high Hertzian contact stresses. For example, Fig. 25 shows a cross-sectional TEM micrograph of a WC/a-C:H coating on a roller taken from a bearing after experiencing about \(20 \times 10^6\) revolutions in boundary lubrication and at 1.5 times rated load.286 The figure shows an uneven fracture-type wear that originated in the \(~150\) nm size columns, which transformed the coating into an abrasive surface. The roughened coating surface abrasively removed the active profile of the bearing raceways and generated high local contact stresses, eventually leading to an infantile failure of the bearing.

To enable the WC/a-C:H coating to withstand high cycle contact stresses, it was necessary to eliminate the columnar growth morphology. After much effort, deposition parameters were optimized that accomplished this task. Due to the absence of the columnar morphology, a maximum coating thickness of about \(1\) \(\mu m\) was required to maintain an intrinsic compressive stress less than \(~1.2\) GPa. Bearings with the optimized WC/a-C:H coating applied to the rolling elements are reported to have significantly longer fatigue lives in debris-containing and thin-film lubrication environments and are highly resistant to other forms of wear.287

The tribological performance of the noncolumnar WC/a-C:H coating in the mixed mode contact was evaluated by Mahmoudi et al.288 Figure 26 displays the Stress Cycle to Failure (S/n) performance of the coating in boundary layer lubrication and with a 2% slide to roll ratio. The open symbols denote run-outs (i.e., no failures) while the filled symbols represent the loss of coating, and the dashed line

![Fig. 25. Cross-sectional TEM micrograph of a WC/a-C:H coating on a roller from a bearing after experiencing about \(20 \times 10^6\) revolutions in boundary lubrication and at 1.5 times rated load.](Image)
represents the estimated S/n failure criteria. Under these test conditions, the optimized WC/a-C:H coating was able to withstand high cycle Hertzian stresses of over 2 GPa without failure.

A WC/a-C:H coating with a noncolumnar morphology that is able to withstand the high cycle Hertzian contact stresses of bearings and gears during normal operation should therefore be able to reduce the frictional heating originating at the roller/cage, roller/rib face, and gear teeth contact during periods of loss of lubricant.

3. Ball bearings and articulating joints in space platforms

Mechanical components on space platforms normally operate their entire lifetime using a single charge of lubricant, where typical lubricant quantities are in the milligram range. The lubricant films are typically thin and permit extensive contact between the mating metal surfaces. Additionally, the interaction of energetic atomic oxygen in low earth orbit promotes a hardening of the lubricant through an oxidation-polymerization process, which in turn intensifies the contact between the mating metal surfaces. Without normal atmospheric oxygen, contacting asperities can easily fuse or weld together, and then, as the surfaces move against each other, the welded asperities break away and form debris particles that can damage the mechanical systems. In these situations, raceways of ball bearings that are in continuous operation become severely damaged, and the bearings may not achieve their designed lifetimes. Ball bearings that undergo oscillatory motion in these environments can develop worn grooves on the raceways that are geometrically spaced with the balls. This type of wear is known as false brinelling and can also lead to premature failures of the bearings. In articulating joints where linear or rotational sliding occurs, poor lubrication conditions greatly elevate the risk of fretting type wear and galling. A representation of the metal-metal contact that can ensue between balls and bearing raceways in a poorly lubricated environment is shown in Fig. 27.

An approach that has been adopted to address the limited lubrication environment of systems in space platforms is to inhibit metal-metal contact through the use of a tribological coating. In the late 1970’s, Hintermann et al. reported on the development of a process to apply titanium carbide (TiC) to 440C bearing balls by CVD. The deposition occurred in a reactor at 900–1050 °C, and the growth rate was about 1 μm h⁻¹. The chemical reaction was

\[
\text{TiCl}_4(g) + \text{CH}_4(g) \xrightarrow{900-1050 \degree C} \text{TiC(s)} + 4\text{HCl(g)}.
\]

During the duration of a 4 μm deposition, diffusion of the steel into the coating occurred, resulting in a metallurgical bonding and a progressive transition of physical and mechanical properties from the steel into the coating. Because of the high deposition temperatures, a significant distortion and tempering of the 440C balls also took place. Therefore, after deposition, it was necessary to retemper 440C and polish the coated balls to the desired smoothness and sphericity values. After polishing, the surface roughness of TiC was \(R_s < 0.007 \mu m\). The difference in the CTE of TiC (8 × 10⁻⁶ K⁻¹) and 440C (12 × 10⁻⁶ K⁻¹) created compressive stresses in the as-deposited coatings. The magnitude of the stress was found to depend upon the coating thickness. For example, coatings with thicknesses of 3 and 15 μm had stresses of ~2.6 and 3.0 GPa, respectively. Quenching
and tempering were not found to significantly affect the intrinsic stress of TiC on 440C. Although coatings with such large amounts of compressive stress probably could not survive the magnitude of Hertzian contact stresses (~1 GPa) required for rolling element bearings in terrestrial applications, the TiC coatings appear to work well with lightly loaded bearings utilized in the mechanical systems of space platforms.

Figure 27 illustrates the functionality of the TiC coating to separate metal contacts in poorly lubricated environments. The excellent performance of TiC coated balls in a gyro spin axis bearing has been demonstrated by McKee.289 The review on lubricants in spacecraft295 reported that the TiC coating is an excellent diffusion barrier and is therefore frequently used to prevent cold-welding and fretting damage on highly stressed stationary-vibrating contacts in hold-down and release mechanisms and latches on deployables.

E. Key aspects and complementary strategies to tailor intrinsic stress

1. Stress sources

As documented in this review, the stress in polycrystalline and amorphous films evolves in a complex way, being in most cases a balance between multiple, competing stress generation, and relaxation mechanisms. It is important to recall that the stress generation is a result of the subsequent atomic rearrangement within the film that is constrained by its attachment to the substrate. If the film is not allowed to change structurally but maintains its “frozen-in” structure, no stress will develop. Sources of intrinsic (growth) stress are numerous and are related to any strained regions due to microstructural/structural modifications taking place (1) within the film bulk (e.g., defect/impurity incorporation at the lattice or GB sites, voids, recrystallization, etc.), (2) at the film/substrate interface (lattice-mismatch, intermixing, etc.), or (3) at the growing film surface (adsorption, surface diffusion, etc.). Dislocation motion as well as diffusion processes can operate to reduce stress in the course of film growth or during subsequent processing.57 The stress can also evolve due to exposure to environmental media, as illustrated in Secs. IV B and IV D, or as a result of other extrinsic effects (chemical reaction, precipitation, plastic or creep deformation, etc.).

The examples provided throughout this article have shown the importance of thermodynamics (miscibility between constitutive elements, chemical ordering, phase stability, and phase changes); however, growth kinetics and growth energetics decisively matter in dictating the stress behavior in films obtained from condensation of vapor fluxes (which occurs under nonequilibrium conditions), especially for sputter-deposited films. Correlation between intrinsic stress and the corresponding film microstructure shows that grain size evolution, which also depends on the adatom mobility, can have a significant effect on the resulting film stress.

From these studies, different mechanisms of stress generation and relaxation have been reviewed. The kinetic model presented in Sec. III A, though not complete, compares favorably with the measured dependence on the grain size for different types of growth, the dependence on the growth rate in several systems, as well as on the thickness-dependence in patterned films. An important benefit of the model is that it provides a quantitative framework with which stress evolution under different conditions can be analyzed. It shows how different parameters may interact so that complex behavior may possibly be understood in terms of the underlying physical mechanisms and provides guidelines to develop stress engineering strategies to control the stress state in functional coatings and nanostructured films. Before addressing some routes to tailor intrinsic stress, let us summarize the possible sources of tensile and compressive stresses.

Tensile stresses set in upon crystallite coalescence31,32,136,137 (corresponding to the second stage in the CTC behavior of high-mobility materials growing in a Volmer–Weber mode) when neighboring islands grow epitaxially on the already strained film. Any film densification, resulting from grain growth,57,130,145 annihilation of excess vacancies,57 or phase transformation,41,188 also contributes to tensile stress generation.

Compressive stresses usually appear at the early growth stages, as a result of capillary forces (or surface stress) acting in the surface plane of small islands rigidly bonded to the substrate.62 The mechanisms at the origin of compressive stress during the late film formation stages (after the film continuity has been reached) are more controversial, but they are experimental evidences that the magnitude of this compression scales with the GB density. Insertion and trapping of excess atoms into GB, driven by either kinetics65,140 or entropic296 effects, reasonably explain much of the experimental findings. However, departure from the morphological equilibrium of the surface profile in the presence of an atomic flux would also be a plausible source of compressive stress build-up.60,297 Nonetheless, whole or part of this stress component is relieved once deposition is ceased, depending on adatom mobility. For the sputter-deposited film, energetic particle bombardment (atomic peening process)146,154 is an additional compressive stress-producing mechanism. Lattice distortion produced by the entrapment of impurity atoms of a size different from the hosting crystal is also a causative source of compression.

For epitaxial systems, lattice mismatch between the deposited material and the substrate will induce either compressive or tensile stress, depending on the material combination and respective lattice spacing. Upon cooling or heating (during deposition or subsequent thermal cycling), tensile or compressive thermal stresses will develop. For the simple case of an elemental layer, the thermal stress may be estimated from the difference in CTE between the film and the substrate.57,298 It is important to point out that CTE for
nanocrystalline films may significantly differ from bulk values.299

In the following, we propose new insights to control and optimize the intrinsic stress in polycrystalline films, based either on developing coating synthesis strategies to minimize growth stress (by the appropriate choice of the process parameters) or by employing interfacial and/or alloy design to purposefully tailor the film microstructure (and related stress state). Obviously, appropriate combination of substrate/coating materials (or the use of compensated layers) is, as far as possible, advisable for minimizing thermal stress.

2. Stress engineering strategies
   a. Optimizing process parameters. Precise control of the intrinsic stress during thin film deposition requires a subtle and simultaneous manipulation of multiple processing parameters, rendering the task not so straightforward. As discussed above, and highlighted in the kinetic model, substrate temperature \( T \) and growth rate \( R \) can be manipulated to modify the intrinsic stress. For high-mobility metals, increasing \( T \) (or decreasing \( R \)) contributes to increasing the compressive stress component, while, for low-mobility metals, it will result in a tensile stress increase. This trend is generally valid for thermal evaporation and electrodeposition conditions. As a rule of thumb, deposition at \( T/T_m \) conditions higher than 0.25–0.3, where \( T_m \) is the melting point of the thin film material, will imply recovery processes that can affect both the microstructure and the stress state (relaxation), at the expense of enhancing thermal stress contribution.

There exist additional processes that operate under bombardment with energetic particles (ions or neutrals), like in MS discharges. In such cases, it is important to remind that their stress state can be altered after venting (see Fig. 10(b)), usually, postgrowth oxygen incorporation generates a compressive stress contribution.146

Depositions at oblique angles tend to promote tensile stresses,25,300 as the fraction of intercolumnar voids increases with substrate tilt angles. The increase in surface roughness and the self-shadowing effect also contribute to decreasing the rate of adatom incorporation at the GB, which lessens the magnitude of the compressive stress component of the intrinsic stress.300

b. Interfacial and alloying design. As illustrated in Fig. 11, the interface chemistry plays a decisive role in the nucleation stage of polycrystalline thin films and consequently in the microstructure (grain size, texture) morphology and related stress state. Controlling the grain size, through the use of template layers,301 surfactant elements,175 or alloying,172,178,182 can be advantageously employed as a robust strategy to tailor the intrinsic stress. Films with coarse microstructures are recommended to minimize the contribution of defects trapping at the GB.

The combination of materials with different intrinsic stresses in a multilayer geometry is also a possible way to achieve low-stress levels.41,119 Specifically, one can tune the overall stress state by the appropriate choice of the modulation period as well as individual layer thicknesses.302 Films that remain in an amorphous state, e.g., below a critical thickness302 or above a certain concentration threshold in the case of alloys,188 are an interesting class of materials. Under conditions of sufficient mobility, they do not manifest any significant morphological change during growth so that a steady-state stress is in most cases observed.34 Therefore, one can take advantage of the disordered isotropic atomic structure to design films with a smooth surface and tailored stress levels by frustrating any grain-related stress changes. Amorphous alloys are used in many applications, but their intrinsic stress evolution during growth remains yet little explored.34,303

c. Tuning ion energy and ion flux. IBAD can be a clever route to control the stress state in thin films. By applying a bias voltage to the substrate, or by controlling the energy (also ideally the flux) of an independent ion source, the energy of ionized species can be adjustably controlled. Increasing the bias voltage will contribute to increasing the compressive stress component due to atomic peening through knock-on implantation processes in the growing layer. To initiate this process, the arriving particles must have energies higher than the energy threshold for atomic displacement of the thin film material, \( E_D \). However, it is interesting to operate at relatively high bias voltages (typically \( > 200–300 \) V) to promote stress relaxation via “thermal spike” processes.34,304 Therefore, with increasing bombardment energy, the compressive stress initially increases, goes through a maximum, and then decreases. Only a few percent (1%–2%) of very energetic impacts can favorably act as a stress relief source. This effect has been capitalized by growing films using plasma immersion ion implantation (PIII), a technique which has been shown to produce low-stress coatings.305

Recently, HiPIMS has proved to be a very promising route to tailor thin film microstructures and residual
By synchronizing the bias voltage to the temporal profile of the HiPIMS pulse delivered to the cathode material, it is possible to considerably reduce the compressive stress levels in TiAlN or VAlN films. Data reported in Fig. 28(b) clearly show that by the appropriate choice of the metal-ion flux during hybrid HiPIMS/DCMS film growth, the stress can be tailored accordingly. In the case of TiAlN films, it is conclusive to operate the Al target under HiPIMS conditions, favoring an Al\(^{+}\)-rich incident flux compared to the DCMS case, while the opposite target configuration (Al-DCMS/Ti-HiPIMS) involves a significant fraction of Ti\(^{2+}\) ions in addition to Ti\(^{+}\) ions [see Fig. 28(a)]. As a consequence, these doubly charged species gain an acceleration energy which is twofold that of Al\(^{+}\) ions. Greczynski et al. have shown that these Ti\(^{2+}\) ions are at the origin of larger compressive stress as well as precipitation of the wurtzite AlN phase at a lower Al content due to larger values of momentum transfer.\(^{43,306}\) Furthermore, these authors have shown that the compressive stress can be further decreased if one operates by synchronizing the bias voltage to the metal-rich portion [see the shaded region in Fig. 28(c)] of the HiPIMS pulse.\(^{307}\) In such a way, it is possible to select Al\(^{+}\) ions as the predominant energetic species reaching the substrate, while synchronizing the bias voltage to the entire HiPIMS pulse implies an Ar\(^{+}\)-rich ion flux. Irradiation with energetic Ar\(^{+}\) ions generates compressive stress due to creation of residual point defects, especially Ar entrapment at interstitial sites, whilst Al atoms can be primarily incorporated into the metallic lattice sites of the TiAlN film. This metal-ion bombardment also favors an extended solubility of Al in the cubic lattice of ternary nitrides.\(^{44}\) Thereby, tuning the time domain of ionized fluxes during HiPIMS deposition appears as a perspective approach to produce metastable layers with controllable stress levels. Finally, Cemin et al.\(^{308}\) reported that, in contrast to common expectations, the compressive stress magnitude in Cu HiPIMS films could be significantly reduced despite the energy increase of the bombarding particles from \(~20\) to \(~100\) eV. This behavior was attributed to ion-assisted grain growth during HiPIMS deposition which involved a large fraction of Cu\(^{+}\) ions.

**F. Living with cracks**

In our daily experience, we often fall victims to unwanted cracks or pernicious fractures. This is especially true for thin films where residual stresses result in film cracks. Film cracking, which is often fatal for film functionality, is
usually seen as a nuisance to avoid. However, it is interesting to note that film cracks frequently assume intriguing morphologies: a small sampler of arresting film crack patterns is provided in Fig. 29 with spiral cracks (a1), crescent shape zig-zag cracks (a2), festooned circular blisters (b1), and an intriguing parrot ladder blister network interspersed with floppy nondescript structures (b2). All these cracks and blisters formed spontaneously in homogeneous films.

Understanding the rationale behind such morphologies is interesting in itself, but there is more: we can build structures with cracks. Controlled crack propagation is widely used as a technological process: as examples we can mention glass cutting and wafer dicing. But these high-tech contemporary processes are by far predated (and in a sense also outperformed) by the craftsmanship of paleolithic artisans who could produce such amazing artefacts as 20 cm long, millimeters thick laurel leaf blades out of flintstone by the sole use of conchoidal fracture.

It turns out that for thin films, controlled cracking is a relevant elaboration process as well. On rigid substrates, controlled cracking of films has been proposed for mask fabrication to deposit submillimetric stochastic electrical grids. The stochastic nature of the mask results from the random process of cracking, and the absence of periodicity produces interesting optical properties. In this case, no film delamination should occur, but only sharp through cracks: wet deposition of the colloidal film material and carefully controlled drying make it possible to obtain the clean sharp crack edges and flat crack bottoms eminently suitable for subsequent metal deposition.

Another area where controlled film cracking is desirable is flexible electronics. Large deformations of a metal coating deposited on a compliant substrate can be obtained if triangular crack structures form, allowing out of plane torsion of the resulting ligaments with no further film rupture during stretching. This mode is essentially an excursion from planar into more 3D deformation. The strategy can be developed further, and in many cases, partial release of the thin film from the substrate (i.e., controlled interfacial cracking or delamination) is required. The film then assumes the form of a network of compliant straps which similarly allows the formation of large substrate strains without unwanted film rupture.

In the last example, the architecture of the coating is defined by lithographic processes, but we could consider carefully controlled cracking, as in the previous example. In fact, film cracking may emerge as a manufacturing tool for the fabrication of coatings with advanced architectures.

Can we thus think about a design with cracks? To meet this challenge, the control of crack propagation is vital. Unfortunately, it appears to be the second most difficult problem in the field of mechanics of materials, as we also know from our daily experience where getting a crack to go along a precisely defined path seems to be a difficult, sometimes impossible, task. Of course, there are physical laws behind film cracking, as the regularity of the patterns in Fig. 29 suggests. Even though our understanding of these laws is incomplete, in this section, we will illustrate some of the fundamental ideas around film cracking.

1. **Concerted film cracking**

In a film, some elastic energy is stored, amounting to \( e = \frac{h^2}{2E} \), where \( E \) is the film stress, \( E \) the biaxial modulus, and \( h \) the film thickness. Consider tensile stresses and one single through crack: since the film is constrained by the substrate, elastic energy is released in a region of size \( h \) on each side, and the crack can propagate if \( 2e > G_c \) (condition 1), where \( G_c \) is the film fracture energy.

In the inorganic sol-gel films of Fig. 28(a), the film thickness is such that condition (1) is not obeyed: one single crack cannot propagate. However, if two cracks propagate one along the other at separation \( W \), then the energy released is much larger as it now involves the full delamination width \( W \). It is given by \( We = WT - 2h G_c \) so that \( e = \Gamma + \frac{2h}{W} G_c \) (condition 2), where \( \Gamma \) is the interfacial rupture energy. A more in-depth analysis shows that \( h/W = 0.04 \). From this low value, we find that if film adhesion \( \Gamma \) is moderate, condition (2) is indeed much less stringent than condition 1. Moreover, one can show that the elastic interaction between the two cracks is stabilizing. If they stray apart, the interaction brings them back, and if they get closer, they are driven apart. As a result, a stable concerted propagation is possible.

Due to the symmetry of the equi-biaxial stress state, the spiral morphology shown in Fig. 28(a1) results from the propagation of the crack in interaction with itself, at a fixed separation \( W \). In the zig-zag/crescent morphology (a2), the crack interacts with itself as well, but when it comes to the end of a preexisting leg, it strays away to fold back onto itself. This behavior points to more subtle effects such as

![Fig. 29. (Color online) Various film crack patterns: spiral cracks (a1) and crescent shape zig-zag cracks (a2) (courtesy Marthelot, Ref. 309), festooned circular blisters (b1) (courtesy Benedetto), and blister network (b2) (courtesy Faou, Ref. 310).](image)
instabilities and bifurcations—the reason why it chose this direction can be better understood from some of the effects presented in Sec. IV F 2, namely, the telephone cord buckle and related morphologies.

2. Film instabilities, festoons, and branches

In the buckling delamination of compressively stressed films, as for the zig-zag concerted crack(s), the usual telephone cord and related morphologies [such as the parrot ladder—Fig. 29(b2)] all break the original axial symmetry.

For buckles (also called blisters), the key ingredients are twofold. First, we have to take into account the geometrical nonlinearities of plate mechanics. As a simple example of the first element of complexity, let us consider a square sheet of cardboard which is pulled apart by diagonally opposite corners (Fig. 30). In contrast to expectations, the deformed state is not the symmetric shape shown in Fig. 30(1a) but one of the two possible configurations where the full curvature is located along one single diagonal [Fig. 30(1b)]. This arbitrary choice of one state, out of two, signals bifurcation. But plate nonlinearities by themselves do not lead to the observed buckle morphologies. Therefore, a second element of complexity needed to understand thin film buckles lies in the physics of the interface. When the film buckles, the interface is loaded by both the opening moment \( M \) due to the buckle [Fig. 29(b1) inset] which prises the interface open and the traction \( T \) from the partially relaxed film within the buckle, which shears the interface. The ratio between these two types of loading (quantified by the mode mixity angle \( \eta \)) depends upon the plate conformation, and it strongly affects the actual interfacial energy of rupture [Fig. 30(2)]. Large shear significantly increases the rupture energy, effectively building up a pinning point. As a result, it is the coupling between plate nonlinearities and loading dependent rupture energy which gives rise to the diverse morphologies of thin film buckles.

With such a high degree of complexity, only the simplest geometries can be tackled analytically. The axi-symmetric case, i.e., the periodic destabilization of a circular blister, has been calculated by Hutchinson et al.\(^{317}\) This morphology has actually been observed [Fig. 29(b1)] although it is very infrequent in practice. For more complex (and more usual) morphologies, such as telephone cords and others [Fig. 29(b2)], one has to resort to numerical calculations. In this direction, interesting results connecting the period of the telephone cord buckles and interfacial rupture energy or predicting branching conditions have been recently published.\(^{310,318}\)

In summary, our improving ability to understand the mechanisms behind film cracking may empower crack-based fabrication processes for architectured films with advanced properties. Of course, there is still a long way to go. As with all architectured systems, one of the limits is the often imperfect match between the desired structures (when we can predict them, of course) and what is feasible. In this respect, as a final word of caution, we should stress that what has been discussed here is only crack propagation, i.e., the extension process of a preexisting crack. For architectured thin film manufacture, one would also need to control where a crack will originate, thus calling upon the most difficult problem in the field of materials science: crack initiation.

V. CONCLUSIONS AND OUTLOOKS

The last decade has seen remarkable progress in the understanding of stress evolution during growth and processing of thin film materials and coatings. This has been driven by the development and emergence of new evaluation methods and gaining insight into the underlying stress mechanisms. Characterization methods, such as wafer curvature and XRD, can be used operando and in real-time, making it possible to achieve a fine tuning and control of the stress level for the design of novel nanostructured materials and functional coatings with enhanced performance and extended durability. Using state-of-the-art x-ray synchrotron and FIB facilities, location and depth-sensing of the microstructural attributes in terms of grain size, grain shape, phases, texture, and residual stress profiles have become possible, with spatial resolution down to 50 nm, opening important avenues in the assessment of stress in coatings with complex microstructures and architectures, such as graded composite layers, multilayered systems, and multicomponent and multiphase thin film materials.

Recent experimental findings, supported by analytical models and atomic-scale simulations, have shown the importance of grain boundaries and deposition flux in governing the intrinsic stress build-up and relaxation processes during thin film growth. While the energetic particle bombardment
often results in the development of compressive residual stress, controlling the ionization degree of film-forming species and the temporal profile of the particle flux may be advantageously employed to mitigate and tailor the intrinsic stress in specific films and coating systems.

Besides the energetic control of the film growth through temperature and ion bombardment as well as interfacial engineering (to ensure good adhesion) and postdeposition stress relaxation (e.g., annealing), complementary approaches to mitigate, control, and manage stress are related to specific film and coating architectures; this includes the use of stressors in microelectronics and micro-/nanosystems, the use of new and novel coating materials (metallic glass films and hybrid organic/inorganic films), and the novel multifunctional approach consisting of implementing and optimizing both optical and thermo-mechanical (stress, curvature component, and CTE) properties in optical coatings and other areas.

One of the key elements in further progress to manage stress relies on our understanding of the relationship between stress and mechanical instabilities, such as buckling or film cracking as a consequence of the stored energy dissipation. This issue clearly calls for in-depth investigations of propagation of preexisting cracks and, particularly, crack initiation, both related to the assessment of toughness. Such a consideration opens up new avenues for both experimental and modeling and simulation studies, including assessment and control of the stress depth profiles and crack-based fabrication processes for architectured films and coatings with advanced properties. These challenges could be used as strain engineering routes to develop nanoscale systems with innovative device-level functionalities through a careful control of the mechanical deformations and applied stresses, such as flexible, stretchable, and shape adaptive devices based on atomically thin materials (graphene, transition metal dichalcogenide monolayers, and Si nanomembranes) or coatings on polymeric substrates.

ACKNOWLEDGMENTS

The authors wish to acknowledge Jonathan Colin and Clarisse Furgeaud, University of Poitiers, and Rostislav Daniel, Montanuniversität Leoben, for their respective contribution to some of the presented work. M.S. would like to thank Edoardo Bemporad (Roma TRE University) and Alexander M. Korsunsky (University of Oxford) for the extremely useful scientific discussions on residual stress depth profiling by FIB-DIC. The work of E.C. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award No. DE-SC0008799 and the NSF-DMR under Award No. 1602491. M.S.’s work has received research funding from the European Union, within the large collaborative project OYSTER, Grant Agreement No. 760827. Most of the FIB activities were carried out at the “Inter-Departmental Laboratory of Electron Microscopy” (LIME), University ROMA TRE (http://www.lime.uniroma3.it). G.B.T. recognizes the support of the Army Research Office, Grant No. W911NF1310436, Michael Bakas program manager. Examples of the results from Polytechnique Montréal have been obtained from the work supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada through numerous university-industry research partnership grants, most recently the NSERC Multisectorial Industrial Research Chair in Coatings and Surface Engineering. Finally, the authors particularly acknowledge the support of the two cooperating organizations, namely, the Society of Vacuum Coaters (SVC) and the Advanced Surface Engineering Division (ASED) of the American Vacuum Society (AVS).

NOMENCLATURE

\begin{align*}
\text{ALD} &= \text{atomic layer deposition} \\
\text{APT} &= \text{atom probe tomography} \\
\text{AR} &= \text{antireflective} \\
\text{ASED} &= \text{Advanced Surface Engineering Division} \\
\text{AVS} &= \text{American Vacuum Society} \\
\text{BEM} &= \text{boundary element method} \\
\text{BMG} &= \text{bulk metallic glasses} \\
\text{BOPET} &= \text{biaxially oriented polyethylene terephthalate} \\
\text{CAD} &= \text{cathodic arc deposition} \\
\text{CCD} &= \text{charge-coupled device} \\
\text{CMOS} &= \text{complementary metal oxide semiconductor} \\
\text{CTC} &= \text{compressive-tensile-compressive} \\
\text{CTE} &= \text{coefficient of thermal expansion} \\
\text{CVD} &= \text{chemical vapor deposition} \\
\text{DCMS} &= \text{direct current magnetron sputtering} \\
\text{DIBS} &= \text{dual ion beam sputtering} \\
\text{DIC} &= \text{digital image correlation} \\
\text{DLC} &= \text{diamondlike-carbon} \\
\text{EBE} &= \text{electron beam evaporation} \\
\text{EBSD} &= \text{electron backscattered diffraction} \\
\text{EC} &= \text{explosive crystallization} \\
\text{FCAD} &= \text{filtered cathodic arc deposition} \\
\text{FE} &= \text{finite element} \\
\text{FET} &= \text{field effect transistor} \\
\text{FIB} &= \text{focused ion beam} \\
\text{FWHM} &= \text{full width at half maximum} \\
\text{GB} &= \text{grain boundary} \\
\text{GIXRD} &= \text{glancing-incidence x-ray diffraction} \\
\text{HiPIMS} &= \text{high-power impulse magnetron sputtering} \\
\text{IBA-CVD} &= \text{ion beam assisted chemical vapor deposition} \\
\text{IBAD} &= \text{ion beam assisted deposition} \\
\text{IBS} &= \text{ion beam sputtering} \\
\text{MD} &= \text{molecular dynamics} \\
\text{MEMS} &= \text{microelectromechanical systems} \\
\text{MOSS} &= \text{multiple-beam optical stress sensor} \\
\text{MS} &= \text{magnetron sputtering} \\
\text{NEMS} &= \text{nanoelectromechanical systems} \\
\text{OC} &= \text{optical coating} \\
\text{OIF} &= \text{optical interference filters} \\
\text{PC} &= \text{polycarbonate} \\
\text{PECVD} &= \text{plasma enhanced chemical vapor deposition} \\
\text{PET} &= \text{polyethylene terephthalate} \\
\text{PMMA} &= \text{polymethyl methacrylate} \\
\text{PVD} &= \text{physical vapor deposition} \\
\text{SDRS} &= \text{surface differential reflectance spectroscopy}
\end{align*}
SEM = scanning electron microscopy
SOI = silicon on insulator
SVC = Society of Vacuum Coaters
SZM = structure zone model
TEM = transmission electron microscopy
TSV = trench silicon vias
UHV = ultrahigh vacuum
VAM = velocity accommodation mode
XEC = X-ray elastic constant
XRD = x-ray diffraction
XRR = x-ray reflection
