ORIGINAL

Vlado A. Lubarda · Yujia Liu

Areal moments of inertia revisited: on the distinction between the principal directions

Abstract Three commonly used methods to determine the principal moments of inertia of a plane area and their directions are based on: (i) the stationarity condition for the axial moment of inertia, (ii) the eigenvalue analysis, and (iii) Mohr's circle. In this paper we provide two new derivations, which are based on: (a) the matrix diagonalization and the invariant tensor properties, and (b) the conjugacy property of the moment of inertia vectors. A new general expression is derived which specifies the principal directions of inertia, as well as the directions of the maximum and minimum product of inertia. A comparative study of the five presented approaches is given, which is of interest from both conceptual and methodological points of view. The connection between the deviatoric part of the moment of inertia tensor and Land's circle of inertia is also given. The presented analysis applies to any two-by-two symmetric second order tensor.

Keywords Diagonalization · Eigenvalue analysis · Inertia vectors · Land's circle · Moment of inertia · Mohr's circle · Principal directions

1 Introduction

The determination of the principal moments of inertia of a plane area and their principal directions is of great importance in the mechanics of rigid and deformable bodies, as evidenced by its detailed coverage in every undergraduate textbook on these subjects. The three commonly used methods to accomplish this are based on: (i) the condition of stationary axial moment of inertia about an arbitrary axis, (ii) the eigenvalue analysis of the moment of inertia tensor, and (iii) Mohr's circle. The objective of this paper is to present two additional derivations, which have not been reported in the literature before. The first is based on the matrix diagonalization of the moment of inertia tensor and its invariant tensor properties, and the second on the conjugacy property of the moment of inertia vectors. These methods also apply to an analogous determination of the principal stresses and their directions in two dimensional continuum mechanics problems (or any other two-by-two symmetric second order tensor), with the main difference that the moment of inertia tensor is a positive definite tensor, while the stress tensor is not necessarily so.

In Sect. 2 we diagonalize the moment of inertia tensor and use its invariant tensor properties to derive the expressions for the principal moments of inertia, and the expressions for the tangens of the angle that specifies each principal direction separately. The well-known derivation based on the eigenvalue analysis is included for comparison. In Sect. 3 we introduce the conjugacy property of the moment of inertia vectors, previously not explored in the literature in the context of moments of inertia, but well-known in the context of stress

V. A. Lubarda (✉) · Y. Liu
Department of Mechanical and Aerospace Engineering, University of California, San Diego,
La Jolla, CA 92093-0411, USA
E-mail: vlubarda@ucsd.edu
Tel.: 858-534-3169
Fax: 858-534-5698

analysis (conjugacy of traction vectors), although not used there to determine the principal stresses and their directions either. We derive the general expression which yields the principal directions of inertia, as well as the directions of the maximum and minimum product of inertia. The geometric interpretation of all analytical results is given based on Mohr's circle of inertia. The moment of inertia tensor is decomposed into its isotropic and deviatoric part in Sect. 4. Since the isotropic part does not have preferred directions, the moment of inertia tensor and its deviatoric part share the same principal directions. As a result, a simpler construction of Mohr's circle for the deviatoric part can be used to determine the principal directions. The connection to Land's circle of inertia is then given. The results are summarized and discussed in the concluding Sect. 5.

2 Diagonalization of the moment of inertia tensor

Consider a fixed coordinate system (x, y) and a coordinate system (u, v) obtained from (x, y) by the counter-clockwise rotation by an angle φ. The two sets of coordinates are related by the matrix transformation

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{Q}^T \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}. \tag{1}$$

The matrix $\mathbf{Q}$ is the rotation matrix that carries the unit vectors $\mathbf{e}_x$ and $\mathbf{e}_y$, along the coordinate directions (x, y), into the unit vectors $\mathbf{e}_u = \mathbf{Q}\mathbf{e}_x$ and $\mathbf{e}_v = \mathbf{Q}\mathbf{e}_y$, along the coordinate directions (u, v). In (1), and in the sequel, the abbreviations $c = \cos\varphi$ and $s = \sin\varphi$ are used.

The tensor of the second moment of inertia of an area A is defined (e.g., [1, p. 417]) as the symmetric tensor whose rectangular components are

$$\mathbb{I}_{x_i x_j} = \mathbb{I}_0 \delta_{ij} - \int_A x_i x_j \, \mathrm{d}A, \quad (i = 1, 2), \tag{2}$$

where $x_1 = x$ and $x_2 = y$, $\mathbb{I}_0 = \int_A (x_1^2 + x_2^2) \, \mathrm{d}A$ is the polar moment of inertia, and δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$, and zero otherwise). The matrix associated with (2) is

$$\mathbb{II} = \begin{bmatrix} \mathbb{I}_{xx} & \mathbb{I}_{xy} \\ \mathbb{I}_{xy} & \mathbb{I}_{yy} \end{bmatrix}, \tag{3}$$

with the elements

$$\mathbb{I}_{xx} = \int_A y^2 \, \mathrm{d}A, \quad \mathbb{I}_{yy} = \int_A x^2 \, \mathrm{d}A, \quad \mathbb{I}_{xy} = -\int_A xy \, \mathrm{d}A. \tag{4}$$

It can be easily verified that, under the rotation of the coordinate system (1), the tensor components in (3) transform according to

$$\begin{bmatrix} \mathbb{I}_{uu} & \mathbb{I}_{uv} \\ \mathbb{I}_{uv} & \mathbb{I}_{vv} \end{bmatrix} = \mathbf{Q}^T \begin{bmatrix} \mathbb{I}_{xx} & \mathbb{I}_{xy} \\ \mathbb{I}_{xy} & \mathbb{I}_{yy} \end{bmatrix} \mathbf{Q}, \tag{5}$$

such that

$$\begin{aligned}
\mathbb{I}_{uu} &= \mathbf{e}_u^T \mathbb{II} \mathbf{e}_u = c^2 \mathbb{I}_{xx} + s^2 \mathbb{I}_{yy} + 2sc \mathbb{I}_{xy}, \\
\mathbb{I}_{vv} &= \mathbf{e}_v^T \mathbb{II} \mathbf{e}_v = s^2 \mathbb{I}_{xx} + c^2 \mathbb{I}_{yy} - 2sc \mathbb{I}_{xy}, \\
\mathbb{I}_{uv} &= \mathbf{e}_u^T \mathbb{II} \mathbf{e}_v = \mathbf{e}_v^T \mathbb{II} \mathbf{e}_u = -sc(\mathbb{I}_{xx} - \mathbb{I}_{yy}) + (c^2 - s^2)\mathbb{I}_{xy},
\end{aligned} \tag{6}$$

where $\mathbf{e}_u^T = [c \quad s]$ and $\mathbf{e}_v^T = [-s \quad c]$ are the transposes of the column vectors $\mathbf{e}_u$ and $\mathbf{e}_v$.

On the other hand, the elements of the matrix

$$\mathbf{I} = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix},$$

where $I_{xx} = \mathbb{I}_{xx}$, $I_{yy} = \mathbb{I}_{yy}$, and $I_{xy} = -\mathbb{I}_{xy}$ do not constitute the components of the second-order tensor, because, under the rotation of the coordinate system (1), they do not transform according to the required tensor

transformation rule $\mathbf{Q}^T \mathbf{I} \mathbf{Q}$, but according to $\mathbf{Q} \mathbf{I} \mathbf{Q}^T$. The component I_{xy} is commonly referred to as the product of inertia (sometimes also mixed or deviatoric moment of inertia).

The subsequent derivation is based entirely on the tensor definition of the moment of inertia, so that there is a complete analogy of the resulting formulas with those for the stress tensor. If I_{xy} is used, rather than $\mathbb{I}_{xy}$, the analogy holds if the correspondence $\sigma_{xx} \leftrightarrow I_{xx}$, $\sigma_{yy} \leftrightarrow I_{yy}$, and $\sigma_{xy} \leftrightarrow -I_{xy}$ is made (e.g., [2,3], although this is sometimes overlooked; [4, p. 436]).

2.1 Principal moments of inertia

Theorem *The maximum and minimum moments of inertia are the diagonal entries of the diagonalized form of the moment of inertia tensor.*

Proof The tensor $\mathbf{I}$ is observed as a diagonal matrix, with positive[1] diagonal entries $\mathbb{I}_1$ and $\mathbb{I}_2$, in the coordinate system $(1, 2)$ obtained from the (x, y) coordinate system by the rotation φ, such that

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix}^T \begin{bmatrix} \mathbb{I}_{xx} & \mathbb{I}_{xy} \\ \mathbb{I}_{xy} & \mathbb{I}_{yy} \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \mathbb{I}_1 & 0 \\ 0 & \mathbb{I}_2 \end{bmatrix}. \tag{7}$$

Since the trace and the determinant of the matrix remain preserved under orthogonal transformations, we can write

$$\mathbb{I}_1 + \mathbb{I}_2 = \mathbb{I}_{xx} + \mathbb{I}_{yy}, \tag{8}$$

$$\mathbb{I}_1 \mathbb{I}_2 = \mathbb{I}_{xx} \mathbb{I}_{yy} - \mathbb{I}_{xy}^2. \tag{9}$$

Eliminating $\mathbb{I}_2$ from these two equations one obtains

$$(\mathbb{I}_1 - \mathbb{I}_{xx})(\mathbb{I}_1 - \mathbb{I}_{yy}) = \mathbb{I}_{xy}^2 \geq 0, \tag{10}$$

which means that the principal direction $\mathbb{I}_1$ is either greater or smaller than both $\mathbb{I}_{xx}$ and $\mathbb{I}_{yy}$. Thus, one of the diagonal entries on the right-hand side of (7) is the maximum moment of inertia and the other is the minimum. This completes the proof. We label $\mathbb{I}_1$ to be the maximum, and $\mathbb{I}_2$ to be the minimum.

The two invariants, the trace and the determinant, appearing in (8) and (9), are known as the principal invariants of $\mathbf{I}$. An invariant (dependent on the previous two) is also the Frobenious norm of $\mathbf{I}$, defined as the sum of the squares of its elements, i.e.,

$$\mathbb{I}_1^2 + \mathbb{I}_2^2 = \mathbb{I}_{xx}^2 + \mathbb{I}_{yy}^2 + 2\mathbb{I}_{xy}^2. \tag{11}$$

By combining (11) with (9), *i.e.*, by subtracting $2\mathbb{I}_1 \mathbb{I}_2$ from $\mathbb{I}_1^2 + \mathbb{I}_2^2$, there follows

$$(\mathbb{I}_1 - \mathbb{I}_2)^2 = (\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2. \tag{12}$$

Since we labeled the principal directions so that $\mathbb{I}_1 > \mathbb{I}_2$, (12) gives

$$\mathbb{I}_1 - \mathbb{I}_2 = \left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2 \right]^{1/2}. \tag{13}$$

The principal moments of inertia follow by adding and subtracting (8) and (13),

$$\mathbb{I}_{1,2} = \frac{1}{2} (\mathbb{I}_{xx} + \mathbb{I}_{yy}) \pm \frac{1}{2} \left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2 \right]^{1/2}. \tag{14}$$

$\square$

---

[1] Axial moments of inertia are positive quantities by their definition and their geometrical meaning. Accordingly, the tensor $\mathbf{I}$ is a positive definite tensor, i.e., for any vector $\mathbf{r} = r\mathbf{n}$, where $r = |\mathbf{r}|$ is the length of $\mathbf{r}$, one has $r^2 \mathbf{n}^T \mathbf{I} \mathbf{n} = r^2 \mathbb{I}_n > 0$. By taking $\mathbf{r}^T = [1 \ \ 1]$ and $\mathbf{r}^T = [1 \ \ -1]$, and combining the results, it follows that $|\mathbb{I}_{xy}| < (\mathbb{I}_{xx} + \mathbb{I}_{yy})/2$, so that the largest element of the matrix $\mathbf{I}$ is on its diagonal (a symmetric positive definite matrix has a "weighty" diagonal; [5, p. 140]).

2.2 Principal directions

The angles that define the principal directions follow from the expanded form of (7), which is

$$c^2 \mathbb{I}_{xx} + s^2 \mathbb{I}_{yy} + 2sc\mathbb{I}_{xy} = \mathbb{I}_1,$$
$$s^2 \mathbb{I}_{xx} + c^2 \mathbb{I}_{yy} - 2sc\mathbb{I}_{xy} = \mathbb{I}_2, \tag{15}$$
$$-sc(\mathbb{I}_{xx} - \mathbb{I}_{yy}) + (c^2 - s^2)\mathbb{I}_{xy} = 0.$$

By using the transformations

$$c^2 = \frac{1}{1+t^2}, \quad s^2 = \frac{t^2}{1+t^2}, \quad t = \tan\varphi, \tag{16}$$

the first of (15) can be rewritten as a quadratic equation for t,

$$(\mathbb{I}_1 - \mathbb{I}_{yy})t^2 - 2t\mathbb{I}_{xy} + \mathbb{I}_1 - \mathbb{I}_{xx} = 0. \tag{17}$$

This has the unique solution

$$t_1 = \tan\varphi_1 = \frac{\mathbb{I}_{xy}}{\mathbb{I}_1 - \mathbb{I}_{yy}}, \tag{18}$$

because the discriminant of the quadratic equation (17) is equal to zero; see (10).

Likewise, the second of (15) can be rewritten as

$$(\mathbb{I}_2 - \mathbb{I}_{xx})t^2 + 2t\mathbb{I}_{xy} + \mathbb{I}_2 - \mathbb{I}_{yy} = 0, \tag{19}$$

which has the unique solution

$$t_2 = \tan\varphi_2 = \frac{\mathbb{I}_{xy}}{\mathbb{I}_{xx} - \mathbb{I}_2}, \tag{20}$$

because

$$\mathbb{I}_{xy}^2 - (\mathbb{I}_2 - \mathbb{I}_{xx})(\mathbb{I}_2 - \mathbb{I}_{yy}) = 0. \tag{21}$$

Finally, the third of (15), requiring that the product of inertia for the principal directions vanishes, gives

$$\tan 2\varphi = \frac{2\mathbb{I}_{xy}}{\mathbb{I}_{xx} - \mathbb{I}_{yy}}, \tag{22}$$

which is satisfied by both φ_1 and φ_2. The expressions (18) and (20) are more appealing than (22), because they unambiguously specify the two principal directions separately, while (22) does not make distinction between the two directions.

2.3 Octant theorem

Expression (22) is a common outcome of the analysis based on the stationarity condition $d\mathbb{I}_{uu}/d\varphi = 0$. Since $d\mathbb{I}_{uu}/d\varphi = 2\mathbb{I}_{uv}$ (which means that, if the area has its maximum or minimum moment of inertia about an axis, the product of inertia for that axis and the axis orthogonal to it vanishes), there follows (22). The analysis of the sign of the corresponding second derivative $d^2\mathbb{I}_{uu}/d\varphi^2$ is more lengthy and tedious, but can be summarized as the following (octant) theorem.[2]

---

[2] The term octant is here used in the sense of one eight of the 360° angle.

Theorem *The principal direction of the maximum moment of inertia passes through*:

$$1^{\text{st}} \text{ and } 5^{\text{th}} \text{ octant, } \quad \text{if } \mathbb{I}_{xx} > \mathbb{I}_{yy} \text{ and } \mathbb{I}_{xy} > 0,$$
$$2^{\text{nd}} \text{ and } 6^{\text{th}} \text{ octant, } \quad \text{if } \mathbb{I}_{xx} < \mathbb{I}_{yy} \text{ and } \mathbb{I}_{xy} > 0,$$
$$3^{\text{rd}} \text{ and } 7^{\text{th}} \text{ octant, } \quad \text{if } \mathbb{I}_{xx} < \mathbb{I}_{yy} \text{ and } \mathbb{I}_{xy} < 0,$$
$$4^{\text{th}} \text{ and } 8^{\text{th}} \text{ octant, } \quad \text{if } \mathbb{I}_{xx} > \mathbb{I}_{yy} \text{ and } \mathbb{I}_{xy} < 0.$$

To make the distinction between the principal directions, it actually suffices to remember a weaker (quadrant rather than octant) form of the theorem: The direction of the maximum moment of inertia passes through the 1$^{\text{st}}$ and 3$^{\text{rd}}$ quadrant if $\mathbb{I}_{xy} > 0$, and through the 2$^{\text{nd}}$ and 4$^{\text{th}}$ quadrant if $\mathbb{I}_{xy} < 0$. Another weaker form of the theorem is also helpful: If $\mathbb{I}_{xx} > \mathbb{I}_{yy}$, the direction of the maximum moment of inertia is closer to the x-axis than the y-axis (and vice versa) [6, p. 555].

Alternatively, to identify the angles φ_1 and φ_2 from (22), we can check which of the corresponding values of $\mathbb{I}_{uu}$ is greater and which one is smaller [7, p. 429]. Or, we can use (22) and the trigonometric relations between the sine, cosine, and tangens, to obtain

$$\sin 2\varphi_1 = \frac{2\mathbb{I}_{xy}}{\left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2\right]^{1/2}}, \quad \cos 2\varphi_1 = \frac{\mathbb{I}_{xx} - \mathbb{I}_{yy}}{\left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2\right]^{1/2}}. \tag{23}$$

The angle φ_1 can then be determined as the angle which satisfies both expressions in (23) [8, p. 550]. To circumvent this double calculation, the use of the octant theorem, or the analytical determination of the angles φ_1 and φ_2 based on (18) and (20), is clearly more convenient.

2.4 Maximum product of inertia

Theorem *Two orthogonal directions for which an area has equal axial moments of inertia are the directions for which that area has the maximum or minimum product of inertia.*

Proof The moment of inertia tensor is observed in the coordinate systems (u, v) and $(1, 2)$ as

$$\begin{bmatrix} \mathbb{I}_{uu} & \mathbb{I}_{uv} \\ \mathbb{I}_{uv} & \mathbb{I}_{vv} \end{bmatrix}, \quad \begin{bmatrix} \mathbb{I}_1 & 0 \\ 0 & \mathbb{I}_2 \end{bmatrix}. \tag{24}$$

The determinant $\mathbb{I}_{uu}\mathbb{I}_{vv} - \mathbb{I}_{uv}^2$ remains unchanged under the rotation of the coordinate system. Thus, if $\mathbb{I}_{uv}^2$ is at its maximum in some coordinate system, the product $\mathbb{I}_{uu}\mathbb{I}_{vv}$ must be at its maximum in the same coordinate system, too. Since the trace remains invariant under the rotation of the coordinate system, $\mathbb{I}_{uu} + \mathbb{I}_{vv} = \mathbb{I}_0$, we search for the maximum of the function $g = \mathbb{I}_{uu}\mathbb{I}_{vv} = \mathbb{I}_{uu}(\mathbb{I}_0 - \mathbb{I}_{uu})$. The extremum condition for this is

$$\frac{\mathrm{d}g}{\mathrm{d}\mathbb{I}_{uu}} = 0 \quad \Rightarrow \quad \mathbb{I}_0 - 2\mathbb{I}_{uu} = 0, \tag{25}$$

which gives $\mathbb{I}_{uu} = \mathbb{I}_0/2$. Thus, in the coordinate system in which $\mathbb{I}_{uu}\mathbb{I}_{vv}$, and thus $\mathbb{I}_{uv}^2$, is at its maximum, the two diagonal entries of the moment of inertia tensor are equal to each other ($\mathbb{I}_{uu} = \mathbb{I}_{vv} = \mathbb{I}_0/2$). This completes the proof.[3] $\qquad \square$

Denoting the coordinate system for which the product of inertia is maximum by $(\bar{u}, \bar{v})$, we have

$$\mathbb{I}_{\bar{u}\bar{u}} = \mathbb{I}_{\bar{v}\bar{v}} = \frac{1}{2}\mathbb{I}_0 = \frac{1}{2}(\mathbb{I}_{xx} + \mathbb{I}_{vv}). \tag{26}$$

---

[3] An alternative simple proof is based on the identity

$$\frac{\mathrm{d}\mathbb{I}_{uv}}{\mathrm{d}\varphi} = -2\left[\mathbb{I}_{uu} - \frac{1}{2}(\mathbb{I}_{xx} + \mathbb{I}_{yy})\right],$$

so that the stationarity condition $\mathrm{d}\mathbb{I}_{uv}/\mathrm{d}\varphi = 0$ gives $\mathbb{I}_{uu} = (\mathbb{I}_{xx} + \mathbb{I}_{yy})/2$.

The corresponding, maximum value of $\mathbb{I}_{uv}^2$ follows from the invariance of the determinant of the moment of inertia tensor, which gives

$$\mathbb{I}_{\bar{u}\bar{v}}^2 = \mathbb{I}_{\bar{u}\bar{u}}^2 - \mathbb{I}_1 \mathbb{I}_2 = \mathbb{I}^2, \quad \mathbb{I} = \frac{1}{2}\left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2\right]^{1/2}, \tag{27}$$

because $\mathbb{I}_1 \mathbb{I}_2 = \mathbb{I}_0^2/4 - \mathbb{I}^2$, by (14).

To find the directions $(\bar{u}, \bar{v})$ for which $\mathbb{I}_{\bar{u}\bar{v}} = \mathbb{I}$, we equate the spectral representations of $\mathbf{II}$ in the $(\bar{u}, \bar{v})$ and $(1, 2)$ coordinate systems. Upon the cancelation of the isotropic parts on the two sides of the equation,[4] it follows

$$\mathbf{e}_{\bar{u}}\mathbf{e}_{\bar{v}}^T + \mathbf{e}_{\bar{v}}\mathbf{e}_{\bar{u}}^T = \mathbf{e}_1\mathbf{e}_1^T - \mathbf{e}_2\mathbf{e}_2^T. \tag{28}$$

This implies that $\mathbf{e}_{\bar{u}} = (\mathbf{e}_1 - \mathbf{e}_2)/2$ and $\mathbf{e}_{\bar{v}} = (\mathbf{e}_1 + \mathbf{e}_2)/2$, i.e., the $(\bar{u}, \bar{v})$ axes are obtained from the principal axes $(1, 2)$ by the $45°$ clockwise rotation.

An alternative derivation based on the stationarity condition of $\mathbb{I}_{uv}$ is well-known; e.g. [9, p. 536]. To find the coordinate axes (u, v) for which the product of inertia $\mathbb{I}_{uv}$ has its maximum or minimum value, we require that

$$\frac{d\mathbb{I}_{uv}}{d\varphi} = -(\mathbb{I}_{xx} - \mathbb{I}_{yy})(c^2 - s^2) - 4sc\mathbb{I}_{xy} = 0, \tag{29}$$

which gives

$$\tan 2\bar{\varphi} = -\frac{\mathbb{I}_{xx} - \mathbb{I}_{yy}}{2\mathbb{I}_{xy}}. \tag{30}$$

By comparing (22) and (30), $\tan 2\varphi \tan 2\bar{\varphi} = -1$, so that φ_1 and $\bar{\varphi}_1$ and $45°$ apart.

2.5 Eigenvalue analysis

For the sake of comparison, we summarize in this section the well-known derivation of the principal moments of inertia and their directions, based on the eigenvalue analysis of the moment of inertia tensor $\mathbf{II}$. The moment of inertia about an axis along the unit vector $\mathbf{n}$ is $\mathbb{I}_n = \mathbf{n}^T \mathbf{II} \mathbf{n}$. The principal direction of inertia is the direction that maximizes or minimizes $\mathbb{I}_n$, over all directions $\mathbf{n}$ subject to the constraint $\mathbf{n}^T \mathbf{n} = 1$. Thus, introducing the Lagrange multiplier λ and the function $\Phi = \mathbb{I}_n - \lambda \mathbf{n}^T \mathbf{n}$, the condition $\partial\Phi/\partial\mathbf{n} = \mathbf{0}$ for the unconstrained extremum of Φ gives

$$\mathbf{II}\mathbf{n} = \lambda \mathbf{n}. \tag{31}$$

Evidently, if (31) is premultiplied by $\mathbf{n}^T$, it follows that $\lambda = \mathbf{n}^T \mathbf{II} \mathbf{n}$ is the maximum or minimum moment of inertia. The corresponding $\mathbf{n}$, satisfying (31), is the principal direction. The scalar λ and the vector $\mathbf{n}$ are referred to as the eigenvalue and eigenvector of $\mathbf{II}$.

By taking $\mathbf{n}^T = [c \quad s]$, which automatically satisfies the constraint condition $\mathbf{n}^T \mathbf{n} = 1$, the expanded form of the eigenvalue problem (31) is

$$\begin{aligned}(\mathbb{I}_{xx} - \lambda)c + \mathbb{I}_{xy}s &= 0, \\ \mathbb{I}_{xy}c + (\mathbb{I}_{yy} - \lambda)s &= 0.\end{aligned} \tag{32}$$

In order that this homogeneous system of linear algebraic equations for c and s has a nontrivial solution, the determinant of the system must vanish. Upon expansion, this gives a quadratic (secular) equation for the

---

[4] The isotropic parts are

$$\frac{1}{2}(\mathbb{I}_{xx} + \mathbb{I}_{yy})\left(\mathbf{e}_{\bar{u}}\mathbf{e}_{\bar{u}}^T + \mathbf{e}_{\bar{v}}\mathbf{e}_{\bar{v}}^T\right) = \frac{1}{2}(\mathbb{I}_1 + \mathbb{I}_2)\left(\mathbf{e}_1\mathbf{e}_1^T + \mathbf{e}_2\mathbf{e}_2^T\right).$$

eigenvalues, whose solutions are the principal moments of inertia, given by (14). The corresponding principal directions follow from (32). For $\lambda = \mathbb{I}_1$, (32) yields

$$\tan \varphi_1 = \frac{\mathbb{I}_1 - \mathbb{I}_{xx}}{\mathbb{I}_{xy}} = \frac{\mathbb{I}_{xy}}{\mathbb{I}_1 - \mathbb{I}_{yy}}, \tag{33}$$

while for $\lambda = \mathbb{I}_2$,

$$\tan \varphi_2 = -\frac{\mathbb{I}_{xx} - \mathbb{I}_2}{\mathbb{I}_{xy}} = -\frac{\mathbb{I}_{xy}}{\mathbb{I}_{yy} - \mathbb{I}_2}, \tag{34}$$

in agreement with the results from Sect. 2.2, and in analogy with the eigenvalue analysis of the plane state of stress [10,11].

To make the connection with the analysis based on the diagonalization of the moment of inertia tensor, we recall from linear algebra [12,13] that, when diagonalizing a matrix $\mathbf{II}$, the column vectors of the modal (diagonalizing) matrix are eigenvectors of $\mathbf{II}$, and the diagonal elements of the diagonalized form are corresponding eigenvalues of $\mathbf{II}$. Thus, the modal matrix in (7) is

$$\mathbf{Q} = \begin{bmatrix} c_1 & c_2 \\ s_1 & s_2 \end{bmatrix}, \tag{35}$$

where $c_1 = \cos \varphi_1$, $s_1 = \sin \varphi_1$, and $c_2 = -s_1$, $s_2 = c_1$ specify the eigenvectors of $\mathbf{II}$.

3 Moment of inertia vectors and their conjugacy property

The moment of inertia vector for the direction u is defined as

$$\mathbf{p}_u = \mathbf{II} \mathbf{e}_u. \tag{36}$$

Its representation in the (u, v) coordinate system is

$$\mathbf{p}_u^T = [\mathbb{I}_{uu} \quad \mathbb{I}_{uv}]. \tag{37}$$

The conjugacy property of the inertia vectors $\mathbf{p}_n$ and $\mathbf{p}_m$ is

$$\mathbf{p}_n^T \mathbf{m} = \mathbf{p}_m^T \mathbf{n}, \tag{38}$$

for any unit direction vectors $\mathbf{m}$ and $\mathbf{n}$, which holds because of the symmetry of $\mathbf{II}$. This conjugacy property, not explicitly stated or explored in the literature in the context of the moments of inertia, is well-known in the stress analysis, where it applies to traction vectors over the surface elements with the unit normals $\mathbf{m}$ and $\mathbf{n}$; [14, p. 79].

If (38) is applied to unit vectors $\mathbf{e}_u$ and $\mathbf{e}_y$, there follows

$$\mathbf{e}_u^T \mathbf{II} \mathbf{e}_y = \mathbf{e}_y^T \mathbf{II} \mathbf{e}_u. \tag{39}$$

The unit vector $\mathbf{e}_u$, expressed in the (x, y) coordinate system, is $\mathbf{e}_u = c\mathbf{e}_x + s\mathbf{e}_y$, while the unit vector $\mathbf{e}_y$, expressed in the (u, v) coordinate system, is $\mathbf{e}_y = s\mathbf{e}_u + c\mathbf{e}_v$. Thus, by evaluating the left-hand side of (39) in the (u, v) coordinate system, and the right-hand side in the (x, y) coordinate system, carrying in mind that the representation of $\mathbf{p}_y = \mathbf{II} \mathbf{e}_y$ in the (x, y) coordinate system is $\mathbf{p}_y^T = [\mathbb{I}_{yx} \quad \mathbb{I}_{yy}]$, there follows

$$s\mathbb{I}_{uu} + c\mathbb{I}_{uv} = c\mathbb{I}_{yx} + s\mathbb{I}_{yy}. \tag{40}$$

This gives a remarkable expression, not reported in the literature before,

$$\tan \varphi = \frac{\mathbb{I}_{xy} - \mathbb{I}_{uv}}{\mathbb{I}_{uu} - \mathbb{I}_{yy}}. \tag{41}$$

Its validity can be verified directly by using the expressions for $\mathbb{I}_{uu}$ and $\mathbb{I}_{uv}$ from the transformation formulas (6), while its geometric interpretation by means of Mohr's circle will be given in Sect. 3.3.

Similarly, if (38) is applied to unit vectors $\mathbf{e}_u$ and $\mathbf{e}_x$, there follows

$$\mathbf{e}_u^T \mathbf{II} \mathbf{e}_x = \mathbf{e}_x^T \mathbf{II} \mathbf{e}_u. \tag{42}$$

Since $\mathbf{e}_x = c\mathbf{e}_u - s\mathbf{e}_v$, and since the representation of $\mathbf{p}_x = \mathbf{II}\mathbf{e}_x$ in the (x, y) coordinate system is $\mathbf{p}_x^T = [\mathbb{I}_{xx} \quad \mathbb{I}_{xy}]$, by evaluating the left-hand side of (42) in the (u, v) coordinate system, and the right-hand side in the (x, y) coordinate system, there follows

$$c\mathbb{I}_{uu} - s\mathbb{I}_{uv} = c\mathbb{I}_{xx} + s\mathbb{I}_{xy}. \tag{43}$$

This gives an alternative or dual expression to (41), which is

$$\tan \varphi = \frac{\mathbb{I}_{uu} - \mathbb{I}_{xx}}{\mathbb{I}_{uv} + \mathbb{I}_{xy}}. \tag{44}$$

3.1 Principal directions

The expression (41) delivers the angle φ_1 by taking $\mathbb{I}_{uv} = 0$ and $\mathbb{I}_{uu} = \mathbb{I}_1$, and the angle φ_2 by taking $\mathbb{I}_{uv} = 0$ and $\mathbb{I}_{uu} = \mathbb{I}_2$. The results are

$$\tan \varphi_1 = \frac{\mathbb{I}_{xy}}{\mathbb{I}_1 - \mathbb{I}_{yy}}, \quad \tan \varphi_2 = -\frac{\mathbb{I}_{xy}}{\mathbb{I}_{yy} - \mathbb{I}_2}, \tag{45}$$

in agreement with (18). The expression (44) delivers the angle φ_1 by taking $\mathbb{I}_{uv} = 0$ and $\mathbb{I}_{uu} = \mathbb{I}_1$, and the angle φ_2 by taking $\mathbb{I}_{uv} = 0$ and $\mathbb{I}_{uu} = \mathbb{I}_2$, such that

$$\tan \varphi_1 = \frac{\mathbb{I}_1 - \mathbb{I}_{xx}}{\mathbb{I}_{xy}}, \quad \tan \varphi_2 = -\frac{\mathbb{I}_{xx} - \mathbb{I}_2}{\mathbb{I}_{xy}}, \tag{46}$$

in agreement with (20).

Both (45) and (46) can also be obtained directly from the conjugacy property (38), by applying it to the direction pairs $(1, x)$ and $(2, x)$, respectively, *i.e.*, from $\mathbf{e}_1^T \mathbf{II} \mathbf{e}_x = \mathbf{e}_x^T \mathbf{II} \mathbf{e}_1$ and $\mathbf{e}_2^T \mathbf{II} \mathbf{e}_x = \mathbf{e}_x^T \mathbf{II} \mathbf{e}_2$ (the left-hand sides being evaluated in the $(1, 2)$ coordinate system, and the right-hand sides in the (x, y) coordinates system, or vice versa).

3.2 Principal moments of inertia

Once (45) or (46) are established, the corresponding principal moments of inertia follow from either $\mathbb{I}_1 = \mathbf{e}_1^T \mathbf{II} \mathbf{e}_1$ or $\mathbb{I}_2 = \mathbf{e}_2^T \mathbf{II} \mathbf{e}_2$. For example, since $\mathbf{e}_1^T = [c_1 \quad s_1]$, we have

$$\mathbb{I}_1 = c_1^2 \mathbb{I}_{xx} + s_1^2 \mathbb{I}_{yy} + 2s_1 c_1 \mathbb{I}_{xy}. \tag{47}$$

The substitution of

$$c_1^2 = \frac{1}{1 + t_1^2}, \quad s_1^2 = \frac{t_1^2}{1 + t_1^2}, \quad s_1 c_1 = \frac{t_1}{1 + t_1^2}, \tag{48}$$

in conjunction with $t_1 = (\mathbb{I}_1 - \mathbb{I}_{xx})/\mathbb{I}_{xy}$, yields

$$(\mathbb{I}_1 - \mathbb{I}_{xx})(\mathbb{I}_1 - \mathbb{I}_{yy}) - \mathbb{I}_{xy}^2 = 0. \tag{49}$$

This is a quadratic equation for $\mathbb{I}_1$, whose greater root is the maximum, and smaller root is the minimum moment of inertia, given by (14). The same results follow from the analysis of the expression $\mathbb{I}_2 = \mathbf{e}_2^T \mathbf{II} \mathbf{e}_2$ and $t_2 = (\mathbb{I}_2 - \mathbb{I}_{xx})/\mathbb{I}_{xy}$.
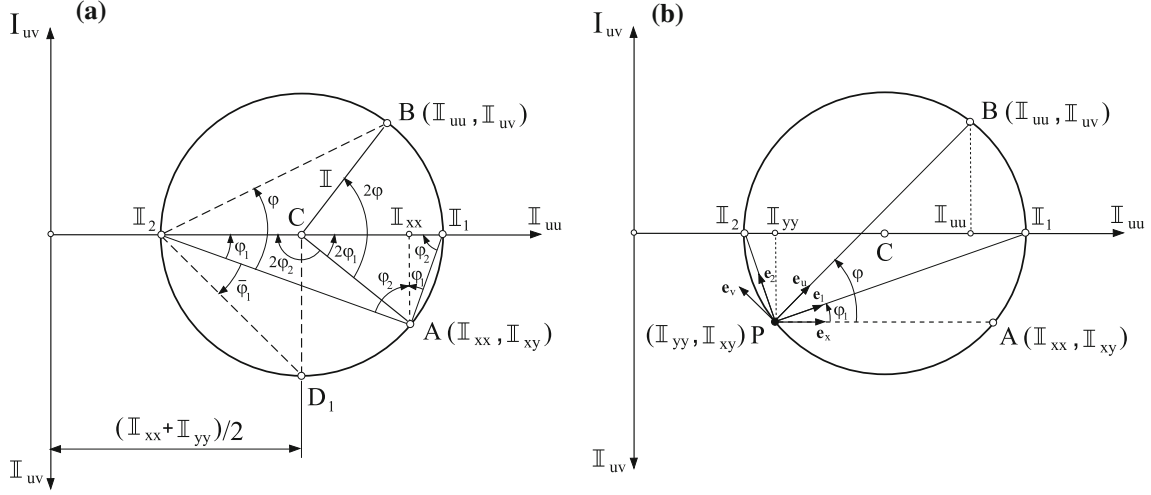
Fig. 1 a Mohr's circle for the moment of inertia tensor $\mathbb{II}$, in the case $\mathbb{I}_{xx} > \mathbb{I}_{yy}$ and $\mathbb{I}_{xy} > 0$. **b** The pole P of Mohr's circle used to identify the orientation of the principal directions of inertia in the physical space (relative to the horizontal direction x)

3.3 Mohr's circle interpretation

The derived results can be given a geometrical interpretation by means of Mohr's circle

$$\left[\mathbb{I}_{uu} - \frac{1}{2}(\mathbb{I}_{xx} + \mathbb{I}_{yy})\right]^2 + \mathbb{I}_{uv}^2 = \mathbb{I}^2, \quad \mathbb{I} = \frac{1}{2}\left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2\right]^{1/2}. \tag{50}$$

By geometric properties of the circle, and the triangles within the circle shown in Fig. 1a, it readily follows that

$$\tan\varphi_{1,2} = \frac{\mathbb{I}_{xy}}{\mathbb{I}_{xx} - \mathbb{I}_{2,1}} = \frac{\mathbb{I}_{1,2} - \mathbb{I}_{xx}}{\mathbb{I}_{xy}}, \tag{51}$$

confirming the results (45) and (46). In addition,

$$\tan 2\varphi_{1,2} = \frac{2\mathbb{I}_{xy}}{\mathbb{I}_{xx} - \mathbb{I}_{yy}}, \tag{52}$$

and

$$\cos 2\varphi_1 = \frac{\mathbb{I}_{xx} - \mathbb{I}_{yy}}{2\mathbb{I}}, \quad \sin 2\varphi_1 = \frac{\mathbb{I}_{xy}}{\mathbb{I}}. \tag{53}$$

Figure 1b illustrates the use of the pole P to identify the actual orientation of the principal directions $\mathbf{e}_1$ and $\mathbf{e}_2$ (in the physical space), and the moments of inertia corresponding to orthogonal directions along the unit vectors $\mathbf{e}_u$ and $\mathbf{e}_v$ [15, p. 227]. The unit vector $\mathbf{e}_u$ is at an angle φ with respect to the horizontal direction $\mathbf{e}_x$, and $\mathbf{e}_v$ is 90° counterclockwise from $\mathbf{e}_u$. The pole P is the point on Mohr's circle obtained from the point $A(\mathbb{I}_{xx}, \mathbb{I}_{xy})$ as its mirror image across the vertical diameter of the circle, so that $P(\mathbb{I}_{yy}, \mathbb{I}_{xy})$. The tangens of the angle φ, defining the point $B(\mathbb{I}_{uu}, \mathbb{I}_{uv})$ and the direction $\mathbf{e}_u$, is

$$\tan\varphi = \frac{\mathbb{I}_{xy} - \mathbb{I}_{uv}}{\mathbb{I}_{uu} - \mathbb{I}_{yy}}, \tag{54}$$

confirming the result (41). The expression (54) delivers both φ_1 and φ_2, by taking $\mathbb{I}_{uv} = 0$, and $\mathbb{I}_{uu}$ equal to $\mathbb{I}_1$ in the first case (for φ_1), and $\mathbb{I}_{uu}$ equal to $\mathbb{I}_2$ in the second case (for φ_2), so that

$$\tan\varphi_{1,2} = \frac{\mathbb{I}_{xy}}{\mathbb{I}_{1,2} - \mathbb{I}_{yy}}. \tag{55}$$

It also delivers the analytical expressions for the angles $\bar{\varphi}_1$ and $\bar{\varphi}_2$, corresponding to maximum and minimum values of $\mathbb{I}_{uv}$. These are obtained by substituting in (54) $\mathbb{I}_{uu} = (\mathbb{I}_{xx} + \mathbb{I}_{yy})/2$ and $\mathbb{I}_{uv} = \pm\mathbb{I}$. The results are

$$\tan\bar{\varphi}_{1,2} = 2\,\frac{\mathbb{I}_{xy} \mp \mathbb{I}}{\mathbb{I}_{xx} - \mathbb{I}_{yy}}, \tag{56}$$

with the connection $\bar{\varphi}_{1,2} = \varphi_1 \mp 45°$.

4 Deviatoric part of the moment of inertia tensor

It is convenient to rewrite the transformation formulas for the components of the moment of inertia tensor (6) by adding and subtracting the expressions for the axial moments of inertia. This gives

$$\begin{aligned}
\mathbb{I}_{uu} + \mathbb{I}_{vv} &= \mathbb{I}_{xx} + \mathbb{I}_{yy}, \\
\mathbb{I}_{uu} - \mathbb{I}_{vv} &= (c^2 - s^2)(\mathbb{I}_{xx} - \mathbb{I}_{yy}) + 4sc\mathbb{I}_{xy}, \\
\mathbb{I}_{uv} &= -sc(\mathbb{I}_{xx} - \mathbb{I}_{yy}) + (c^2 - s^2)\mathbb{I}_{xy}.
\end{aligned} \tag{57}$$

The last two of these expressions can be grouped into the complex form

$$\frac{\mathbb{I}_{uu} - \mathbb{I}_{vv}}{2} - i\mathbb{I}_{uv} = \left(\frac{\mathbb{I}_{xx} - \mathbb{I}_{yy}}{2} - i\mathbb{I}_{xy}\right) e^{2i\varphi}. \tag{58}$$

The difference $(\mathbb{I}_{xx} - \mathbb{I}_{yy})/2$ is the axial component of the deviatoric part $\mathbf{J}$ of the moment of inertia tensor $\mathbf{\mathbb{I}}$, whose components are defined by

$$\mathbb{J}_{ij} = \mathbb{I}_{ij} - \frac{1}{2}\mathbb{I}_0\,\delta_{ij}, \quad \mathbb{I}_0 = \mathbb{I}_{xx} + \mathbb{I}_{yy}, \tag{59}$$

i.e.,

$$\begin{bmatrix} \mathbb{J}_{xx} & \mathbb{J}_{xy} \\ \mathbb{J}_{xy} & \mathbb{J}_{yy} \end{bmatrix} = \begin{bmatrix} (\mathbb{I}_{xx} - \mathbb{I}_{yy})/2 & \mathbb{I}_{xy} \\ \mathbb{I}_{xy} & -(\mathbb{I}_{xx} - \mathbb{I}_{yy})/2 \end{bmatrix}. \tag{60}$$

Clearly,

$$\mathbb{J}_{xx} + \mathbb{J}_{yy} = 0, \quad \mathbb{J}_{xx}\mathbb{J}_{yy} - \mathbb{J}_{xy}^2 = -\mathbb{I}^2. \tag{61}$$

An alternative complex form to (58) is

$$\frac{\mathbb{I}_{uu} - \mathbb{I}_{vv}}{2} + i\mathbb{I}_{uv} = \left(\frac{\mathbb{I}_{xx} - \mathbb{I}_{yy}}{2} + i\mathbb{I}_{xy}\right) e^{-2i\varphi}. \tag{62}$$

By multiplying (58) and (62), there follows

$$\mathbb{J}_{uu}^2 + \mathbb{J}_{uv}^2 = \mathbb{J}^2, \quad \mathbb{J} = \frac{1}{2}\left[(\mathbb{I}_{xx} - \mathbb{I}_{yy})^2 + 4\mathbb{I}_{xy}^2\right]^{1/2}, \tag{63}$$

which is the equation of the circle of radius $\mathbb{J} = \mathbb{I}$ in the $(\mathbb{J}_{uu}, \mathbb{J}_{uv})$ plane (Fig. 2). Since the isotropic part $\mathbb{I}_0\,\delta_{ij}/2$ of the tensor $\mathbf{\mathbb{I}}$ does not have any preferred directions, the principal directions of the deviatoric part $\mathbf{J}$ coincide with those of $\mathbf{\mathbb{I}}$. Thus, from the triangles in Fig. 2,

$$\tan\varphi_{1,2} = \frac{\mathbb{J}_{xy}}{\mathbb{J}_{xx} \pm \mathbb{J}} = \frac{\pm\mathbb{J} - \mathbb{J}_{xx}}{\mathbb{J}_{xy}}, \tag{64}$$

where $\mathbb{J}_1 = \mathbb{J}$ is the positive eigenvalue of $\mathbf{J}$ (radius of the circle). In addition,

$$\tan 2\varphi_{1,2} = \frac{\mathbb{J}_{xy}}{\mathbb{J}_{xx}} = \frac{2\mathbb{I}_{xy}}{\mathbb{I}_{xx} - \mathbb{I}_{yy}}. \tag{65}$$
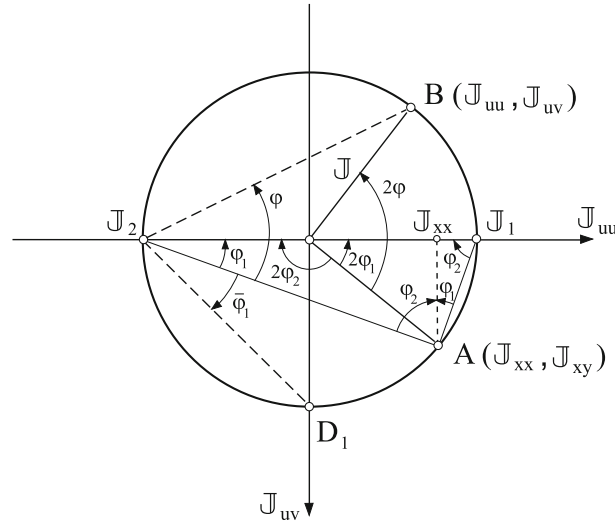
Fig. 2 Mohr's circle for the deviatoric part **J** of the moment of inertia tensor $\mathbb{II}$. The two tensors share the same principal directions (the angles φ_1 and φ_2), while their principal values are related by $\mathbb{I}_{1,2} = \frac{1}{2}\mathbb{I}_0 + \mathbb{J}_{1,2}$, with $\mathbb{J}_2 = -\mathbb{J}_1 = -\mathbb{J}$
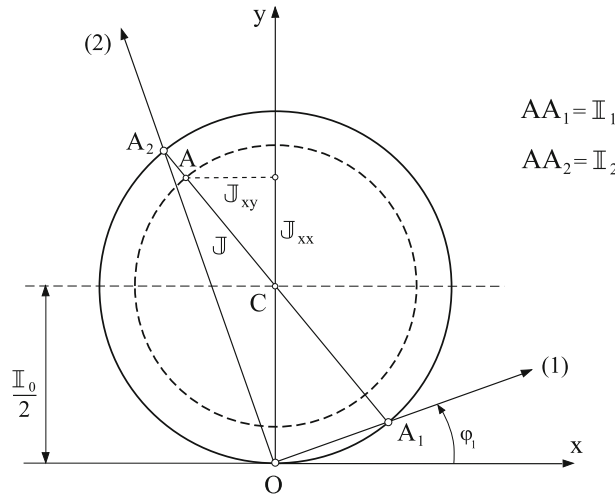

Fig. 3 Land's circle of the moment of inertia with its center at C and with the radius $\mathbb{I}_0/2$

The magnitude of the maximum product of inertia is $\mathbb{I}_{uv}^{\max} = \mathbb{J}_1$, and it is for the pair of orthogonal axes at $45°$ relative to the principal axes of inertia (point D_1 in Fig. 2, with $\bar{\varphi}_1 = \varphi_1 - 45°$).

The presented analysis of the deviatoric part of the moment of inertia can be related to Land's circle, which is another graphical method, albeit less frequently utilized than Mohr's circle [4,16]. This is shown in Fig. 3. The center of Land's circle is at the point C on the y-axis, and its diameter is $\mathbb{I}_0 = \mathbb{I}_{xx} + \mathbb{I}_{yy}$. The point A has the coordinates $(\mathbb{J}_{xx}, \mathbb{J}_{xy})$ relative to the center, as indicated. By extending the line AC (of length $\mathbb{J}$) to intersect the outer (Land's) circle, the points A_1 and A_2 are obtained, such that $AA_1 = \mathbb{I}_1 = \mathbb{I}_0/2 + \mathbb{J}$ and $AA_2 = \mathbb{I}_2 = \mathbb{I}_0/2 - \mathbb{J}$. The direction OA_1 specifies the principal direction 1, at the angle φ_1 relative to the horizontal x-axis, while OA_2 specifies the principal direction 2.

5 Conclusions

We have presented two new derivations of the expressions for the principal moments of inertia and their directions, which also applies to any two-by-two symmetric second order tensor. The first is based on the diagonalization of the moment of inertia tensor and its invariant tensor properties, and the second on the conjugacy property of the moment of inertia vectors. Both derivations deliver separate expressions for the tangens of the

angle that identifies the direction of the maximum and minimum moment of inertia, thus circumventing the ambiguity associated with the use of tan 2φ expression, which is a common outcome of the analysis based on the stationarity condition for the axial moment of inertia. New appealing expressions (41) and (44) are derived, which are capable of specifying the principle directions of inertia, as well as the directions of the maximum and minimum product of inertia. All analytical results are given geometric interpretations based on Mohr's circle.

Of the five presented derivations, the simplest is probably the derivation based on Mohr's circle, which delivers all relevant information by simple geometric considerations. The classical derivation based on the stationarity condition for the axial moment of inertia is most direct and conceptually least involved, but the analysis of the sign of its second derivative, which governs the distinction between the directions of the maximum and minimum moment of inertia, is lengthy and tedious. The derivation based on the diagonalization of the moment of inertia tensor is mathematically elegant and effective, but it requires a familiarity with the matrix algebra and the invariant tensor properties. Similar comments apply to derivations based on the eigenvalue analysis, and the conjugacy property of the moment of inertia vectors. Only the last derivation yields the general expressions (41) and (44).

We believe that the presented analysis sheds some new light on this classical topic, from both conceptual and pedagogical points of view, which may be of interest to the applied mechanics community, in particular to mechanics instructors and their students, and to other interested readers.

References

1. Thornton, S.T., Marion, J.B.: Classical Dynamics od Particles and Systems, 5th edn. Brooks/Cole-Thomson Learning, Belmont (2004)
2. Barber, J.R.: Intermediate Mechanics of Materials. McGraw-Hill, New York (2001)
3. Ugural, A.C.: Mechanics of Materials. John Wiley, New York (2008)
4. Hearn, E.J.: Mechanics of Materials, 2nd edn. Pergamon Press, Oxford (1985)
5. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. The Johns Hopkins University Press, Baltimore (1996)
6. Boresi, A.P., Schmidt, R.J.: Advanced Mechanics of Materials, 6th edn. John-Wiley, New York (2003)
7. Beer, F.P., Johnston, E.R., DeWolf, J.T., Mazurek, D.F.: Mechanics of Materials, 5th edn. McGraw-Hill, New York (2009)
8. Gere, J.M., Goodno, B.J.: Mechanics of Materials, 7th edn. Cengage Learning, Toronto (2009)
9. Craig, R.R.: Mechanics of Materials, 2nd edn. John Wiley, New York (2000)
10. Cook, R.D., Young, W.C.: Advanced Mechanics of Materials, 2nd edn. Prentice Hall, Upper Saddle River (1999)
11. Vable, M.: Intermediate Mechanics of Materials. Oxford University Press, New York (2008)
12. Leon, S.J.: Linear Algebra with Applications, 6th edn. Prentice Hall, Upper Saddle River (2002)
13. Strang, G.: Introduction to Linear Algebra, 3rd edn. Wellesley-Cambridge Press, Wellesley (2003)
14. Malvern, L.E.: Introduction to the Mechanics of a Continuous Medium. Prentice Hall, Upper Saddle River (1969)
15. Drucker, D.C.: Introduction to Mechanics of Deformable Solids. McGraw-Hill, New York (1967)
16. Teodorescu, P.P.: Mechanical Systems, Classical Models (Vol. 1, Particle Mechanics). Springer, Berlin (2007)