Determining a Surrogate Contact Pair in a Hertzian Contact Problem

Anthony P. Sanders
Ortho Development Corp.,
12187 S. Business Park Dr.,
Draper, UT 84020
e-mail: tsanders@odev.com

Rebecca M. Brannon
Department of Mechanical Engineering,
University of Utah,
50 S. Central Campus Dr.,
2134 Merrill Engineering Bldg.,
Salt Lake City, UT 84112
e-mail: brannon@mech.utah.edu

Laboratory testing of contact phenomena can be prohibitively expensive if the interacting bodies are geometrically complicated. This work demonstrates means to mitigate such problems by exploiting the established observation that two geometrically dissimilar contact pairs may exhibit the same contact mechanics. Specific formulas are derived that allow a complicated Hertzian contact pair to be replaced with an inexpensively manufactured and more easily fixtured surrogate pair, consisting of a plane and a spheroid, which has the same (to second-order accuracy) contact area and pressure distribution as the original complicated geometry. This observation is elucidated by using direct tensor notation to review a key assertion in Hertzian theory; namely, geometrically complicated contacting surfaces can be described to second-order accuracy as contacting ellipsoids. The surrogate spheroid geometry is found via spectral decomposition of the original pair's combined Hessian tensor. Some numerical examples using free-form surfaces illustrate the theory, and a laboratory test validates the theory under a common scenario of normally compressed convex surfaces. This theory for a Hertzian contact substitution may be useful in simplifying the contact, wear, or impact testing of complicated components or of their constituent materials. [DOI: 10.1115/1.4003492]

Keywords: contact mechanics, Hertzian theory, elliptical contact, contact testing, wear testing, substitute contact

1 Introduction

Hertz's theory of elastic contact was originally introduced in 1882 [1]. Johnson [2] noted that the theory met with early appreciation and has "stood the test of time." Notwithstanding myriad pertinent developments in the intervening century, Hertz's geometric propositions are somewhat obfuscated by the common, but cumbersome, use of scalar notation where tensor notation would be clearer [3]. Also, the theory's geometric details lead quite naturally to the concept that the local deformation and stress fields of a complicated contact pair may be replicated to second-order accuracy by a simpler pair having the same relative curvatures. This is a known, but not yet developed or implemented, aspect of the prior work that offers practical benefits of simplifying potentially difficult contact and wear testing projects. Accordingly, the present article has two specific aims:

1. To show that direct tensor notation improves clarity of the geometric propositions in the Hertz contact theory. The theory is underpinned by two propositions that describe the geometry of two smooth contacting surfaces:

- each surface can be described with two principal curvatures lying on mutually orthogonal planes
- the distance between opposing points on the contacting surfaces can be described by a quadratic form, which implies that lines of constant separation are ellipses

Despite seeming to be very limiting, these propositions are, in fact, second-order accurate for arbitrary differentiable surfaces. In this work, a novel tensor description of the geometric problem will be used to clarify and justify these propositions.

2. To prove that an arbitrarily shaped Hertzian contact pair can be substituted with a simpler pair consisting of a spheroid and a flat plane, without altering the original contact area or contact pressure. As shown by upcoming Eqs. (21), (19), and (13), the dimensions of the spheroid can be readily computed using simple functions of the original pair's geometry. Verification is offered by geometric analysis and contact stress analysis of a pair of complicated 3D surfaces. A laboratory validation demonstrates the substitution theory with convex surfaces. This substitution concept clearly has potential to reduce costs of contact and wear testing by using simply manufactured and easily fixtured surrogate geometries.

2 Geometry in Hertzian Contact

2.1 The Assumptions Concerning Geometry. Assumptions behind Hertzian theory have been amply justified [2,4,5]. Two of the key geometric propositions that describe the contacting surfaces are mathematical statements:

1. Each surface may be described by an equation that includes the surface's two principal curvatures at the contact point
$$z = \tfrac{1}{2}\rho_1 x^2 + \tfrac{1}{2}\rho_2 y^2 \tag{1}$$
2. Closed curves on the two surfaces that are separated by a constant distance h have a projection onto the tangent plane that is an ellipse:
$$h = Ax^2 + By^2 \tag{2}$$

The focus of this section is to elucidate these descriptions via a novel tensor formulation of Hertzian geometry and to illustrate their second-order accuracy in describing the geometry of contact.

2.2 Description of a Single Surface. We begin with a description of a general surface in R^3 and several vectors associated with a point on that surface (Fig. 1(*a*)). An orthonormal coordinate basis with base vectors $\mathbf{e}_i$ is introduced for convenience. Vector $\mathbf{x}$ describes a variable point on the surface, and vector $\mathbf{p}$ describes the contact point. An analytic function f describes the surface as the set of all points $\mathbf{x}$ satisfying the equation $f(\mathbf{x})=0$. The normal at $\mathbf{p}$ is $\hat{\mathbf{n}}=\nabla f/\|\nabla f\|$. The vector from $\mathbf{p}$ to another point on the surface is $(\mathbf{x}\text{-}\mathbf{p})$, and the vector $\mathbf{r}$ is defined as the part of $(\mathbf{x}\text{-}\mathbf{p})$

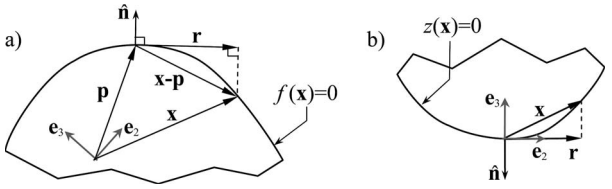

Fig. 1 (*a*) Section view through a general surface. (*b*) Section view through a surface with the laboratory basis positioned at the point of contact.

Contributed by the Tribology Division of ASME for publication in the JOURNAL OF TRIBOLOGY. Manuscript received March 21, 2010; final manuscript received January 12, 2011; published online March 17, 2011. Assoc. Editor: Ilya I. Kudish.

lying in the plane tangent to the surface at **p**.

For convenience, the basis is positioned at the initial contact point, with one base vector aligned opposite to the surface normal (Fig. 1(*b*)). The other two base vectors span the tangent plane. Disallowing cusps in the surface, the scalar-valued function for the height $z(\mathbf{r})$ of the surface above the tangent plane may be expanded in a Taylor series about the origin as follows:

$$z(\mathbf{r}) = z(\mathbf{0}) + \mathbf{z}'(\mathbf{0}) \bullet \mathbf{r} + \tfrac{1}{2}\mathbf{z}''(\mathbf{0}):\mathbf{rr} + \cdots \tag{3}$$

Here, ":" denotes the tensor inner product operation. The vector- and tensor-valued functions are

$$\mathbf{z}' = \nabla z \text{ and } \mathbf{z}'' = \nabla\nabla z \text{ or: } z_i = \frac{\partial z}{\partial x_i} \text{ and } z_{ij} = \frac{\partial^2 z}{\partial x_i\, \partial x_j} \quad i,j = 1,2 \tag{4}$$

The higher-order terms are neglected in view of the typical assumption that the contact region is small. The first term in the series is zero because the surface is positioned at the origin; i.e., $z(\mathbf{0})=0$. Moreover, the second term is zero because $\mathbf{z}'(\mathbf{0})$ is normal to the surface while **r** is tangent to the surface. Thus, Eq. (3) reduces to one quadratic form $z(\mathbf{r}) \approx \frac{1}{2}\mathbf{z}'':\mathbf{rr}$. Here, $\mathbf{z}''$ is understood to be $\mathbf{z}''(\mathbf{0})$. The matrix and indicial forms are

$$z(\mathbf{r}) \approx \tfrac{1}{2}\{\mathbf{r}\}^{\mathrm{T}}[\mathbf{z}'']\{\mathbf{r}\} \quad \text{and} \quad z(r_1,r_2) = \tfrac{1}{2} r_i (\mathbf{z}'')_{ij} r_j \tag{5}$$

where $\{\mathbf{r}\}$ is the 2×1 array of in-plane components of **r**, and $[\mathbf{z}'']$ is the 2×2 component matrix of $\mathbf{z}''$, and repeated indices are implicitly understood to be summed from 1 to 2. Since $[\mathbf{z}'']$ is real and symmetric, there exists an orthonormal principal basis in which it will be diagonal with its eigenvalues as the diagonal components. This matrix $[\mathbf{Z}]$ is related to $[\mathbf{z}'']$ via the basis transformation

$$[\mathbf{z}''] = [\mathbf{Q}][\mathbf{Z}][\mathbf{Q}]^{\mathrm{T}} \tag{6}$$

Here, $[\mathbf{Q}]$ is an orthogonal 2×2 direction cosine matrix whose columns contain the components of the orthonormalized eigenvectors of $[\mathbf{z}'']$. The matrix $[\mathbf{z}'']$ is also known as the *Hessian* matrix [6]. Its eigenvalues are the principal curvatures (η_1 and η_2) of the surface at $\mathbf{r}=\mathbf{0}$. Symmetry of $[\mathbf{z}'']$ ensures that principal curvatures are located on two mutually orthogonal planes.

Thus, Eq. (1) is an approximation based on a Taylor series expansion [3], and the appearance of the principal curvatures results from a spectral analysis of $[\mathbf{z}'']$. The orthogonal orientation between the principal curvature directions stems from the fact that the curvatures are eigenvalues of a symmetric Hessian tensor $\mathbf{z}''$. Substituting Eq. (6) into Eq. (5) yields an equation of the form proposed in Eq. (1) as follows:

$$z = \tfrac{1}{2}\{\mathbf{r}\}^{\mathrm{T}}([\mathbf{Q}][\mathbf{Z}][\mathbf{Q}]^{\mathrm{T}})\{\mathbf{r}\} = \tfrac{1}{2}\{\mathbf{y}\}^{\mathrm{T}}[\mathbf{Z}]\{\mathbf{y}\} \tag{7}$$

where $\{\mathbf{y}\}=[\mathbf{Q}]^{\mathrm{T}}\{\mathbf{r}\}$ so that, because $[\mathbf{Z}]$ is diagonal

$$z = \tfrac{1}{2}(\eta_1 y_1^2 + \eta_2 y_2^2) \tag{8}$$

The y_i are the components of **r** with respect to an orthonormal basis aligned with the eigenvectors of $[\mathbf{z}'']$.

2.3 Distance Between Surfaces. Equation (2) pertains to two contacting surfaces (Fig. 2). The principal basis of Surface *K* is the **a**-basis, and that of Surface *P* is the **b**-basis. The following shows in direct tensor notation that points (e.g., *k* and *p*) separated by a fixed distance, $z=z_K+z_P$, form a locus given by a quadratic form analogous to Eq. (2).

Let $\boldsymbol{\kappa}$ represent the Hessian ($\mathbf{z}''$) tensor of Surface *K*, and $\boldsymbol{\rho}$ that of *P*. Thus, Eq. (5) for Surface *K* is $z_K=\frac{1}{2}\mathbf{r}\bullet\boldsymbol{\kappa}\bullet\mathbf{r}$, and for Surface *P* is $z_P=\frac{1}{2}\mathbf{r}\bullet\boldsymbol{\rho}\bullet\mathbf{r}$. Thus, the total separation distance of two points *k* and *p*, $z=z_K+z_P$, is

$$z = \mathbf{r}\bullet\mathbf{A}\bullet\mathbf{r} \quad \text{where} \quad \mathbf{A} = \tfrac{1}{2}(\boldsymbol{\kappa}+\boldsymbol{\rho}) \tag{9}$$

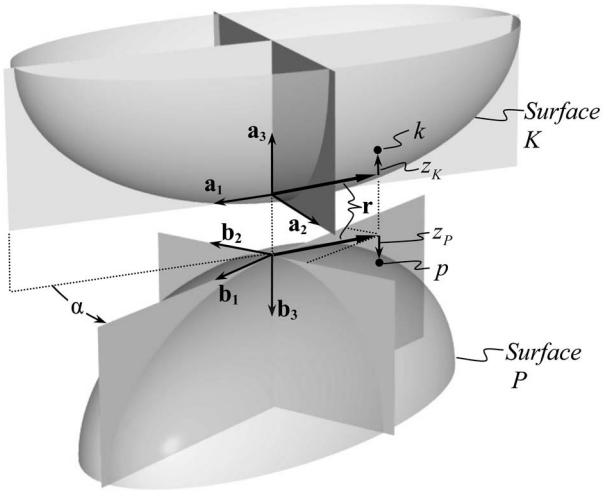

Fig. 2 Contacting curved surfaces (separated for clarity). Section planes contain principal curvatures. For clarity, points *k* and *p* are shown at an exaggerated distance from the origin.

Each Hessian tensor is symmetric; therefore, tensor **A** in Eq. (9) is symmetric. The eigenvectors of **A** are not generally aligned with those of $\boldsymbol{\kappa}$ or $\boldsymbol{\rho}$. Since Eq. (9) is in direct notation, it is valid in any basis; with respect to the principal basis of **A**, it may be written in component form as

$$z = \{\mathbf{y}\}^{\mathrm{T}}[\mathbf{D}]\{\mathbf{y}\} = \lambda_1 y_1^2 + \lambda_2 y_2^2 \tag{10}$$

Here, $[\mathbf{D}]$ is a diagonal matrix whose diagonal components, λ_1 and λ_2, are the eigenvalues of $[\mathbf{A}]$, and y_1 and y_2 are the components of **r** with respect to an orthonormal basis aligned with the eigenvectors of $[\mathbf{A}]$. Equation (10) describes an ellipse; ordering the eigenvalues $\lambda_1<\lambda_2$, and then substituting

$$c = \sqrt{z/\lambda_1} \quad \text{and} \quad d = \sqrt{z/\lambda_2} \tag{11}$$

yields the common form $y_1^2/c^2+y_2^2/d^2=1$, where $c>d$. This describes an ellipse centered at $\mathbf{r}=\mathbf{0}$, with its major semiaxis *c* aligned with the eigenvector associated with λ_1 and its minor semiaxis *d* aligned with the eigenvector of λ_2. We term a locus of points satisfying Eq. (10) for a particular *z* as a *tangent ellipse*. For all such points **r**, there are corresponding points $z_K(\mathbf{r})$ and $z_P(\mathbf{r})$ on *K* and *P*. We call these loci *separation curves* (Fig. 3). Viewed along the normal axis, the separation curves are coincident with the tangent ellipse.

2.4 Spectral Analysis. Spectral analysis of $[\mathbf{A}]$ is used to find its eigenvalues. Selecting a basis that coincides with the principal basis of Surface *K* (the **a**-basis), the component form of $[\mathbf{A}]$ is

$$[\mathbf{A}] = \frac{1}{2}\left(\begin{bmatrix}\kappa_1 & 0\\ 0 & \kappa_2\end{bmatrix} + \begin{bmatrix}\cos\alpha & \sin\alpha\\ \sin\alpha & -\cos\alpha\end{bmatrix}\begin{bmatrix}\rho_1 & 0\\ 0 & \rho_2\end{bmatrix}\begin{bmatrix}\cos\alpha & \sin\alpha\\ \sin\alpha & -\cos\alpha\end{bmatrix}\right) \tag{12}$$

Here, the principal curvatures of $\boldsymbol{\kappa}$ and $\boldsymbol{\rho}$ are ordered such that $|\kappa_1|<|\kappa_2|$ and $|\rho_1|<|\rho_2|$. The orientation angle α shown in Fig. 2 is the angle between principal directions of $\boldsymbol{\kappa}$ and $\boldsymbol{\rho}$. The eigenvalues of $[\mathbf{A}]$ are

$$\lambda_1 = \tfrac{1}{4}(\Sigma - \Delta) \quad \text{and} \quad \lambda_2 = \tfrac{1}{4}(\Sigma + \Delta) \quad \text{where}$$

$$\Sigma = \kappa_1 + \kappa_2 + \rho_1 + \rho_2 = 2\ \mathrm{tr}(\mathbf{A})$$

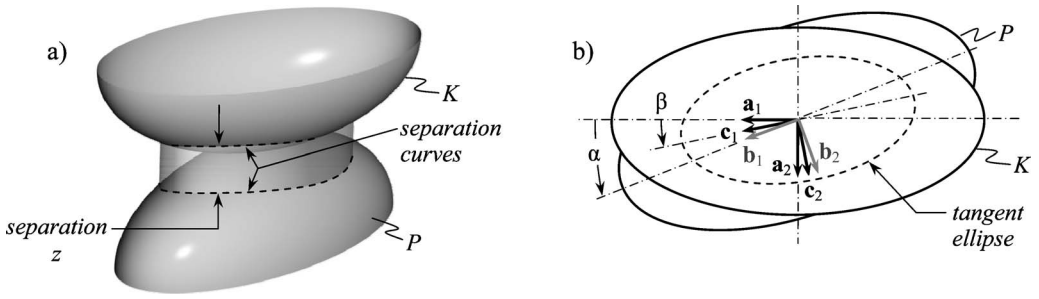

Fig. 3 (*a*) Separation curves, joined by a transparent surface for visual reference. (*b*) In plan view, separation curves overlie the tangent ellipse; also, a c-basis is aligned with the principal axes of the tangent ellipse.

$$\Delta = [(\kappa_1 - \kappa_2)^2 + (\rho_1 - \rho_2)^2 + 2(\kappa_1 - \kappa_2)(\rho_1 - \rho_2)\cos(2\alpha)]^{1/2} \tag{13}$$

In summary, Eq. (2) is derivable using a symmetric tensor **A** composed of the principal curvatures of the two surfaces. Spectral analysis of this tensor reveals the key dimensions of the tangent ellipse: its major and minor semiaxis lengths. This derivation illustrates how the separation distance and the tangent ellipse can be determined from the surface configuration using direct tensor notation, which is clearer than scalar analysis because of the intrinsic invariance of tensors to basis change.

2.5 Verification. Numerical simulations were performed to examine the accuracy of the derived formulas. The simulation objects were arbitrarily curved surfaces created using computer-aided design (CAD) software. First, a curved surface was created, as in Fig. 4(*a*). Then, portions of the surface were extracted and transformed into an initial contact configuration, as in Fig. 4(*b*). Both surfaces were measured in the software to determine the normal vectors, principal curvatures, and principal curvature orientations (Table 1). The laboratory basis was positioned at the contact point and aligned with one surface's normal and principal curvatures.

Next, one surface was translated a small distance δ into the other along the normal. The intersection curve between the surfaces was equivalent to a separation curve for the distance δ. This 3D curve was projected onto the tangent plane at the origin, forming a 2D curve that was termed a *quasi-ellipse* because corresponding semiaxes were unequal in length. A sequence of such translations (Sim 1) was compiled for a range of δ. For each, the lengths of both of the quasi-ellipse's major semiaxes were recorded. In a second simulation (Sim 2), one member was concave.

For comparison with each constructed quasi-ellipse, the corresponding tangent ellipse was determined. Its major semiaxis length was computed using Eqs. (13) and (11) along with the data in Table 1 and the values of δ as data for z in Eq. (11). Figure 5 illustrates the second-order accuracy of the computed dimensions; it shows that the difference between the quasi-ellipse (generated from arbitrary surfaces) and the tangent ellipse (from the theory) tends toward zero as the separation distance decreases.

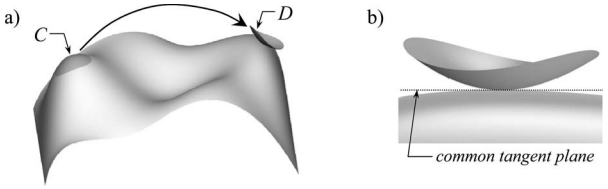

Fig. 4 (*a*) Free-form virtual surface; region at *C* extracted and transformed to contact the original surface at *D*. (*b*) Surfaces at *D* (viewed along tangent plane) initially touch at a point.

3 Systematic Determination of a Surrogate Contact Pair

We seek to replace an original, presumably complicated, Hertzian contact pair with a surrogate pair that will exhibit the same Hertzian contact mechanics as the original pair to second-order accuracy. The replacing pair should meet two design conditions to fulfill this purpose. Condition 1 is that prior to loading, the replacing pair should exhibit the same separation curves as the original pair. Thus, from Eq. (10), the replacing pair should have the same constants λ_1 and λ_2 as the original pair; hence, these are combined into an equality expressing the following condition:

$$\lambda_{1R} + \lambda_{2R} = \lambda_{1O} + \lambda_{2O} \tag{14}$$

Here, subscripts R and O denote the replacing pair and the original pair, respectively.

Condition 2 is that upon loading, the replacing pair must exhibit the same contact area to second-order accuracy, given identical normal force, materials, etc. The tangent ellipse due to a constant separation distance z in Eq. (10) is not identical to the ellipse that borders the loaded contact area (termed the *contact ellipse*). Instead, the two ellipses are related by the following equation, written in terms of the original pair [2]:

$$\frac{\lambda_{2O}}{\lambda_{1O}} = \frac{(a/b)^2\mathbf{E}(e) - \mathbf{K}(e)}{\mathbf{K}(e) - \mathbf{E}(e)} \tag{15}$$

Here, a and b are the major and minor semiaxes of the original

Table 1 Principal curvatures and orientation for each simulation

Sim	Surface 1 (mm^{-1}) κ_1	Surface 1 (mm^{-1}) κ_2	Surface 2 (mm^{-1}) ρ_1	Surface 2 (mm^{-1}) ρ_2	α (rad)
1	0.1821	0.2849	0.0454	0.3076	$\pi/6$
2	−0.0298	−0.0995	0.1014	0.2932	$\pi/9$

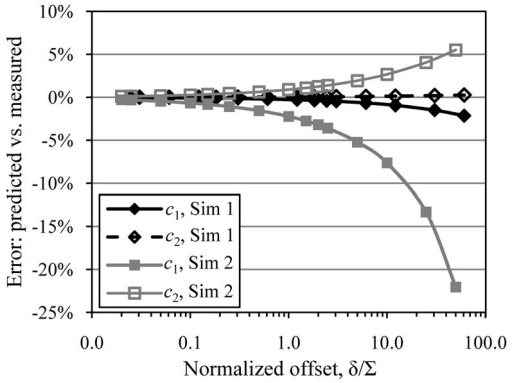

Fig. 5 Major semiaxis lengths from Sims 1 and 2, compared with predictions. There were two values for each semiaxis (i.e., c_1 and c_2) since the quasi-ellipses' complementary semiaxes were not identical.

pair's contact ellipse, $e=\sqrt{1-(b^2/a^2)}$ is the *eccentricity*, and **K** and **E** are the complete elliptic integrals of the first and second kinds, respectively. Condition 2 means that the replacing pair must also satisfy Eq. (15), from which a second equality is deduced as follows:

$$\lambda_{2R}/\lambda_{1R}=\lambda_{2O}/\lambda_{1O} \tag{16}$$

In the replacing pair, the second body is chosen to be a plane; thus, $\rho_{1R}=\rho_{2R}=0$. Then, using Eq. (16), Eq. (14) reduces to

$$\kappa_{1R}+\kappa_{2R}=\Sigma_O \tag{17}$$

and Eq. (16) reduces to

$$\frac{\kappa_{2R}}{\kappa_{1R}}=\frac{\Sigma_O+\Delta_O}{\Sigma_O-\Delta_O} \tag{18}$$

Solving these simultaneously yields the following:

$$\kappa_{1R}=\tfrac{1}{2}(\Sigma_O-\Delta_O),\quad \kappa_{2R}=\tfrac{1}{2}(\Sigma_O+\Delta_O) \tag{19}$$

where κ_{1R} and κ_{2R} are the principal curvatures of the first body in the replacing pair. The quantities Σ_O and Δ_O are simple functions of the original pair's configuration, per Eq. (13). Equation (19) shows that the principal curvatures of the first body in the replacing pair can be expressed entirely in terms of known geometric properties of the original surfaces. By using a lathe, a prolate spheroid can be inexpensively produced to provide the two required curvatures. The spheroid is described by the equation

$$\frac{x^2}{r_1^2}+\frac{y^2}{r_2^2}+\frac{z^2}{r_1^2}=1 \quad \text{where} \quad r_2>r_1 \tag{20}$$

To form the replacing pair, this spheroid would be placed in contact with its planar counterpart, with its z axis normal to the plane. For this configuration, the spheroid's defining radii can be computed as

$$r_1=1/\kappa_{2R},\quad r_2=\sqrt{1/(\kappa_{1R}\kappa_{2R})} \tag{21}$$

The result in Eq. (21) provides a means to substitute an arbitrary contact pair with a surrogate pair, consisting of a spheroid and a plane, which will provide the same contact area and pressure to second-order accuracy. This concept of Hertzian substitution is illustrated in Fig. 6.

It is alternatively possible to determine the dimensions of a surrogate pair consisting of a cylinder and a spheroid. Choosing a value for the radius of the replacing cylinder r_2 yields the cylinder principal curvatures $\rho_{1R}=0$ and $\rho_{2R}=1/r_2$. Then, the curvatures of the replacing spheroid can be computed as

$$\kappa_{1R}=\tfrac{1}{2}(\Sigma_O-\Delta_O),\quad \kappa_{2R}=\tfrac{1}{2}(\Sigma_O+\Delta_O-2\rho_{2R}) \tag{22}$$

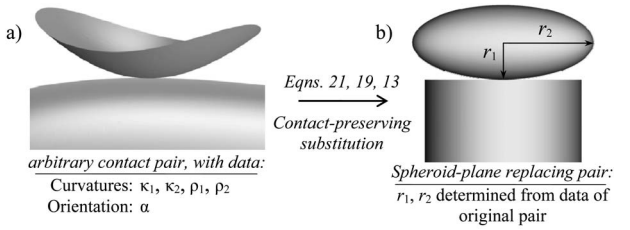

Fig. 6 Hertzian substitution concept: An arbitrary contact pair (*a*), with given principal curvatures and orientation, is substituted with a simpler contact pair (*b*) consisting of a spheroid and a plane

3.1 Verification. To verify these analytical results, the free-form virtual contact pair of Fig. 6(*a*) was taken as an original pair, and it was compared with its replacing pair of Fig. 6(*b*). In the following simulations, the surfaces were built, manipulated, and measured in the CAD software.

Sim A. Comparison of tangent ellipse dimensions to demonstrate the second-order equivalence of the replacing pair as a surrogate for the original pair.

Sim B. Comparison of contact ellipse dimensions to substantiate use of the replacing pair as a surrogate for the original pair in testing scenarios.

The geometry of the replacing pair was computed from the original pair's data using Eqs. (21), (19), and (13). The data and results are given in Table 2.

In Sim A, the dimensions of the tangent ellipses were computed using Eqs. (13) and (11) with the data in Table 2. Results were computed across a range of z, and at every value, the difference between the semiaxis lengths was zero (Fig. 7(*a*)). This parity between the tangent ellipses demonstrates that the original and replacing pairs are equivalent contact pairs to second-order accuracy.

In Sim B, the virtual surfaces were assumed to be solid bodies, and both contact pairs were examined by Hertzian contact analysis for normal loading. The first body in each pair was assigned an elastic modulus of 200 GPa and Poisson's ratio of 0.3, and the second body was assigned values of 150 GPa and 0.3. The dimensions of the contact ellipse and the peak contact pressure were computed for a range of loads, using formulas described elsewhere [2]. The results, given in Fig. 7(*b*), were identical for the two pairs, which further demonstrates their equivalence as Hertzian contact pairs.

3.2 Validation. Experimental validation was performed by comparing the contact ellipses of several different original pairs with those of their replacing pairs. Results from one such trial are presented here.

A novel "fingerprinting" technique was developed for recording a contact patch. Specifically, one surface was dabbed with grease, and the grease was wiped repeatedly (e.g., 20 times) with clean paper towels to leave a scant film. When the two surfaces were

Table 2 Geometry of original and replacing pairs for Sims A and B (curvatures in mm^{-1} and radii in mm)

Original pair, with $\alpha=\pi/6$			
κ_1	κ_2	ρ_1	ρ_2
0.1820980	0.2848800	0.0453731	0.3076450
Replacing pair			
κ_{1R}	κ_{2R}	r_1	r_2
0.246974	0.573022	1.745133	2.658207

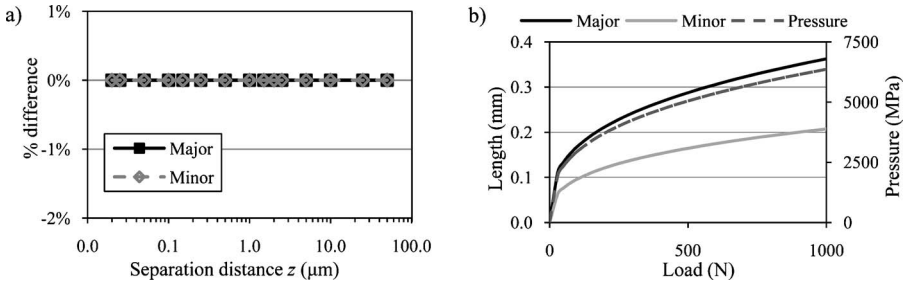

Fig. 7 Simulation results. (*a*) Sim A tangent ellipses: results show zero difference. (*b*) Sim B: Contact analysis: results for semiaxis lengths and pressure were identical for the pairs.

loaded, a thin layer of the grease was transferred onto the originally clean surface at the contact patch. After the test, this surface was sprinkled with photocopier toner powder, and then the powder was blown off with dry compressed air. This left black powder adhered to the transferred grease, which provided a record of the contact patch with better defined boundaries than could be obtained with commercially available methods for marking contact patches. The contact patch was then measured and photographed using an optical coordinate measuring machine (Nexiv, Nikon, Japan).

In one validation test, the original pair consisted of two $\varnothing 37.5 \times 76$ mm cylinders made of extruded polyetherimide plastic. The cylinders were arranged one on top of the other, with their axes mutually angled at 45 deg. The replacing pair consisted of a spheroid mated with a flat disk, with the spheroid's dimensions computed using Eq. (21). All surfaces were polished to a surface texture of ~ 0.2 μm roughness average (R_a). Each pair was compressed under loads from 200 N to 600 N. The major and minor axes of the contact patches were measured, and the results are compared in Fig. 8.

4 Discussion

The key geometric propositions in Hertz's theory, which are sometimes obfuscated through the use of scalar notation, have been presented here using a concise and relatively intuitive tensor notation that groups related properties that would otherwise appear as separate variables. For example, the Hessian tensor in Eq. (4) is represented compactly by $\mathbf{z}''$, and trigonometric terms of the orientation angle α are not introduced until the combined Hessian tensor $\mathbf{A}$ is expressed in component form in Eq. (12). The use of a truncated, multivariate Taylor series expansion to represent the distance between the contacting surfaces rigorously showed the Hertzian approximation of the contacting surfaces to be second-order accurate.

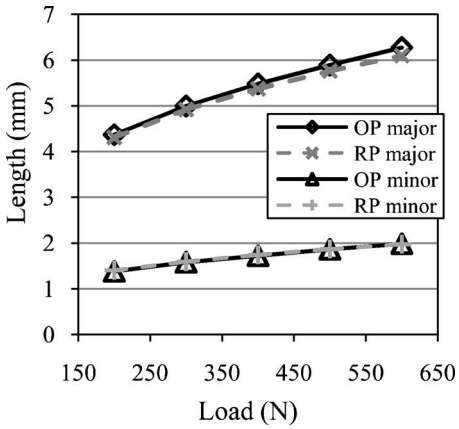

Fig. 8 Results of one validation test: comparison of contact patch major and minor axes. Legend: "OP" is the original pair, and "RP" is its replacing pair. Results for minor axes overlie one another.

Two simulations have been here described to potentially set a standard for verification of codes that employ the Hertzian geometric propositions. Differences between the quasi-ellipses and the tangent ellipses are shown to arise because the separation distance, as represented by the tangent ellipse equation, is second-order accurate, not exact. By reducing the normal offset (δ) of one surface into the other, the Hertzian formulation was shown to become increasingly accurate in the sense that the quasi-ellipse approached an ellipse. This exercise nicely illustrates the small-displacement limitation of Hertzian theory.

The second part of this article developed simple equations to compute the dimensions of a spheroid-plane contact pair that can substitute for a complicated contact pair and exhibit the same contact behavior, which is appealing for reducing costs in wear testing. Sims A and B demonstrated that the original and replacing pairs have separation distances, contact ellipses, and contact pressures that are equal to second-order accuracy. A simple laboratory experiment validated the substitution theory in a convex contact pair. Johnson showed [2] that the coefficients A and B describing the separation of contacting surfaces (Eq. (2)) determine in part the contact ellipse. The present article has thoroughly demonstrated a consequence of that observation via a validated analysis giving explicit equations for a simple surrogate contact pair.

5 Summary and Conclusions

The established notion that two geometrically dissimilar contact pairs may exhibit the same contact mechanics can be exploited to advantage. An original, complicated Hertzian contact pair may be substituted by a simpler, surrogate contact pair that will replicate the original pair's contact mechanics to second-order accuracy.

The surrogate contact pair can take the form of a spheroid-plane pair or of a spheroid-cylinder pair. The geometric simplicity of these replacing pairs may make them useful surrogates for expensive or complicated components in contact testing, such as ranking of materials for their wear resistance.

The geometric propositions of Hertzian contact theory are elucidated by using direct tensor notation. This approach clarifies the simplifying representation of an arbitrary surface by its two principal curvatures. Further, the approach shows that the Hertzian formulation of surface separation as a quadratic form results from a second-order Taylor series approximation. The coefficients in the quadratic form are the eigenvalues of the combined Hessian tensors of the two contacting surfaces.

Acknowledgment

This work was supported by Award No. R21AR056374 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS).

Nomenclature

Vectors, Tensors, and Matrices

Vector and tensor quantities are in bold; in matrix form they are enclosed in brackets or braces. For example, a tensor $\mathbf{T}$ in matrix form is denoted as $[\mathbf{T}]$. Likewise, a vector $\mathbf{v}$ in array form is denoted as $\{\mathbf{v}\}$.

- $\mathbf{A}$ = Hessian tensors combined under a common laboratory basis
- $[\mathbf{D}]$ = diagonalized component matrix of $\mathbf{A}$
- $[\mathbf{Q}]$ = rotation matrix formed from components of eigenvectors of $\mathbf{z}''$
- $[\mathbf{Z}]$ = diagonal matrix of principal curvatures of a surface
- $\mathbf{a}_i, \mathbf{b}_i$ = ith base vector of a basis aligned with principal curvatures of Surfaces K and P
- $\mathbf{c}_i$ = ith base vector of a basis aligned with the principal axes of the tangent ellipse
- $\mathbf{e}_i$ = ith base vector of a fixed orthonormal laboratory reference basis
- $\hat{\mathbf{n}}$ = unit vector normal to a surface
- $\mathbf{p}$ = the initial point of contact on a surface
- $\mathbf{r}$ = the part of $(\mathbf{x}\text{-}\mathbf{p})$ that lies in the mutual tangent plane
- $\mathbf{x}$ = vector to a point on a surface
- $\mathbf{y}$ = coordinate vector of basis aligned with surface principal curvatures
- $\mathbf{z}'$ = gradient of z
- $\mathbf{z}''$ = double gradient of z; tensor of surface curvatures
- $\boldsymbol{\kappa}, \boldsymbol{\rho}$ = Hessian tensors of Surfaces K and P
- $\mathbf{0}$ = the zero vector

Scalars

- A, B = coefficients in ellipse equation
- K, P = labels for two contacting surfaces
- R^3 = real, three-dimensional space
- a, b = major and minor semiaxes of the contact ellipse
- c, d = major and minor semiaxes of the tangent ellipse
- e = eccentricity of an ellipse
- h = distance between opposing points on contacting surfaces
- k, p = opposing points on contacting surfaces
- r_1, r_2 = principal radii of a prolate spheroid
- z = distance of a surface point from the tangent plane; alternately, distance between opposing points on two surfaces
- α = angle between planes of minimum absolute principal curvature, measured relative to the surface initially expressed in the $\mathbf{a}$-basis
- β = orientation angle of the tangent ellipse relative to the laboratory basis
- Δ = function of principal curvatures and their relative orientation
- Σ = sum of principal curvatures of both surfaces
- δ = a displacement along the common normal axis
- κ_1, κ_2 = principal curvatures at the contact point on the surface aligned with the $\mathbf{a}$-basis
- λ_1, λ_2 = eigenvalues of $[\mathbf{A}]$
- η_1, η_2 = principal curvatures at the contact point on a lone surface
- ρ_1, ρ_2 = principal curvatures at the contact point on the surface aligned with the $\mathbf{b}$-basis

Special Functions

- $\mathbf{K}, \mathbf{E}$ = complete elliptic integrals of first and second kinds of modulus e

Subscripts

- K, P = surfaces K and P
- O, R = original pair and replacing pair

References

[1] Hertz, H., 1882, "Uber Die Beruhrung Fester Elastischer Korper (on the Contact of Elastic Solids)," J. Reine Angew. Math., **92**, pp. 156–171.

[2] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge.

[3] Lur'e, A. I., 1964, *Three-Dimensional Problems in the Theory of Elasticity*, Wiley, New York.

[4] Antoine, J. F., Visa, C., Sauvey, C., and Abba, G., 2006, "Approximate Analytical Model for Hertzian Elliptical Contact Problems," ASME J. Tribol., **128**(3), pp. 660–664.

[5] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York.

[6] Frankel, T., 1997, *The Geometry of Physics: An Introduction*, Cambridge University Press, Cambridge.