The metric description of elasticity in residually stressed soft materials

Efi Efrati,*a Eran Sharonb and Raz Kupfermanc

Living tissue, polymeric sheets and environmentally responsive gel are often described as elastic media. However, when plants grow, plastic sheets deform irreversibly and hydrogels swell differentially the different material elements within an object change their rest lengths often resulting in objects that possess no stress-free configuration making the standard elastic description inappropriate. In this paper we review an elastic framework based on Riemannian geometry devised to describe such objects lacking a stress-free configuration. In this framework the growth or irreversible deformation are associated with the change of a reference Riemannian metric that prescribes local distances within the body, and the elastic problem is one of optimal embedding. We discuss and resolve points of controversy regarding the Riemannian metric formulation. We give examples for dimensionally reduced theories, such as plates and shells theories, which arise naturally and discuss the relation between geometric frustration and residual stress.

1 Introduction

In the past few years there has been a revived interest in the continuum description of growth and plasticity in amorphous materials, and the generation of residual stress.1–4 This interest may be attributed, in part, to the availability of new responsive materials in which spatially inhomogeneous swelling and shrinking can be controlled by simple external triggers, for example by a uniform heating.5 In addition, there has been an increasing interest in the mechanics of biological systems and the precise control of biomimetic devices.6–7 These interests in the quantitative mechanical description of soft and responsive matter have been accompanied by an effort to formulate the existing elastic theories within a purely geometric framework, which facilitates a more rigorous approach to dimensionally reduced theories.8

The Riemannian-metric description of residually stressed soft materials is aimed to describe a wide class of amorphous materials, in which local irreversible deformations determine local rest lengths rather than prescribe a configuration. These non-elastic deformations (which in fact can be reversible, however depend on external conditions, such as temperature of humidity, and are therefore irreversible from an elastic point of view) include growth in plants (both by cell proliferation9 and cell expansion), hygroscopic motion,7,11 environmentally responsive gels,5,12 and plastic deformations in amorphous solids.13 There are fundamental differences between these different types of irreversible deformations. Thermodynamic inequalities (e.g. entropy production) that govern plastic deformations do not hold for general irreversible processes. In growth processes, even mass conservation may not hold (due to nutrient flow, which is not considered within the elastic framework). Yet, despite these differences, the elastic theories that treat the bodies’ response after the irreversible deformation took place follow similar lines. Some of the ideas presented below in the context of growth and other irreversible deformations can be traced back to ideas and notions presented over fifty years ago in the context of plasticity. These include the concepts of inclusions and eigenstrains,14 continuous distribution of dislocations,15–18 and material inhomogeneities.19 For brevity, from this point onward we will refer to the non-elastic portion of the deformations as irreversible deformations, a term which in our context should be interpreted as any of the processes described above: growth, plasticity, and external stimuli response (temperature, humidity, etc.).

In standard elastic theories the main descriptor of the system is the strain tensor, which quantifies the deviation of the configuration that the body assumes from its stress-free rest configuration. The constitutive relations assign to every strain a stress tensor, from which one obtains the equations of elastic equilibrium; in hyper-elasticity these constitutive relations can be derived from an elastic energy density functional. As we demonstrate in Section (5.1) below, local irreversible deformations, which prescribe changes in the local rest lengths, often lead to bodies that possess no stress-free configuration, making such a rest configuration-based formulation inappropriate.
There are numerous different elastic theories accounting for irreversible deformations and aimed at describing bodies lacking a stress-free configuration. These theories adopt different approaches yet share some common traits. In particular all theories distinguish between deformations that are due to irreversible processes and deformations that are due to the elastic response. The decomposition of the total deformation into a reversible and an irreversible component is, however, done differently in different theories.

One of the most successful and widely adopted elastic frameworks for describing irreversibly deformed soft media is due to Rodriguez, Hoger and McCulloch;20 it adopts the Bilby–Kröner–Lee multiplicative decomposition of deformation gradients.21–23 The deformation of an initially stress-free object into the (possibly) residually stressed final configuration is considered as a composition of two deformations. The first deformation is irreversible and maps the stress-free state into a “virtual configuration”. The second deformation is due to the elastic response and maps the “virtual configuration” of the body to its deformed final configuration. The gradient of the full mapping becomes a product of two “gradient like” terms. We define this approach mathematically and review further details of this decomposition in Appendix B. The deformed configuration of the body, the “virtual configuration” of the body and the mapping between these two configurations form the basic components of the elastic description in this approach. The constitutive laws in this framework, even in the small strain limit, are non-linear in the components of the elastic deformation gradient-like term.

In contrast, in the metric description no explicit use of the configuration of the body is made. The theory is formulated in terms of metric tensors (prescribing local distances) and employs intrinsic Riemannian geometric quantities. While capable of describing large strains, it is particularly attractive when applied to cases of small strain and arbitrary rotations, in which the constitutive laws become linear in the unknown metric. The non-linear nature of the problem remains in the form of the non-linear compatibility conditions imposed on the metric of the deformed body. The decomposition of an observed deformation into an irreversible contribution and an elastic contribution resembles an additive decomposition of strains. While some aspects of such additive decomposition have been shown to lead to inconsistencies,24 we show (in Appendix B) that this is not the case in the metric description. Moreover, as the metric description is fully covariant, it does not make use of the initial stress-free configuration of the body, and only refers to the local attempted rest-lengths which incorporate the results of the growth process. While in almost all cases the results obtained from the metric approach agree with results obtained via the multiplicative decomposition, the metric approach has two main advantages: (i) it is purely geometric, allowing for the immediate adaptation of many embedding results from Riemannian geometry, and (ii) it forms a non-linear extension of the standard small displacement theory, thus makes most of the quantities involved easy to interpret.

We next review qualitatively the central notions that necessitate the departure from the standard elastic description: incompatibility and residual stress.

2 Incompatibility and residual stress

In the metric formulation, an irreversible deformation results in new local “rest distances”. The prescription of local distances in a body, endows it with a geometric structure captured by a reference metric, $g$. This geometric structure may however be incompatible with the known laws of Euclidean space (for example, the sum of internal angles in a triangle may not equal 180 degrees). As the body assumes its current configuration in Euclidean space, its adapted geometric state needs to obey the laws of Euclidean space. In particular, the metric, $g$, associated with its adapted configuration reflects the geometric structure and laws of Euclidean space. Hence, the rest distances dictated by the reference metric $\bar{g}$ cannot be everywhere satisfied (by the metric $g$), giving rise to geometric incompatibility with Euclidean space, or for short incompatibility.

The notion of incompatibility is central in the geometric formulation of irreversible deformations. The mechanical manifestation of incompatibility is the presence of residual stress, which is the stress present within a body in the absence of external forces or constraints.

Whenever the metric $\bar{g}$ induced by the growth is incompatible, there exists no Euclidean† metric $g$ that coincides with $\bar{g}$ almost everywhere, (see for example ref. 25 p. 26). Thus, if a growth profile induces an incompatible metric, the prescribed rest lengths cannot be obeyed everywhere simultaneously, giving rise to residual strain. This residual strain in turn gives rise to residual stress through the specific constitutive relations of the material. Even though incompatibility and residual stress are intimately connected, they are not synonymous as they are not related by any local law. Incompatibility is a local property which is estimated through the metric (encoding distances) and its first and second derivatives, whereas the residual stress also depends on global properties, such as the shape, topology and size of the body.

In the appendix we give an explicit example of the above distinctions in a two dimensional setup. We demonstrate that uniform (constant) incompatibility gives rise to residual stress whose magnitude increases as we consider larger domains.

3 The signature of residual stress

The notion of geometric incompatibility was introduced in the context of elasticity in the description of plastic deformations and defects in solids. Kröner26 provides the following explanation:

“If the [residually] stressed body is cut into small elements, in which the [residual] stress is then relaxed, we obtain an assembly of elements which do not fit together. Thus, an incompatible deformation implies that a non-fitting collection of elastic elements is united to a compact body. If the non-fitting occurs on

† Continuous and non-singular.
the infinitesimal scale, then we can obtain an internal stress state which varies continuously through the body."

This statement can also be interpreted the other way around: the residual stress present in the equilibrium configuration of a body can be revealed by dissecting the body into small elements. Once dissected, the pieces of the body relax the residual stress through deformations that were unaccessible when the body was a whole. In general, the relaxed fragments no longer fit together; if they do, then the residual stress can be neglected in the elastic description of the body. See Fig. 1 for an example.

4 The metric tensor and the Riemannian curvature

In order to quantify incompatibility we resort to tools from differential geometry, namely the metric tensor and the Riemannian curvature tensor. We endow a body with a set of material curvilinear coordinates \( x \in \Omega \subset \mathbb{R}^3 \), i.e. we associate every material point in the body with a triplet \((x^1, x^2, x^3)\) which varies continuously throughout the body and deforms with the body (keeping the identity of the material point with which it is identified). We associate the configuration the body assumes in space with a mapping \( \mathbf{r}: \Omega \to \mathcal{D} \). This mapping induces on the body a metric \( g \) whose components are

\[
g_{ij} = \partial_i \mathbf{r} \cdot \partial_j \mathbf{r},
\]

where \( \partial_i = \partial/\partial x^i \) and Latin indices take the values \( \{1, 2, 3\} \). The generalization of the Pythagorean equality gives local distances by

\[
dx^2 = g_{ij} dx^i dx^j,
\]

where we have used the Einstein summation convention whereby any product of a repeating upper and lower index is summed over its range. The components of the mapping gradient are given by

\[
F_{ij} = \partial_i \mathbf{r}.
\]

The metric can be represented using the mapping gradient as \( g = F^T F \). Note that not every 3-by-3 matrix, \( F \), can represent a mapping gradient. A necessary and sufficient condition is that its rows have a vanishing curl. For the metric tensor, the condition that \( g \) is of the form (1) is non-linear and slightly more complicated: the necessary and sufficient condition for a symmetric positive definite 3-by-3 matrix \( g \) to be the metric tensor of a body in Euclidean space is that all (six) independent components of the Riemannian curvature tensor associated with it vanish.\(^{25}\) The components of the Riemannian curvature tensor are given by

\[
R^k_{ij} = \partial_i F^k_j + \partial_j F^k_i - \partial_k F^l_i \delta^j_l - \partial_k F^l_j \delta^i_l,
\]

where the components of the Christoffel symbols, \( \Gamma \), are given by

\[
\Gamma^i_{jk} = \frac{1}{2} g^{il} (\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk}).
\]

The vanishing of the independent components of the Riemannian curvature tensor are also called the compatibility conditions, and a metric that complies with these compatibility conditions is called a Euclidean (or compatible) metric. For simply connected domains, every Euclidean metric determines a configuration, \( \mathbf{r} \), which is unique up to rigid motions.

5 Irreversible deformation profiles and metric prescription

Formulating irreversible deformations in terms of a metric tensor provides a natural description of the geometry of processes such as growth, plasticity and temperature response. Using this formulation we can easily determine which deformation profiles will result in a residually stressed body, and conversely, characterize the intrinsic geometric structure of the body independently of the underlaying deformation profile leading to it. In this context we conclude that (i) a generic irreversible deformation will result in a residually stressed body, and (ii) different irreversible deformation profiles may lead to the exact same intrinsic geometry. Below we exemplify these two principles by specific examples.
reference metric, *i.e.* are realizable by an Euclidean metric, $g$, and therefore do not induce residual stress?

To answer this question we write down the components of the Riemannian curvature tensor (4) of the metric $\bar{g}$ in terms of the expansion factor $\lambda$ and its derivatives. Taking independent linear combination of the covariant components of the Riemannian curvature tensor yields the following compatibility conditions:

\[
2(\hat{\partial}_i \hat{\partial}_j \lambda) - \hat{\partial}_i \hat{\partial}_j \lambda - \Delta \lambda = 0, \quad 2\hat{\partial}_i \hat{\partial}_j \lambda - \lambda \hat{\partial}_i \hat{\partial}_j \lambda = 0, \\
2(\hat{\partial}_i \hat{\partial}_j \lambda) - \hat{\partial}_i \hat{\partial}_j \lambda - \Delta \lambda = 0, \quad 2\hat{\partial}_i \hat{\partial}_j \lambda - \lambda \hat{\partial}_i \hat{\partial}_j \lambda = 0, \\
2(\hat{\partial}_i \hat{\partial}_j \lambda) - \hat{\partial}_i \hat{\partial}_j \lambda - \Delta \lambda = 0, \quad 2\hat{\partial}_i \hat{\partial}_j \lambda - \lambda \hat{\partial}_i \hat{\partial}_j \lambda = 0,
\]

where $\Delta = \nabla \cdot \nabla = \partial^2 + \partial^2 + \partial^2$ is the standard Laplacian operator. It takes straightforward algebra and integration to find that the only non-constant solution of the above equations is

\[
\lambda = \frac{C^2}{|X - X_0|},
\]

for some constants $C$ and $X_0$. Every other isotropic expansion profile of an initially Euclidean 3D body will give rise to a non-Euclidean metric and inevitably result in a residually stressed body. This result, which also appears in ref. 27 and 28, may be surprising when considering growth profiles. However it is a consequence of a well-known geometric result whereby all conformal mappings in $\mathbb{R}^3$ are inversions of a sphere. It implies that any growth that does not result in residual stress requires delicate global control, or some mechanical feedback.

### 5.2 Different growth profiles leading to the same intrinsic geometry

We now give an example of three different growth profiles of an initially Euclidean geometry that lead to the same final reference metric. We endow the body, prior to its growth, with cylindrical coordinates, $(r, \theta, z)$. As in eqn (2) we express the corresponding metric as the square of an infinitesimal length element:

\[
ds_0^2 = dr^2 + r^2 d\theta^2 + dz^2,
\]

defined in the cylindrical domain $0 \leq r \leq 1/\sqrt{2}$.

- **Planar isotropic expansion:**

  Consider an isotropic expansion in the $(r, \theta)$ plane by a factor $\lambda(r) = 2c/(1 + cr^2)$, where $c = 2 - \sqrt{2}$. As a result, the infinitesimal length element becomes $ds_0^2 = \lambda^2 dr^2 + \lambda^2 r^2 d\theta^2 + dz^2$. Setting $\rho(r) = 2\arctan(cr)$ the new length element takes the form,

\[
ds_\rho^2 = dr^2 + \sin^2(\rho) d\theta^2 + dz^2, \quad 0 \leq \rho \leq \pi/4.
\]

- **Radial expansion:**

  Consider now an expansion only in the radial direction by a factor of $\eta(r) = (1 - r^2)^{-1/2}$. As a result, the infinitesimal length element becomes $ds_\eta^2 = \eta^2 dr^2 + r^2 d\theta^2 + dz^2$. Setting $\rho(r) = \arcsin(r)$, the new length element takes the form,

\[
ds_\rho^2 = dr^2 + \sin^2(\rho) d\theta^2 + dz^2, \quad 0 \leq \rho \leq \pi/4.
\]

- **Azimuthal shrinkage:**

  Lastly, consider a uniform planar isotropic expansion by a factor $\alpha = \pi/(2\sqrt{2})$ followed by an azimuthal shrinkage by a factor of $\psi(r) = \sin(\alpha r)/(\alpha r)$. The resulting length element is $ds_{\text{III}}^2 = \alpha^2 d\theta^2 + \alpha^2 r^2 d\theta^2 + dz^2$. An explicit substitution of $\psi$ and setting $\rho = \alpha r$ gives

\[
ds_{\text{III}}^2 = d\rho^2 + \sin(\rho) d\theta^2 + dz^2, \quad 0 \leq \rho \leq \pi/4.
\]

Note that $\bar{g}_1 = \bar{g}_n = \bar{g}_{\text{III}}$, thus all three growth profiles lead to the exact same intrinsic geometry. While the equivalence of the growth profiles is apparent in the $(\rho, \theta)$ coordinates, this is not easily identified when the metrics are given with respect to the original coordinates $(r, \theta)$.

### 6 Hyper-elasticity in the metric formulation

When the reference metric of an unconstrained body is Euclidean, it determines a unique (up to rigid motions) stress-free configuration which, in the absence of external forces, the body will trivially adopt. However, whenever the reference metric is non-Euclidean, the equilibrium configuration of the body will depend on the form of the elastic energy. Specifically, two bodies possessing the exact same intrinsic geometry but having two different constitutive relations will result in different equilibrium configurations. We now turn to study the elastic energy within the framework of a geometric description of hyper-elasticity.

The basic principle of hyper-elasticity states that the elastic work done by a body is derived from a local potential, *i.e.,* (i) it depends on the configuration of the body but is independent of the path (in configuration space) leading to the specific configuration and (ii) it is a sum over local energy contributions. At this point one may derive the form of the elastic energy from a Lagrangian–Riemannian approach, or within an Eulerian setting starting from the embedded body.

While the purely geometric Lagrangian approach is more natural to the problem, the Eulerian approach is closer in spirit to the classical derivations of the theory for hyper-elastic solids, and is more physically intuitive. We next present the latter.

The assumption of hyper-elasticity states that we may write the elastic energy as a volume integral:

\[
E = \int_{\Omega} \hat{W} dV,
\]

where $\hat{W}$ is the elastic energy density. For an unconstrained body, the local energy contribution is independent of the absolute position of the material element in $\mathbb{R}^3$, and depends only on the deformation gradient, $\hat{W} = \hat{W}(\nabla r(x), x)$. Moreover, the energy density in an unconstrained body is also invariant under rigid motions, therefore, the metric tensor, $g$, which determines the configuration up to rigid motions, may be used as the unknown describing the configuration in the elastic energy,† We therefore write an energy density function with

† This assumption fails if the ambient space is not isotropic or not homogeneous or in cases where the embedding is not orientation preserving.
respect to the variable \( g \) (which in the elastic context is known as the right Cauchy–Green deformation tensor\(^{25,32} \)),

\[
E = \int_\Omega \mathcal{H}(g, x) \, dV.
\]

We further assume that at every point \( x \) the energy density, \( \mathcal{H} \), possesses a unique zero which we denote by \( g = \tilde{g}(x) \), i.e.

\[
\mathcal{H}(g, x) = 0 \iff g = \tilde{g}(x).
\]

The tensor \( \tilde{g} \) is the reference metric, which would be assumed by the body in a state of zero stress (in ref. 31 the term target metric was used). \( \tilde{g} \) is a local property of the body, which in the present case is determined by the growth process. We may now write an energy density per unit of reference volume \( \mathcal{H} = \mathcal{H}/|\tilde{g}| \) where \( |\cdot| \) denotes a determinant. If the elastic properties of an amorphous growing body are homogenous, the only spatially varying properties in the body are those captured by the reference metric. We may therefore write

\[
E = \int_\Omega \mathcal{H}(g, \tilde{g}) \sqrt{\tilde{g}} \, dV \, dV = \int_\Omega \mathcal{H}(g, \tilde{g}) \sqrt{\tilde{g}} \, dx \, dx \, dx,
\]

where both the reference metric, \( \tilde{g} \), and the body’s metric \( g \) are assumed continuous and non-singular. If either is allowed to be singular, then non-orientation preserving realizations of the metric become possible thus invalidating the above considerations. For a non-singular reference metric \( \tilde{g} \) obtaining an elastic equilibrium consists of finding a Riemannianly flat, continuous, and non-singular metric, \( g \) that minimizes \( \mathcal{H} \).

### 6.1 The stress distribution at equilibrium

Identifying the second Piola–Kirchhoff stress tensor as\(^{25,32} \)

\[
S^i\!_j = 2 \frac{\partial \mathcal{H}}{\partial g_{ij}},
\]

we obtain the Euler–Lagrange equations

\[
\frac{1}{\sqrt{|\tilde{g}|}} \frac{\partial}{\partial x} \left( \sqrt{|\tilde{g}|} S^i\!_j \right) + \Gamma^i\!_{jk} S^k\!_j = 0. \tag{8}
\]

An alternative form of these equations in which their contravariant nature is more apparent is

\[
\tilde{g}^i\!_j S^i\!_j + (\Gamma^i\!_{jk} - \tilde{\Gamma}^i\!_{jk}) S^k\!_j = 0,
\]

where \( \tilde{g} \) denotes the covariant derivative with respect to the metric \( \tilde{g} \), (for details see ref. 3). Whenever \( \tilde{g} \) is non-Euclidean, there must be a metric discrepancy between \( g \) and \( \tilde{g} \) which, as discussed above, will manifest as residual stress.

In the absence of external forces the field of residual stress must be self balancing. This property poses restrictions on the possible states of residual stress within a body.\(^{23} \) The restrictions for the field of residual stress may be obtained by considering a scalar field \( \chi \) with a non vanishing gradient \( \nabla \chi \), which satisfies

\[

\nabla \chi \nabla v_j \mathcal{H} = 0, \tag{9}
\]

where the covariant derivative above is taken with respect to the metric \( g \). For all such test functions \( \chi \) the quadratic form \( S^i\!_j \nabla \chi \) must average to zero when integrated over the entire body, i.e.,

\[
\int_\Omega \nabla \chi \nabla \chi \mathcal{H} \sqrt{|\tilde{g}|} \, dx \, dx \, dx = 0. \tag{10}
\]

The proof follows immediately from integration by parts and explicit substitution of the divergence equation for the stress \( \mathcal{H} \). In particular, as the integrand is a quadratic form in the gradient, \( \nabla \chi \), every non trivial residual stress field must contain both tension and compression, as discussed in ref. 34. This result can be further understood if we consider the Eulerian coordinates, \( (x, y, z) \), which form a Cartesian set in the deformed configuration. For such a choice of coordinates the metric is trivial, i.e. \( g = I \), and all test function gradients \( \nabla \chi \) which comply with \( \epsilon \) are constant, and can therefore be taken outside the integral. As the components of the constant gradients \( \nabla \chi \) can be chosen arbitrarily, the above result \( \epsilon \) implies that each of the Cartesian components of the second Piola–Kirchhoff stress tensor must average to zero.

### 6.2 Linear (Hookean) constitutive relations

The Green-St. Venant strain tensor can be expressed using the metric tensor\(^{25} \) by

\[
\varepsilon = \frac{1}{2} (g - \tilde{g}).
\]

The assumptions in the previous subsection imply that

\[
\mathcal{H} = 0 \iff \varepsilon = 0.
\]

Thus for small strain

\[
\mathcal{H} = A^{ijkl} \varepsilon_i \varepsilon_j \varepsilon_k \varepsilon_l + C(x^2).
\]

For an isotropic and homogenous solid the rank-four elasticity tensor is fully determined by the reference metric \( \tilde{g} \) and two additional constants, which we identify as the material’s Young’s modulus, \( Y \), and Poisson ratio, \( \nu \):

\[
A^{ijkl} = \frac{Y}{1 + \nu} \left[ \frac{1}{2} \left( \tilde{g}^k \!_j \tilde{g}^l \!_j \tilde{g}^m \!_k + \tilde{g}^m \!_k \tilde{g}^l \!_j \tilde{g}^k \!_j \right) + \frac{\nu}{1 - 2\nu} \tilde{g}^j \!_i \tilde{g}^j \!_k \right]. \tag{11}
\]

where \( \tilde{g}^i \!_j \) are the components of the reciprocal reference metric, i.e. \( g^i \!_j \tilde{g}^j \!_k = \delta_k \). If we define the raising and lowering of indices with respect to the reference metric, the elasticity tensor \( A^{ijkl} \) gives rise to a stress–strain relation of the form

\[
S^i \!_j = A^{ijkl} \varepsilon_k \!_l = \frac{Y}{1 + \nu} \left( \varepsilon_i \!_j \varepsilon_j \!_l + \frac{\nu}{1 - 2\nu} \tilde{g}^j \!_i \tilde{g}^j \!_l \right), \tag{12}
\]

and an energy density

\[
\mathcal{H} = \frac{1}{2} S^i \!_j \varepsilon_i \!_j = \frac{Y}{1 + \nu} \left( \frac{1}{2} \varepsilon_i \!_j \varepsilon_j \!_l + \frac{\nu}{1 - 2\nu} \tilde{g}^j \!_i \tilde{g}^j \!_l \right). \tag{13}
\]

**Comments:**

(1) In cases where \( g \) is Euclidean, eqn (12) coincides with that of standard non-linear Hookean elasticity.\(^{25} \)
the additive decomposition of elastic–plastic strains as
\[ g - I = (g - \bar{g}) + (g - I), \]
there are vast differences between the two. In fact, in some aspects, the metric description is closer in spirit to the multiplicative decomposition of deformation gradients as it does not make use of a plastic strain. In Appendix B we give a thorough account of the points of differences and similarities between these two elastic frameworks.

(3) The Hookean description (13) is expected to be valid for all small elastic strains, \( \varepsilon \), regardless of how much irreversible deformation (change in the reference metric, \( \bar{g} \)) has occurred. In fact, in the metric description there is no cumulative plastic strain, and the system has no memory of the irreversible deformations it underwent; only the final rest distances are retained. However, plastic history dependent effects such as stiffness variations due to plastic flow may be easily incorporated through equations of the form
\[ \hat{Y} = f(Y, \bar{g}, \vec{g}, S), \]
which relates the material’s Young’s modulus, \( Y \), to the plastic flow rate \( \vec{g} \), rather than the total plastic flow. In general, the time scales associated with changes in the reference metric and with variations in the Young’s modulus are much longer than any typical elastic time scale. We therefore assume that the metric \( g \) at every instant solves the elastic problem as given by the Young’s modulus and reference metric at that time.

7 Application to thin growing elastic sheets

In general, observing the effect of residual stress in a three-dimensional body without compromising its integrity is difficult. In thin bodies, however, bending deformations have a low energetic cost, hence large deformations may occur even for small residual stress fields. One of the main motivations that has driven the formulation and study of the metric formulation is its application to thin sheets.\(^3,25\)

Natural growth processes that result in slender structures often allow an elastic description as thin elastic sheets. Plant leaves,\(^3,5\) flower’s petals,\(^9\) algae blades,\(^36\) and seed pods valves are a few examples of such structures. While tissue growth may be sensitive to mechanical stress, the time scale associated with this response is much longer than the elastic relaxation time of these structures, thus the elastic problem posed by a given growth profile can be considered decoupled from the growth process. As growth is a local process that occurs simultaneously at distant points within the tissue, it is likely to generate geometric incompatibility, which will manifest as residual stress. The geometric incompatibility enables very simple growth profiles to generate very complex three dimensional shapes. In particular such systems may present patterns over scales much smaller than the spatial length scale associated with the growth process\(^3\) and may also exhibit stress focusing.\(^37\)

Determining the expected shape and mechanical state produced by a given incompatible growth profile of a thin body provides a better understanding of the growth process, and also allows for new design methods for the production of thin structures.

7.1 Geometry of surfaces and incompatibility

When considering a thin elastic body, it is natural to construct dimensionally reduced theories by identifying the body with its mid-surface and describing its elastic behavior in terms of properties of its mid-surface. Mathematically, a surface embedded in Euclidean space is a mapping \( r: \mathcal{S} \to \mathbb{R}^3 \), where \( \mathcal{S} \subset \mathbb{R}^2 \). Endowing the surface with material two dimensional coordinates \( x^2 \) (Greek indices assume the values \{1, 2\}), induces a first and second fundamental forms on the surface through
\[ a_{ab} = \partial_a r \cdot \partial_b r, \quad b_{ab} = \partial_a \partial_b r \cdot \hat{n}, \]
where \( \hat{n} \) is the normal to the surface. The first fundamental form is simply the metric tensor of the surface. The second fundamental form measures the normal curvature of the surface. The two fundamental forms of a surface are not independent and must comply with three algebraic differential equations called the Gauss–Peterson–Mainardi–Codazzi (GPMC) equations:\(^18\)
\[ \begin{align*}
|b| &= g_{11} R_{111}^2 = |\nu| K, \\
\partial_1 b_{22} + \Gamma^2_{22} b_{11} &= \partial_2 b_{12} + \Gamma^2_{12} b_{22}, \\
\partial_2 b_{11} + \Gamma^1_{11} b_{22} &= \partial_1 b_{21} + \Gamma^1_{21} b_{11},
\end{align*} \tag{14} \]
where \( K \) is the Gaussian curvature. The GPMC equations are necessary and sufficient conditions for a symmetric positive definite matrix \( a \), and a symmetric matrix \( b \), to be the first and second fundamental forms of an immersed surface. The GPMC equations are related to the vanishing of the Riemanian curvature tensor in three dimensions, and are also called compatibility conditions (see ref. 25 for further discussion on this subject). Compatible first and second fundamental forms define, up to rigid motions, a unique surface.

7.2 Non Euclidean plates and shells

The covariant nature of the metric formulation provides a description that is invariant under different choices for the coordinates used to parameterize the body. We may therefore choose a particular convenient set of coordinates for the parametrization of slender bodies. Following\(^3,25,39\) we choose a semi-geodesic set of coordinates with respect to the mid-surface. In such a parametrization the \( x^3 \) parametric curves (characterized by constant values for \( x^1 \) and \( x^2 \)) form geodesic curves which intersect the mid-surface \( (x^3 = 0) \) perpendicularly (see also ref. 40 p. 136). Moreover, \( x^3 \) may be chosen as an arc-length parameter for these curves resulting in the following form for the reference metric:
\[ \bar{g} = \begin{pmatrix}
\bar{g}_{11} & \bar{g}_{12} & 0 \\
\bar{g}_{21} & \bar{g}_{22} & 0 \\
0 & 0 & 1
\end{pmatrix}. \]
Here \( x^3 \in [-t/2, t/2] \), where \( t \) is the local thickness, and the \( x^3 = 0 \) plane corresponds to the mid-surface. We define the reduced two dimensional reference fundamental forms.
\[ \bar{\sigma}_{ab} = \tilde{\sigma}_{ab}, \quad \bar{\sigma}_{ab} = -\frac{1}{2} \partial_i \bar{\sigma}_{aij} |_{x^i = 0} \]

Using the reference fundamental forms we may express the three dimensional elastic energy in terms of a two dimensional energy density

\[ E = \int \int \mathcal{W}(g, \mathcal{G}) \sqrt{g} |dx^1 dx^2 dx^3 \]
\[ \approx \int \int \mathcal{W}_2D(a, \sigma, \tilde{\sigma}, \mathcal{N}) \sqrt{|\sigma|} dx^1 dx^2. \]

Carrying out a formal expansion of the elastic energy density in powers of the thickness, we obtain a reduced energy density,

\[ \mathcal{W}_2D(x^1, x^2) = \mathcal{W}_S(x^1, x^2) + \mathcal{W}_b(x^1, x^2), \quad (15) \]

where

\[ \mathcal{W}_S(x^1, x^2) = \frac{I}{8} \mathcal{A}^{ab\bar{b}} (a_{ab} - \bar{a}_{ab})(a_{\bar{a}\bar{b}} - \bar{a}_{\bar{a}\bar{b}}) \]
\[ \mathcal{W}_b(x^1, x^2) = \frac{E^3}{24} \mathcal{A}^{ab\bar{b}} (b_{ab} - \bar{b}_{ab})(b_{\bar{a}\bar{b}} - \bar{b}_{\bar{a}\bar{b}}), \]

and the reduced two dimensional elastic tensor is

\[ \mathcal{A}^{ab\bar{b}} = \frac{Y}{1 + \nu} \left[ \frac{1}{2} \left( \mathcal{G}^{ab} \mathcal{G}^{\alpha\beta} + \mathcal{G}^{\alpha\beta} \mathcal{G}^{ab} \right) + \frac{\nu}{1 - \nu} \mathcal{G}^{ab} \mathcal{G}^{ab} \right]. \]

Comments:

(1) The above reduced energy density is valid for slow spatially varying metric and thickness, and for small thickness. The correction terms to (15) are \( O(t^3) \), and \( O(t^2 |a - \bar{a}|) \).

(2) For compatible reference fundamental forms, \( \bar{a} \) and \( \bar{b} \), the above equations reduce to the Koiter shell energy.\(^{41}\)

(3) When \( b = 0 \), there are no structural variations across the thin dimensions. If the two dimensional metric, \( \bar{a} \), is associated with a non-vanishing Gaussian curvature, then \( \bar{a} \) and \( \bar{b} \) are incompatible. Such thin bodies were considered in detail in ref. 3 and 42, and were termed non-Euclidean plates.\(^{41}\)

(4) The above energy density was obtained through a formal asymptotic expansion. In particular the assumption that \( |a - \bar{a}| \ll 1 \) can only be justified in an \( L^2 \) sense rather than locally. The limit for non-Euclidean plates was proved rigorously by Lewicka and Pakzad;\(^{44}\) Kupferman and Solomon\(^{46}\) generalized this proof for all slender metrically incompatible bodies, including non-Euclidean shells and rods.

7.3 Examples of incompatible thin growing sheets

The shaping of thin living tissue can occur via various mechanisms that induce differential growth\(^{35}\) and differential or homogenous shrinkage and swelling.\(^{36}\) These processes endow the thin sheet with reference first and second fundamental forms, which in view of Section 5.1, need not be compatible with each other. The elastic equilibrium of such bodies is determined by the minimization of the total elastic energy as given in eqn (15). In the general case, the equilibrium configuration, where both the reference curvatures and the reference metric are not trivial, cannot be easily deduced from (15). We next present two simplified types of incompatible thin sheets.

The first type, termed non-Euclidean plates, exhibits a vanishing reference curvature but a non-Euclidean 2D metric. The second type is associated with an Euclidean 2D metric, but possess non-trivial reference curvatures.

7.3.1 Non Euclidean plates. If an initially flat stress-free sheet grows laterally, such that the growth profile does not vary across its thin dimension, then all sections parallel to the mid-surface of the sheets share the exact same geometry. In such a case, as can be easily deduced from symmetry consideration, the sheet possess zero reference curvature and is therefore plate-like. However, the two dimensional geometry of the mid-surface section may be non-Euclidean determining non-zero reference Gaussian curvature. In such cases, the generation of curvatures is required in order to relieve plane strain. Such bodies were termed non-Euclidean plates, and were treated extensively in ref. 3 and 44–46. Non-Euclidean plates can be used to model growing thin sheets for which the dominant shaping mechanism is planar growth. For very thin sheets, the elastic equilibrium configuration constitutes an isometry of the reference metric \( \bar{a} \). Such an isometry is usually not unique, and the equilibrium configuration is obtained by the bending energy minimizing isometry.

7.3.2 Euclidean sheets with non-trivial curvatures. Endowing a geometrically Euclidean sheet with a single non-zero constant principal curvature makes it into a simple cylindrical shell. If, however, two non-trivial principal curvatures are prescribed, then there is no Euclidean surface that can simultaneously obey them both. Consider a thin growing sheet composed of two laminae. If one of the laminae shrinks homogeneously and uni-axially (leaving lengths on one of the lateral dimensions unchanged), then a single reference principal curvature is induced and the sheet will adopt a cylindrical configuration. If, however, both laminae shrink uni-axially with different principal axes, then the sheet is endowed with two non-trivial uni-axial curvatures. Such a scenario occurs naturally in some seed pods dehiscence, as has been demonstrated for legumes.\(^{47}\)

If the sheet is very thin, the elastic equilibrium must be an isometry of the reference metric and will fully obey one of the reference curvatures, while setting the other to zero.\(^{47}\) Two such local equilibria exist, corresponding to the two different curvatures. If the sheet is not thin, its 2D metric may be violated and a wide variety of solutions exist.

8 Discussion

The metric description of elastic residually stressed solids was suggested by Kondo as early as 1955.\(^{35}\) In this work, the first of a series of memoirs, he writes about tensor calculus and Riemannian geometry.

“...it is strange that the first practical field of application was not the theory of elasticity, especially of residual strains.”
More than half a decade later the concept of residual stress is still commonly favored over the fundamental geometric concept of residual strain, and the isotropic invariants of the deformation gradient are commonly used instead of a Riemannian geometric description of elasticity.

Kondo blames general relativity for giving “a metaphorical appearance to Riemannian expression, banishing it for a time from the attention of engineers”. It is, however, general relativity that rendered the present covariant formulation accessible and popular. Another factor which contributed to the lack of use of the metric formulation is the immediate departure from Riemannian geometry in the description of plasticity in crystals.16,17

Two recent independent research themes have contributed to the revival of the metric description. The first is the advancements in dimensionally reduced theories. Due to the freedom of choice of natural curvilinear coordinates, and the ability to properly describe large rotations, the description of dimensionally reduced theories becomes relatively simple and transparent when done via the Riemannian metric description. As a result the metric description has been recently employed in “text book treatments” of compatible elasticity focusing on thin elastic bodies.25 The second theme is the success of mechanical modeling of some living tissue as amorphous elastic media.7 In such living tissues, due to the generation of residual stress, the Riemannian description is also very natural.

In this work we have attempted to elucidate the structure of the metric formulation of elasticity and show how it relates to other commonly used elastic formulations. We demonstrated different types of irreversible deformations using a Riemannian formulation, and showed the natural emergence of geometric frustration. We have also discussed the relation between the local property of geometric frustration and the resulting global property of residual stress.

While to some extent the choice between the metric formulation and the multiplicative formulation is a matter of taste and habit, we find the metric description natural to the elastic description of irreversibly deforming amorphous bodies, in that it provides an invaluable relation between the concept of residual stress and geometric frustration. We hope that this work render this approach more accessible.

A Two dimensional residual stress calculation in a uniformly frustrated system

In this section we demonstrate the main difference between residual stress and residual strain. While the latter is a local property, the former depends also on global properties such as the domain size and shape.

Let $\bar{g}$ be a 2D reference metric corresponding to a uniform Gaussian curvature, $\bar{K}$. We now seek an elastic energy minimizing configuration in the space of constant Gaussian curvature, $K = \text{const}$. We assume the domain of consideration and the parametrization such that both the reference metric and the embedding metric $g$ are uniformly close to the Euclidean metric, $|\bar{g} - I| \leq \delta$, and $|g - I| \leq \delta$.

To leading order in $\delta$ eqn (8) reduces to the Cartesian divergence equation.

$$\partial_i S^{ij} = 0,$$

which implies the existence of a scalar function $\Phi$ (Airy stress potential) such that

$$\partial_i \partial_i \Phi = S_{22}, \quad \partial_2 \partial_2 \Phi = S_{11}, \quad \partial_1 \partial_2 \Phi = -S_{12}.$$  

For simplicity we now consider a material with a vanishing Poisson ratio and set the Young’s modulus to unity. In such a case the bi-laplacian of the scalar function reads

$$\Delta^2 \Phi = K - K = -\Delta K,$$

where we have made use of the linearized Gaussian curvature (in leading order in $\delta$) where by

$$K = -\frac{1}{2} (\partial_i g_{12} + \partial_2 g_{i1} - 2 \partial_1 \partial_2 g_{12}).$$

Note the symmetry between embedding a hyperbolic surface in Euclidean space and the embedding of a flat surface on a positively curved space. We now consider a strip of length $L$ and width $w$ such that $w \ll L \ll L_{\text{geo}}$ where $L_{\text{geo}}$ is the smallest geometric lengthscale associated with the curvatures; $\frac{1}{\sqrt{K}}$ and $\frac{1}{\sqrt{-K}}$.

Assuming that away from the boundaries the solution will not depend on the coordinate along the long direction, $x^1$, we obtain

$$S_{11} = S_{12} = 0, \quad S_{22} = -\frac{\Delta K}{6} \left( \frac{1}{12} x^2 - \frac{w^2}{12} \right).$$

Upon integration we obtain for the elastic energy

$$E \propto (\Delta K)^2 w^5 L.$$  

The above scaling reads $E \propto w^4 A$ for strips of constant width and varying area, and scales as $E \propto \alpha^2 A^2$ for strips of constant aspect ratio $\alpha = w/L$. One can also solve the above equations is cylindrical geometry to recover the constant aspect ratio scaling $E \propto A^1$ under the assumption of axial symmetry. These results obtained for almost flat geometries and general incompatibility extent the results of Schneider and Gompper obtained for the energy scaling of defect free crystalline domains of positively curved surfaces.48

B The multiplicative decomposition of deformation gradients and its relation to the metric description

In the Bilby–Kroner–Lee multiplicative decomposition of deformation gradients31-33 the deformation gradient, $F$, (defined in eqn (3)) is decomposed into a product of an elastic part $F^e$ and a plastic part $F^p$, namely $F = F^e F^p$. The plastic “gradient
like" term, \( F^0 \) is associated with the irreversible deformation of an initial rest configuration into a virtual configuration, whereas the elastic "gradient like" term, \( F^e \), is associated with an elastic relaxation from the virtual configuration to the actual configuration. In cases where the virtual configuration is not realizable, neither \( F^e \) nor \( F^0 \) are gradients of a map in \( \mathbb{R}^3 \), which is why we call them "gradient like". Note however, that their product \( F \) is always a gradient of a map in \( \mathbb{R}^3 \). This multiplicative decomposition has also been adopted for the description of growth in living tissue.\(^{1,20}\)

Lee claimed that in general one cannot consistently decompose additively a given strain into an elastic strain and a plastic strain.\(^{24}\) In what follows we interpret the metric description in view of these claims (aimed primarily against a description put forward by Green and Naghdi\(^{49}\)). We show the equivalence between the multiplicative decomposition of growth strains and the metric description. We then consider two of the difficulties and inconsistencies that arise in the additive decomposition of strains as elaborated in ref. 24 and show how a proper interpretation of the metric description alleviates these inconsistencies. We do not claim any novelty of the arguments presented here, but rather set to show that they become transparent and sometimes trivial when considered within the metric description of hyper-elasticity.

### B.1 Relating the various measures of deformation

A common description of growth and plasticity follows the multiplicative decomposition of deformation gradients.\(^{1,20}\) Within this description an initially unstressed configuration, \( x \in \mathcal{D} \), is deformed to assume a final configuration \( r \in \mathcal{M} \). The deformation is considered to be composed of an irreversible process mapping the unstressed configuration to a "virtual configuration" followed by an elastic relaxation from the "virtual configuration" to the final configuration.

The total deformation is associated with the mapping \( r(x) : \mathcal{D} \rightarrow \mathcal{M} \). The deformation gradient is given by \( F = \nabla_r x \). The mathematical formulation of the above decomposition reads

\[
F = F^0 F^e,
\]

where we have denoted the irreversible (taken in the context of ref. 23 to be plastic) deformation gradient like term by \( F^0 \), and the elastic response gradient like term by \( F^e \). Within the multiplicative decomposition of deformations the elastic energy of an amorphous body is assumed to be a function of the isotropic invariants of the elastic part of the deformation gradient, \( F^e \).

In the treatment by Green and Naghdi three strains are defined; the total strain \( \varepsilon^{tot} = \frac{1}{2} ((F^e)^T (F^e) - I) \), the plastic strain \( \varepsilon^{pl} = \frac{1}{2} ((F^0)^T (F^0) - I) \), and the elastic strain defined through their difference which can be identified with the Green St-Venant strain:

\[
\varepsilon^e = \varepsilon^{tot} - \varepsilon^{pl} = \frac{1}{2} ((F)^T (F) - (F^e)^T (F^e)) = \frac{1}{2} (g - \bar{g})
\]

### B.2 The path independent definition of the strain

Consider the a solid body at rest as in Fig. 2(a). For convenience we set the height and width of the body to unity. Now allow the body to be plastically deformed into the shape (c) by a shear deformation quantified by \( \delta = \tan(\theta) \). Now stretch the body by a factor of two along the y direction to arrive at the configuration (d).

Consider now the alternative route by which the body is first elastically stretched by a factor of two along the y direction to yield the configuration denoted in (b), and only then introduced with the plastic shear bringing it into (d). In this second path the plastic shear angle is \( \alpha = \arctan(\frac{1}{2} \tan(\theta)) \neq \theta \), as the same horizontal displacement occurs at twice the height. Thus, the two paths mapping the un-deformed configuration (a) into (d), associate this mapping with two different plastic strains, although the total strain and the elastic strain are the same in both paths.

Now consider the same two paths in the metric description. In a pure elastic deformation the reference metric, \( \bar{g} \), is constant. Whereas in a pure plastic deformation the elastic strain \( \varepsilon = \frac{1}{2} (g - \bar{g}) \) remains constant. We choose the Lagrangian coordinates only once as Cartesian in the configuration (a) and follow these coordinates as the body deforms. In the configuration (a) we have

\[
\varepsilon^e = \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \quad \bar{g}^e = \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \quad \varepsilon^a = \left( \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right).
\]

We may calculate the metric of the configuration at (c) explicitly. As the deformation is purely plastic \( \varepsilon^p = \varepsilon^a \) which determines \( \bar{g}^p \).

\[
\varepsilon^e = \left( \begin{array}{cc} \tan(\theta) & \cos^2(\theta) \\ \cos(\theta) & \sin^2(\theta) \end{array} \right), \quad \bar{g}^e = \left( \begin{array}{cc} 1 & \tan(\theta) \\ \tan(\theta) & \cos^2(\theta) \end{array} \right), \quad \varepsilon^e = \varepsilon^a = \left( \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right).
\]

Fig. 2 A combined elastic-plastic deformation: (a) the rest configuration of a body. (b) Uniaxial elastic stretching by a factor of two of the rest configuration. (c) Plastic shearing of the rest configuration. (d) A combined uniaxial elastic stretching and a plastic shearing of the rest configuration.
As the deformation (c) → (d) is purely elastic $\tilde{g}$ remains unaltered. We calculate the metric explicitly to obtain

$$
\tilde{g}^d = \begin{pmatrix}
1 & 2 \tan(\alpha) \\
2 \tan(\alpha) & 4 \cos^2(\alpha)
\end{pmatrix},
$$

$$
\tilde{g}^d = \tilde{g} = \begin{pmatrix}
1 & \tan(\theta) \\
\tan(\theta) & \cos^2(\theta)
\end{pmatrix},
$$

$$
\epsilon^d = \frac{1}{2} (g^d - \tilde{g}^d).
$$

Taking the alternative route, the deformation (a) → (b) is purely elastic, thus $\tilde{g}$ remains unaltered.

$$
\tilde{g}^b = \begin{pmatrix}
1 & 0 \\
0 & 4
\end{pmatrix},
\tilde{g}^b = \tilde{g} = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix},
\epsilon^b = \frac{1}{2} \begin{pmatrix}
0 & 0 \\
0 & 3
\end{pmatrix}.
$$

The purely elastic deformation (b) → (d) leaves the elastic strain unchanged thus we have

$$
\tilde{g}^d = \begin{pmatrix}
1 & 2 \tan(\alpha) \\
2 \tan(\alpha) & 4 \cos^2(\alpha)
\end{pmatrix},
$$

$$
\epsilon^d = \frac{1}{2} (g^d - \tilde{g}^d) = \epsilon^b = \frac{1}{2} \begin{pmatrix}
0 & 0 \\
0 & 3
\end{pmatrix}.
$$

We may therefore deduce $g^d = g^d - 2 \epsilon^d$ through

$$
\tilde{g}^d = \begin{pmatrix}
1 & 2 \tan(\alpha) \\
2 \tan(\alpha) & 4 \cos^2(\alpha) - 3
\end{pmatrix} = \begin{pmatrix}
1 & \tan(\theta) \\
\tan(\theta) & \cos^2(\theta)
\end{pmatrix},
$$

where in the last equality we have made use of the fact that, $2\tan(\alpha) = \delta = \tan(\theta)$. The two paths lead to the exact same reference metric.

### B.3 Strain decoupling and memory loss in the metric description

Consider an elastic body whose energy density is given by (13) undergoing a uniaxial uniform deformation in three stages. In the first stage only elastic extension occurs, i.e. $\tilde{g} = 0$. In the second stage the elastic extension is accompanied by a plastic response of the form

$$
\tilde{g}_{ij} = \frac{1}{2Y} S_{ij} = \frac{1}{2} (g_{ij} - \tilde{g}_{ij}).
$$

(16)

In the third stage again only elastic deformations are used to bring the body to a state of zero stress (while $\tilde{g} = 0$).

We choose the uniaxial elongation to occur along the $\hat{x}$ direction, and impose the $x = 0$ surface of the body to remain stationary. Only the face opposite to the face at $x = 0$ is loaded, and we denote its $x$ coordinate by $X$. The total force exerted on the body may be given by $f = f_{x}\hat{x}$, where

$$
f_{x} = -\partial E/\partial X,
$$

and $E$ above is the elastic energy (13) expressed as a function $X$ alone.

The stress is naturally defined as a contravariant tensor (through the Frechet derivative of a scalar with respect to a covariant tensor in (7)). It therefore seems natural to study a uniaxial deformation through comparing $\tilde{S}^{xx}$ and $g_{xx}$ or $\tilde{e}^{xx}$. However, from the tensorial formulation we know that such a relation will not yield a simple scalar but a four rank tensor. This is of course due to the different units of length associated with each type of tensor. When we seek a simple scalar (such as local material stiffness), we must therefore compare tensors of the same type as done below in Fig. 3.

### Acknowledgements

We would like to thank R. Knops for bringing the early works of Kondo to our attention. E.E. acknowledges the supported of the Simons foundation. R.K. was partially funded by the Israel Science Foundation and by the Israel-US Binational Foundation. E.S. was supported by the European Research Council SoftGrowth project.

### References