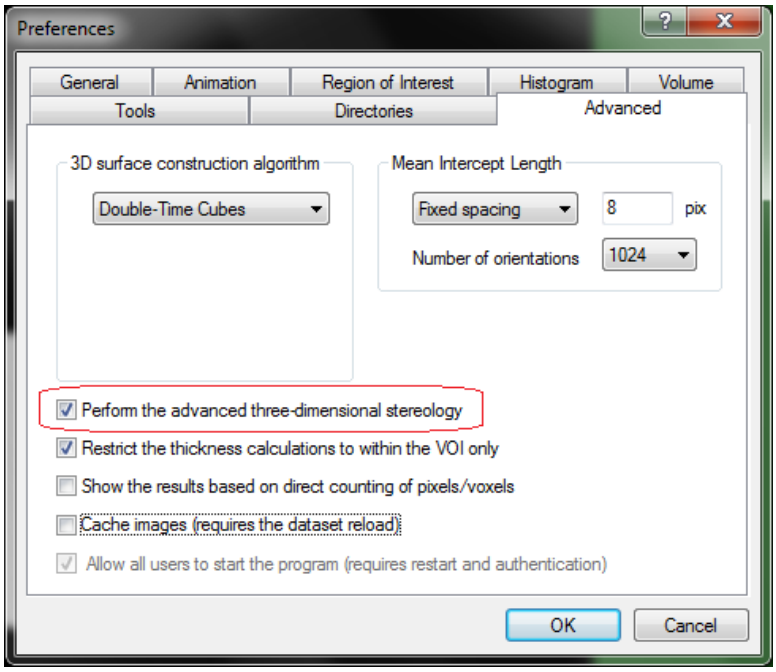
Anisotropy, Mean Intercept Length (MIL) and stereology calculation in CT-Analyser

Method note

MCT-031

1. Stereology and orientation: which way is up, down, XYZ etc.?

The French mathematician Rene Descartes lay ill in bed during a period of military service (for the Bavarian army) in the early 1600's. Watching a house fly buzzing around near the corner of the ceiling, he hit upon the idea of defining the position of the fly in three-dimensional space, as the distances from the corner to the fly, projected orthogonally and separately on the three intersecting walls. The concept of vectors and the XYZ 3D spatial coordinates was born.



There are now several ways to define a location and an orientation in 3D space. Two of these which are used in CT-Analyser ("CTAn") are spherical angular coordinates (polar and azimuth angles theta and phi, or Θ and φ) and the three-vector XYZ system derived from Descartes ("Cartesian" vector coordinates).

In CTAn go to preferences and the advanced tab, and select "perform the advanced 3D stereology".

One of the important outputs of the 3D stereology analysis done in CTAn (using the MIL method) is the predominant orientation of object alignment. (Actually 3 dominant orientations are reported. The first, orientation 1, represents the mean spatial alignment of objects, and the following orientations 2 and 3 are defined orthogonally relative to 1. These three orientations define the ellipsoid "tensor", like the three orthogonal – i.e.

right angles to eachother – axes of an American football or rugby ball, or a bar of soap.)

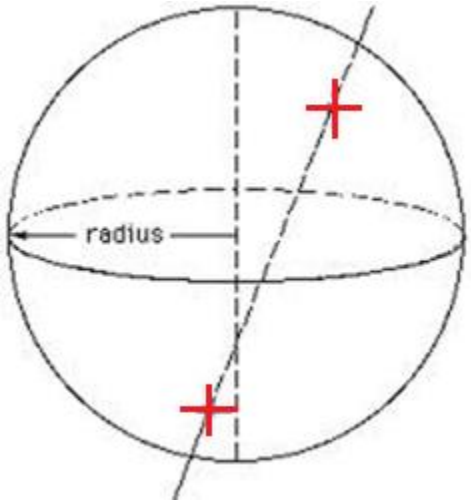


Figure 1. A line in 3D space can be considered as intersecting a reference sphere at two points.

A line in 3D space can be considered as intersecting a reference sphere at two points (figure 1, above). A reconstructed microCT scan of an object or group of objects, might have a preferential or predominant orientation in a certain direction. Identifying and defining this direction is one of the purposes of 3D stereology. This 3D "direction" or orientation can be considered as the line in figure 1.

1.1. Defining 3D orientation by the spherical coordinate system, angles Θ and φ

A description of the spherical coordinate system is described at: http://en.wikipedia.org/wiki/Spherical_coordinate_system

In this system two angles, the polar angle Θ and the azimuth angle φ are equivalent to latitude and longitude or the north-south and east-west respectively in global navigation (figure 2).

The definition of theta Θ as polar angle, phi φ as azimuth angle (and r as radius) corresponds to the ISO 31-11 convention which is described at: http://en.wikipedia.org/wiki/ISO_31-11

To complicate matters, some mathematicians invert these definitions so that theta Θ is the azimuth angle and phi φ the polar angle. Bruker-microCT CT-Analyser follows the ISO 31-11 definition however where theta Θ is the polar angle.

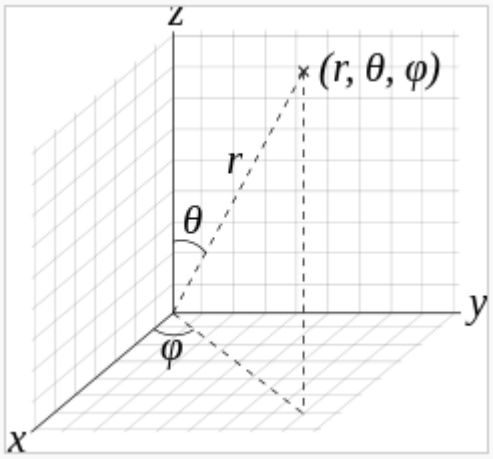


Figure 2. Spherical geometry uses the polar angle and the azimuth angle to define an orientation from the origin point in 3 dimensions (X, Y, Z). The r denotes distance of the line from the origin but is not relevant to the definition of direction.

1.2. Defining 3D orientation by Cartesian XYZ coordinates

In this system, three distances, X, Y and Z from an origin, define 3D orientation (see figure 3). This derives from Rene Descartes' fly and the distances along the 3 walls from the corner to the fly. The eponymous Cartesian vectorial coordinate system is described at: http://en.wikipedia.org/wiki/Cartesian_coordinate_system

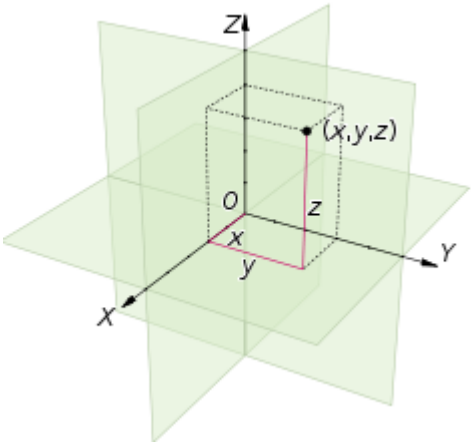


Figure 3. In Cartesian coordinates three orthogonal (at right-angles to eachother) distances x, y and z from origin to end-point, define the 3d orientation of that line.

1.3. Looking at a dataset in CTAn, which direction is "north" (Θ=zero) and which is "east" (φ =zero)?

The polar axis in a Bruker-microCT dataset corresponds to the Z axis, that is the axis perpendicular ("normal") to, or "coming out of" the X-Y plane of reconstructed crossections. Scrolling through crossections is movement in the Z direction. Going up in crossection number, toward the "top" of the dataset, means going "north", i.e. in the direction defined by the polar angle Θ of zero.

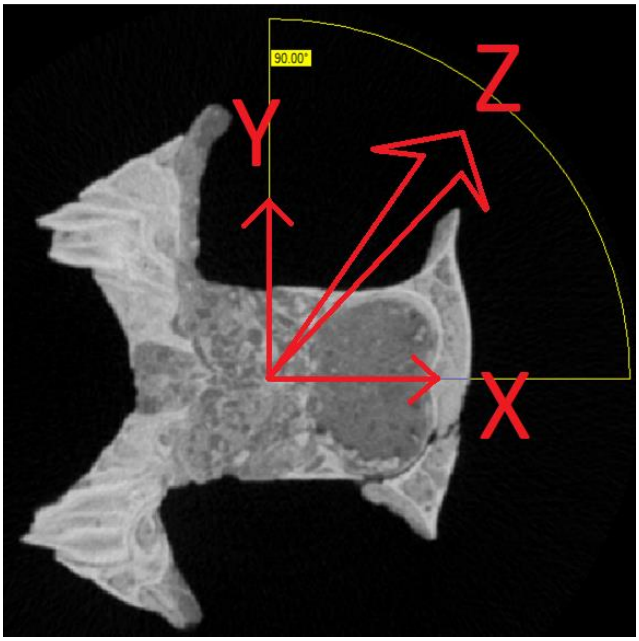


Figure 4. In the crossectional plane the zero φ azimuth angle is to the left ("3 o'clock", the X direction) while the top of the image or "12 o'clock" is 90 degrees (Y). Out of plane toward the viewer is Z.

The azimuth angle φ is the angle in the XY plane of each crossection. By analogy, 3 o'clock or "east" is the zero φ azimuth angle. 12 o'clock is 90 degrees, 11 o'clock is 120 degrees and so on (see figure 4 and 5).

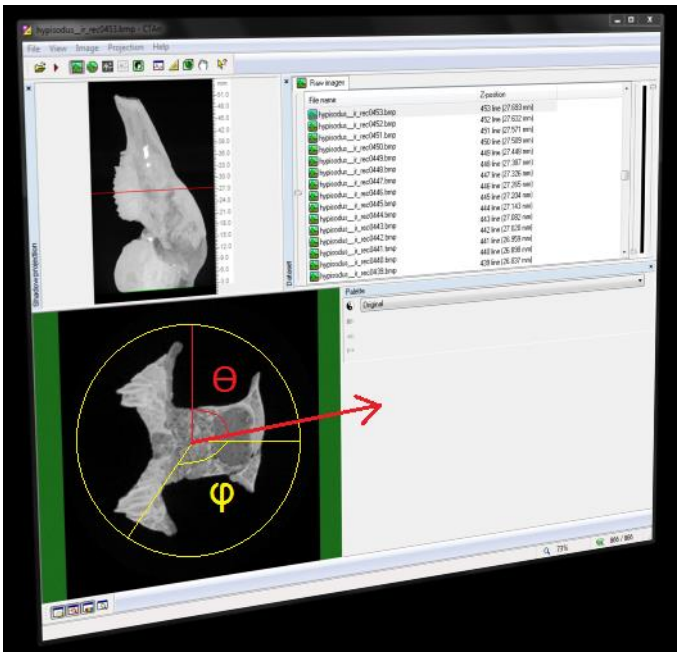
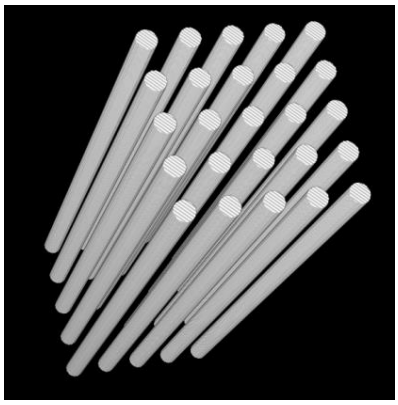


Figure 5. The polar and azimuthal angles in a dataset as viewed in CT-Analyser. The polar angle φ is in the Z direction perpendicular to the crossection image plane (XY). Polar angle Θ = 0 is at 90 degrees to the XY plane. The azimuthal angle is the angle in the XY plane from φ = zero at the "3 o'clock" or due east (right) direction.

1.4. How to interpret the matrix of Cartesian "eigenvectors" in the stereology results

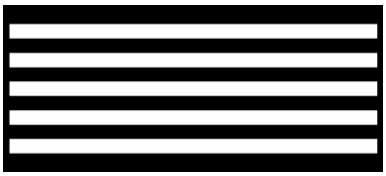


In the stereology results from CTAn 3d analysis the predominant direction defined as Cartesian XYZ coordinates is reported under the name "eigenvectors". The predominant direction is "E-vector 1". (As with the principal orientations, E-vectors 2 and 3 are the orthogonal second and third ellipsoid vectors which follow from vector 1).

For each E-vector, the three numbers listed left to right are X, Y and Z respectively. The E-vectors are a number from 0 to 1.

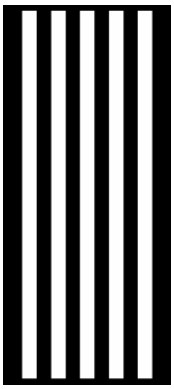
Consider a microCT dataset of an array of parallel rods (image right). We can illustrate what the E-vectors would be for reconstructed datasets of these rods oriented in the X, Y and Z directions respectively.

For objects aligned horizontally in the crossection XY plane (azimuth angle φ = 0), with no Z inclination, the crossections in CTAn would look like this:



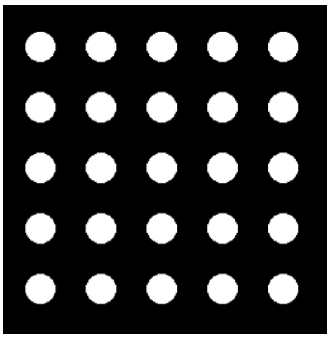
The XYZ values for E-vector 1 in this case will be: 1, ~0, ~0

For objects aligned vertically in the crossection XY plane (azimuth angle φ = 90), with no Z inclination, the crossections in CTAn would look like this:



The XYZ values for E-vector 1 in this case will be: ~0, 1, ~0

For objects aligned in the Z direction (polar angle Θ = 0) perpendicular to the crossection XY plane, the crossections in CTAn would look like this:



The XYZ values for E-vector 1 in this case will be: ~0, ~0, 1

2. Description of the MIL tensor calculation method

This MIL calculation is a function to scan the image array using a three-dimensional version of the directed secant method. The image is scanned by a 3-D test grid at randomly determined orientations. The orientations are defined by spherical angles (theta, phi). Theta is the rotation about the x-axis and phi is the rotation about the z-axis. Based on the threshold value, the image is converted to a binary format and the bone volume fraction is determined. The image is then systematically scanned and intersections are recorded when the binary value of the current voxel differs from the binary

value of the previous voxel. Standard morphology parameters are calculated based on the parallel plate model:

BV/TV = # bone voxels/total # voxels in test sphere

Tb.N = total # intersections/total length of test lines

Tb.Th = BV/TV / Tb.N

Tb.Sp = (1-BV/TV) / Tb.N

BS/BV = 2 * Tb.N / BV/TV

Three-dimensional mean intercept length vectors are calculated for each rotation as:

MIL(theta, phi) = (BV/TV) / ((1/2) * #intersections per unit test line length)

This function fits the 3-D MIL data to an ellipsoid, and determines the goodness of the fit. These methods are based on 2-D methods of Whitehouse (J Microscopy 101:153-168, 1974) and Harrigan and Mann (JMatSci 19:761-767, 1984). The approach is to plot the locus of the end points of the MIL vectors issuing from a common center and fit them to an ellipsoid of general formula:

A*n1^2 + B*n2^2 + C*n3^2 + D*n1*n2 + E*n1*n3 + F*n2*n3 = 1/L^2

where L is the length of the MIL, ni are the direction cosines between L and the base vectors in an arbitrary coordinate system and A...F are the ellipsoid coefficients.

For this code, a multivariable linear least squares fitting technique is used to fit the data by solving the linear system

A * x = b,

where A contains the projection data for each rotation, x is a column vector of the ellipsoid coefficients and b is a column vector of 1/L^2, where L is the magnitude of the MIL vector. The solution is formulated as:

x = (At * A)^-1 * At * b

The sum of the squares of the residuals and the correlation coefficient of the fit are also calculated.

This function defines the MIL tensor based on the ellipsoid fit and determines the principal eigenvectors and eigenvalues, which are used to define the principal MIL vector orientations and magnitudes, respectively. The MIL tensor assumes material orthotropy and is defined as (Harrigan and Mann, JMatSci 19:761-767, 1984):

| ae de/2 ee/2 |

M = | de/2 be fe/2 |

| ee/2 fe/2 ce |

where ae, be, ..,fe are the ellipsoid coefficients

The orientations of the principal MILs are defined by spherical angles, a and b, where a and b correspond to theta and phi, respectively. The magnitudes of the principal MILs are calculated as the inverse of the square root of the absolute values of the eigenvalues. The three degrees of anisotropy are

defined as the relative differences between the three prinicpal MILs and the mean of the principals.

The eigenvalues and eigenvectors are returned as arguments.

This function determines the statistical independence of the eigenvalues and the confidence intervals for the principal MILs. The variance in each of the eigenvalues is determined from the the variance of the estimate of the fit ellipsoid (Se) weighted by the inverse of the appropriate Gaussian multiplier term, Cii. The inverse Gaussian multipliers are coefficients from the inverseof the sum-of-squares matrix. The test statistic is defined as:

$$t = |evali - evalj| / [(Cii+Cjj-2*Cij)*Se]^{0.5}$$

The confidence interval for each eigenvalue is given by:

$$-/+ t(0.05) * (Cii*Se^2)^{0.5}$$

and the 95% confidence intervals are calculated for the principcal MILs.

The statistical independence of the eigenvalues is tested at a 0.05 confidence level (alpha), assuming 128 MIL vectors (rotations). The critical t-value for these conditions is 2.429. Eigenvalues are statistically independent if the test statistic is greater than the critical t-value. The test statistics are calculated for each of eigenvalue 1 versus 2, eigenvalue 2 versus 3, and eigenvalue 1 versus 3. If all three eigenvalues are different, the structure is orthotropic. If two eigenvalues are statistically equivalent, the structure is transversely isotropic. If all three are equivalent, the structure is isotropic. Note that this classification will depend on the significance level you use to determine statistical independence.

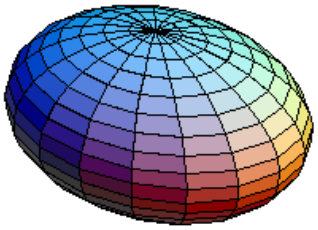
3. References

Whitehouse, J Microscopy 101:153-168, 1974.

Harrigan and Mann, J Mat Sci 19: 761-767, 1984.

APPENDIX 1

The ellipsoid: a description from Wolfram Mathworld (http://mathworld.wolfram.com/Ellipsoid.html)



The general ellipsoid, also called a triaxial ellipsoid, is a quadratic surface which is given in Cartesian coordinates by

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \tag{1}$$

where the semi-axes are of lengths a, b, and c. In spherical coordinates, this becomes

$$\frac{r^2\cos^2\theta\sin^2\phi}{a^2}+\frac{r^2\sin^2\theta\sin^2\phi}{b^2}+\frac{r^2\cos^2\phi}{c^2}=1. \tag{2}$$

If the lengths of two axes of an ellipsoid are the same, the figure is called a spheroid (depending on whether $c<a$ or $c>a$, an oblate spheroid or prolate spheroid, respectively), and if all three are the same, it is a sphere. Tietze (1965, p. 28) calls the general ellipsoid a "triaxial ellipsoid."

There are two families of parallel circular cross sections in every ellipsoid. However, the two coincide for spheroids (Hilbert and Cohn-Vossen 1999, pp. 17-19). If the two sets of circles are fastened together by suitably chosen slits so that they are free to rotate without sliding, the model is

movable. Furthermore, the disks can always be moved into the shape of a sphere (Hilbert and Cohn-Vossen 1999, p. 18).

In 1882, Staude discovered a "thread" construction for an ellipsoid analogous to the taut pencil and string construction of the ellipse (Hilbert and Cohn-Vossen 1999, pp. 19-22). This construction makes use of a fixed framework consisting of an ellipse and a hyperbola.

The parametric equations of an ellipsoid can be written as

$$x = a \cos u \sin v \tag{3}$$

$$y = b \sin u \sin v \tag{4}$$

$$z = c \cos v. \tag{5}$$

for $u \in [0, 2\pi)$ and $v \in [0, \pi]$.

In this parametrization, the coefficients of the first fundamental form are

$$E = \left(b^2 \cos^2 u + a^2 \sin^2 u\right) \sin^2 v \tag{6}$$

$$F = \left(b^2 - a^2\right) \cos u \sin u \cos v \sin v \tag{7}$$

$$G = \left(a^2 \cos^2 u + b^2 \sin^2 u\right) \cos^2 v + c^2 \sin^2 v, \tag{8}$$

and of the second fundamental form are

$$e = \frac{a\, b\, c \sin^2 v}{\sqrt{a^2 b^2 \cos^2 v + c^2 \left(b^2 \cos^2 u + a^2 \sin^2 u\right) \sin^2 v}} \tag{9}$$

$$f = 0 \tag{10}$$

$$g = \frac{a\,b\,c}{\sqrt{a^2\,b^2\cos^2 v + c^2\left(b^2\cos^2 u + a^2\sin^2 u\right)\sin^2 v}}. \tag{11}$$

Also in this parametrization, the Gaussian curvature is

$$K = \frac{a^2\,b^2\,c^2}{\left[a^2\,b^2\cos^2 v + c^2\left(b^2\cos^2 u + a^2\sin^2 u\right)\sin^2 v\right]^2} \tag{12}$$

and the mean curvature is

$$H = \frac{a\,b\,c\left[3\left(a^2+b^2\right)+2\,c^2+\left(a^2+b^2-2\,c^2\right)\cos(2\,v)-2\left(a^2-b^2\right)\cos(2\,u)\sin^2 v\right]}{8\left[a^2\,b^2\cos^2 v + c^2\left(b^2\cos^2 u + a^2\sin^2 u\right)\sin^2 v\right]^{3/2}}. \tag{13}$$

The Gaussian curvature can be given implicitly by

$$K(x,y,z) = \frac{a^2\,b^6\,c^6}{\left[c^4\,b^4 + c^4\left(a^2-b^2\right)y^2 + b^4\left(a^2-c^2\right)z^2\right]^2} \tag{14}$$

$$= \frac{a^6\,b^2\,c^6}{\left[a^4\,c^4 + a^4\left(b^2-c^2\right)z^2 + c^4\left(b^2-a^2\right)x^2\right]^2} \tag{15}$$

$$= \frac{a^6\,b^6\,c^2}{\left[a^4\,b^4 + b^4\left(c^2-a^2\right)x^2 + a^4\left(c^2-b^2\right)y^2\right]^2}. \tag{16}$$

The surface area of an ellipsoid is given by

$$S = 2\,\pi\,a\,b\,\mathrm{ns}\,\theta\int_0^{\theta}\left(\frac{\mathrm{dn}^2\,\theta}{\mathrm{dn}^2\,u} + \frac{\mathrm{cn}^2\,\theta}{\mathrm{cn}^2\,u}\right)d\,u \tag{17}$$

$$= 2\pi\left[c^2 + \frac{b\,c^2}{\sqrt{a^2 - c^2}}\,\theta + b\sqrt{a^2 - c^2}\;E\left(\mathrm{am}\left(\theta\right), k\right)\right], \tag{18}$$

where $\mathrm{ns}\left(\theta\right)$, $\mathrm{dn}\left(\theta\right)$, and $\mathrm{cn}\left(\theta\right)$ are Jacobi elliptic functions with modulus k,

$$k \equiv \frac{e_2}{e_1} \tag{19}$$

$$e_1 \equiv \sqrt{\frac{a^2 - c^2}{a^2}} \tag{20}$$

$$e_2 \equiv \sqrt{\frac{b^2 - c^2}{b^2}}, \tag{21}$$

$E\left(\phi, k\right)$ is an incomplete elliptic integral of the second kind, $\mathrm{am}\left(\phi\right)$ is the Jacobi amplitude with modulus k, and θ is given by inverting the expression

$$e_1 = \mathrm{sn}\left(\theta, k\right), \tag{22}$$

where $\mathrm{sn}\left(\theta\right)$ is another Jacobi elliptic function with modulus k (Bowman 1961, pp. 31-32; error corrected).

Another form of the surface area equation is

$$S = 2\pi\left[c^2 + \frac{b\,c^2}{\sqrt{a^2 - c^2}}\,F\left(\phi, k\right) + b\sqrt{a^2 - c^2}\;E\left(\phi, k\right)\right], \tag{23}$$

where

$$\phi = \sin^{-1}\left(\sqrt{1-\frac{c^2}{a^2}}\right). \tag{24}$$

The surface area can also be obtained directly from the first fundamental form as

$$S = \int_0^{\pi}\int_0^{2\pi}\sqrt{E\,G - F^2}\; d\theta\, d\phi \tag{25}$$

$$= \int_0^{\pi}\sin\phi\int_0^{2\pi}\sqrt{a^2\,b^2\cos^2\phi + c^2\left(b^2\cos^2\theta + a^2\sin^2\theta\right)\sin^2\phi}\; d\theta\, d\phi \tag{26}$$

$$= 2\sqrt{2}\; b\int_0^{\pi}\sqrt{a^2 + c^2 + \left(a^2 - c^2\right)\cos\left(2\phi\right)} \tag{27}$$

$$\sin\phi \times E\left(\frac{c}{b}\sqrt{\frac{2\left(b^2 - a^2\right)}{a^2 + c^2 + \left(a^2 - c^2\right)\cos\left(2\phi\right)}}\;\sin\phi\right) d\phi.$$

The volume of an ellipsoid is

$$V = \frac{4}{3}\pi\, a\, b\, c. \tag{28}$$

The geometric centroids of the half-ellipsoids along the x-, y-, and z-axes are

$$\bar{x} = \frac{3}{16}\,a \tag{29}$$

$$\bar{y} = \frac{3}{16}\,b \tag{30}$$

$$\bar{z} = \frac{3}{16}\,c. \tag{31}$$

The moment of inertia tensor of an ellipsoid is given by

$$I=\begin{bmatrix} \frac{1}{5}M\left(b^2+c^2\right) & 0 & 0 \\ 0 & \frac{1}{5}M\left(a^2+c^2\right) & 0 \\ 0 & 0 & \frac{1}{5}M\left(a^2+b^2\right) \end{bmatrix}. \quad (32)$$

A different parameterization of the ellipsoid is the so-called stereographic ellipsoid, given by the parametric equations

$$x\,(u,v)=\frac{a\left(1-u^2-v^2\right)}{1+u^2+v^2} \quad (33)$$

$$y\,(u,v)=\frac{2\,b\,u}{1+u^2+v^2} \quad (34)$$

$$z\,(u,v)=\frac{2\,c\,v}{1+u^2+v^2}. \quad (35)$$

A third parameterization is the Mercator parameterization

$$x\,(u,v)=a\,\operatorname{sech} v\cos u \quad (36)$$

$$y\,(u,v)=b\,\operatorname{sech} v\sin u \quad (37)$$

$$z\,(u,v)=c\tanh v \quad (38)$$

(Gray 1997).

The support function of the ellipsoid is

$$h=\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}+\frac{z^2}{c^4}\right)^{-1/2}, \quad (39)$$

and the Gaussian curvature is

$$K = \frac{h^4}{a^2 b^2 c^2} \tag{40}$$

(Gray 1997, p. 296).

SEE ALSO:

Confocal Ellipsoidal Coordinates, Confocal Quadrics, Convex Optimization Theory, Ellipsoid Packing, Goursat's Surface, Oblate Spheroid, Prolate Spheroid, Sphere, Spheroid, Superellipsoid

REFERENCES

Beyer, W. H. *CRC Standard Mathematical Tables, 28th ed.* Boca Raton, FL: CRC Press, pp. 131 and 226, 1987.

Bowman, F. *Introduction to Elliptic Functions, with Applications.* New York: Dover, 1961.

Fischer, G. (Ed.). Plate 65 in *Mathematische Modelle aus den Sammlungen von Universitäten und Museen, Bildband.* Braunschweig, Germany: Vieweg, p. 60, 1986.

Gray, A. "The Ellipsoid" and "The Stereographic Ellipsoid." §13.2 and 13.3 in *Modern Differential Geometry of Curves and Surfaces with Mathematica, 2nd ed.* Boca Raton, FL: CRC Press, pp. 301-303, 1997.

Harris, J. W. and Stocker, H. "Ellipsoid." §4.10.1 in *Handbook of Mathematics and Computational Science.* New York: Springer-Verlag, p. 111, 1998.

Hilbert, D. and Cohn-Vossen, S. "The Thread Construction of the Ellipsoid, and Confocal Quadrics." §4 in *Geometry and the Imagination.* New York: Chelsea, pp. 19-25, 1999.

JavaView. "Classic Surfaces from Differential Geometry: Ellipsoid." http://www-sfb288.math.tu-berlin.de/vgp/javaview/demo/surface/common/PaSurface_Ellipsoid.html.

Tietze, H. *Famous Problems of Mathematics: Solved and Unsolved Mathematics Problems from Antiquity to Modern Times.* New York: Graylock Press, pp. 28 and 40-41, 1965.

CITE APPENDIX 1 AS:

Weisstein, Eric W. "Ellipsoid." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/Ellipsoid.html

APPENDIX 2

The Eigenvector: a description from Wolfram Mathworld (http://mathworld.wolfram.com/Eigenvector.html)

Eigenvectors are a special set of vectors associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic vectors, proper vectors, or latent vectors (Marcus and Minc 1988, p. 144).

The determination of the eigenvectors and eigenvalues of a system is extremely important in physics and engineering, where it is equivalent to matrix diagonalization and arises in such common applications as stability analysis, the physics of rotating bodies, and small oscillations of vibrating systems, to name only a few. Each eigenvector is paired with a corresponding so-calledeigenvalue. Mathematically, two different kinds of eigenvectors need to be distinguished: left eigenvectors and right eigenvectors. However, for many problems in physics and engineering, it is sufficient to consider only right eigenvectors. The term "eigenvector" used without qualification in such applications can therefore be understood to refer to a right eigenvector.

The decomposition of a square matrix A into eigenvalues and eigenvectors is known in this work as eigen decomposition, and the fact that this decomposition is always possible as long as the matrix consisting of the eigenvectors of A is square is known as the eigen decomposition theorem.

Define a right eigenvector as a column vector $\mathbf{X}_R$ satisfying

$$\mathsf{A}\,\mathbf{X}_R = \lambda_R\,\mathbf{X}_R, \tag{1}$$

where A is a matrix, so

$$(\mathsf{A} - \lambda_R\,\mathsf{I})\,\mathbf{X}_R = \mathbf{0}, \tag{2}$$

which means the right eigenvalues must have zero determinant, i.e.,

$$\det(\mathbf{A} - \lambda_R \mathbf{I}) = 0. \tag{3}$$

Similarly, define a left eigenvector as a row vector $\mathbf{X}_L$ satisfying

$$\mathbf{X}_L \mathbf{A} = \lambda_L \mathbf{X}_L. \tag{4}$$

Taking the transpose of each side gives

$$(\mathbf{X}_L \mathbf{A})^{\mathrm{T}} = \lambda_L \mathbf{X}_L{}^{\mathrm{T}}, \tag{5}$$

which can be rewritten as

$$\mathbf{A}^{\mathrm{T}} \mathbf{X}_L{}^{\mathrm{T}} = \lambda_L \mathbf{X}_L{}^{\mathrm{T}}. \tag{6}$$

Rearrange again to obtain

$$\left(\mathbf{A}^{\mathrm{T}} - \lambda_L \mathbf{I}\right) \mathbf{X}_L{}^{\mathrm{T}} = \mathbf{0}, \tag{7}$$

which means

$$\det\left(\mathbf{A}^{\mathrm{T}} - \lambda_L \mathbf{I}\right) = 0. \tag{8}$$

Rewriting gives

$$0 = \det\left(\mathbf{A}^{\mathrm{T}} - \lambda_L \mathbf{I}\right) = \det\left(\mathbf{A}^{\mathrm{T}} - \lambda_L \mathbf{I}^{\mathrm{T}}\right) \tag{9}$$

$$= \det(\mathbf{A} - \lambda_L \mathbf{I})^{\mathrm{T}} \tag{10}$$

$$= \det(\mathbf{A} - \lambda_L \mathbf{I}), \tag{11}$$

where the last step follows from the identity

$$\det(\mathbf{A}) = \det\left(\mathbf{A}^{\mathrm{T}}\right). \tag{12}$$

Equating equations (◇) and (11), which are both equal to 0 for arbitrary $\mathbf{A}$ and $\mathbf{X}$, therefore requires that $\lambda_R = \lambda_L \equiv \lambda$, i.e., left and right *eigenvalues* are equivalent, a statement that is not true for *eigenvectors*.

Let $\mathbf{X}_R$ be a matrix formed by the columns of the right eigenvectors and $\mathbf{X}_L$ be a matrix formed by the rows of the left eigenvectors. Let

$$\mathbf{D} \equiv \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}. \tag{13}$$

Then

$$\mathbf{A}\,\mathbf{X}_R = \mathbf{X}_R\,\mathbf{D} \tag{14}$$

$$\mathbf{X}_L\,\mathbf{A} = \mathbf{D}\,\mathbf{X}_L \tag{15}$$

and

$$\mathbf{X}_L\,\mathbf{A}\,\mathbf{X}_R = \mathbf{X}_L\,\mathbf{X}_R\,\mathbf{D} \tag{16}$$

$$\mathbf{X}_L\,\mathbf{A}\,\mathbf{X}_R = \mathbf{D}\,\mathbf{X}_L\,\mathbf{X}_R, \tag{17}$$

so

$$\mathbf{X}_L \mathbf{X}_R \mathsf{D} = \mathsf{D} \mathbf{X}_L \mathbf{X}_R. \quad (18)$$

But this equation is of the form

$$\mathsf{C}\mathsf{D} = \mathsf{D}\mathsf{C} \quad (19)$$

where D is a diagonal matrix, so it must be true that $\mathsf{C} \equiv \mathbf{X}_L \mathbf{X}_R$ is also diagonal. In particular, if A is a symmetric matrix, then the left and right eigenvectors are simply each other's transpose, and if A is a self-adjoint matrix (i.e., it is Hermitian), then the left and right eigenvectors are adjoint matrices.

Eigenvectors may not be equal to the zero vector. A nonzero scalar multiple of an eigenvector is equivalent to the original eigenvector. Hence, without loss of generality, eigenvectors are often normalized to unit length.

While an $n \times n$ matrix always has n eigenvalues, some or all of which may be degenerate, such a matrix may have between 0 and n linearly independent eigenvectors. For example, the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has only the single eigenvector $(1, 0)$.

Eigenvectors may be computed in *Mathematica* using Eigenvectors[*matrix*]. This command always returns a list of length n, so any eigenvectors that are not linearly independent are returned as zero vectors. Eigenvectors and eigenvalues can be returned together using the command Eigensystem[*matrix*].

Given a 3×3 matrix A with eigenvectors $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ and corresponding eigenvalues λ_1, λ_2, and λ_3, then an arbitrary vector $\mathbf{y}$ can be written

$$\mathbf{y} = b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + b_3 \mathbf{x}_3. \quad (20)$$

Applying the matrix $\mathbf{A}$,

$$\mathbf{A}\,\mathbf{y} = b_1\,\mathbf{A}\,\mathbf{x}_1 + b_2\,\mathbf{A}\,\mathbf{x}_2 + b_3\,\mathbf{A}\,\mathbf{x}_3 \tag{21}$$

$$= \lambda_1\left(b_1\,\mathbf{x}_1 + \frac{\lambda_2}{\lambda_1}\,b_2\,\mathbf{x}_2 + \frac{\lambda_3}{\lambda_1}\,b_3\,\mathbf{x}_3\right), \tag{22}$$

so

$$\mathbf{A}^n\,\mathbf{y} = \lambda_1^n\left[b_1\,\mathbf{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^n b_2\,\mathbf{x}_2 + \left(\frac{\lambda_3}{\lambda_1}\right)^n b_3\,\mathbf{x}_3\right]. \tag{23}$$

If $\lambda_1 > \lambda_2, \lambda_3$, and $b_1 \neq 0$, it therefore follows that

$$\lim_{n\to\infty} \mathbf{A}^n\,\mathbf{y} = \lambda_1^n\,b_1\,\mathbf{x}_1, \tag{24}$$

so repeated application of the matrix to an arbitrary vector amazingly results in a vector proportional to the eigenvector with largest eigenvalue.

SEE ALSO:

Eigen Decomposition, Eigen Decomposition Theorem, Eigenfunction, Eigenvalue, Left Eigenvector, Matrix, Matrix Diagonalization, Matrix Equation, Right Eigenvector

REFERENCES

Arfken, G. "Eigenvectors, Eigenvalues." §4.7 in *Mathematical Methods for Physicists, 3rd ed.* Orlando, FL: Academic Press, pp. 229-237, 1985.

Marcus, M. and Minc, H. *Introduction to Linear Algebra.* New York: Dover, p. 145, 1988.

Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; and Vetterling, W. T. "Eigensystems." Ch. 11 in *Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed.* Cambridge, England: Cambridge University Press, pp. 449-489, 1992.

Referenced on Wolfram|Alpha: Eigenvector

CITE APPENDIX 2 AS:

Weisstein, Eric W. "Eigenvector." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/Eigenvector.html

APPENDIX 3

The Eigenvalue: a description from Wolfram Mathworld (http://mathworld.wolfram.com/Eigenvalue.html)

Eigenvalues are a special set of scalars associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic roots, characteristic values (Hoffman and Kunze 1971), proper values, or latent roots (Marcus and Minc 1988, p. 144).

The determination of the eigenvalues and eigenvectors of a system is extremely important in physics and engineering, where it is equivalent to matrix diagonalization and arises in such common applications as stability analysis, the physics of rotating bodies, and small oscillations of vibrating systems, to name only a few. Each eigenvalue is paired with a corresponding so-called eigenvector (or, in general, a corresponding right eigenvector and a corresponding left eigenvector; there is no analogous distinction between left and right for eigenvalues).

The decomposition of a square matrix A into eigenvalues and eigenvectors is known in this work as eigen decomposition, and the fact that this decomposition is always possible as long as the matrix consisting of the eigenvectors of A is square is known as the eigen decomposition theorem.

The Lanczos algorithm is an algorithm for computing the eigenvalues and eigenvectors for large symmetric sparse matrices.

Let A be a linear transformation represented by a matrix A. If there is a vector $\mathbf{X} \in \mathbb{R}^n \neq \mathbf{0}$ such that

$$\mathsf{A}\mathbf{X} = \lambda \mathbf{X} \qquad (1)$$

for some scalar λ, then λ is called the eigenvalue of A with corresponding (right) eigenvector $\mathbf{X}$.

Letting A be a $k \times k$ square matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} \tag{2}$$

with eigenvalue λ, then the corresponding eigenvectors satisfy

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \tag{3}$$

which is equivalent to the homogeneous system

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4}$$

Equation (4) can be written compactly as

$$(\mathbf{A} - \lambda\, \mathrm{I})\, \mathbf{X} = \mathbf{0}, \tag{5}$$

where I is the identity matrix. As shown in Cramer's rule, a linear system of equations has nontrivial solutions iff the determinant vanishes, so the solutions of equation (5) are given by

$$\det(\mathbf{A} - \lambda\, \mathrm{I}) = 0. \tag{6}$$

This equation is known as the characteristic equation of A, and the left-hand side is known as the characteristic polynomial.

For example, for a 2×2 matrix, the eigenvalues are

$$\lambda_{\pm} = \frac{1}{2}\left[(a_{11} + a_{22}) \pm \sqrt{4\,a_{12}\,a_{21} + (a_{11} - a_{22})^2}\,\right], \tag{7}$$

which arises as the solutions of the characteristic equation

$$x^2 - x\,(a_{11} + a_{22}) + (a_{11}\,a_{22} - a_{12}\,a_{21}) = 0. \tag{8}$$

If all k eigenvalues are different, then plugging these back in gives $k-1$ independent equations for the k components of each corresponding eigenvector, and the system is said to be nondegenerate. If the eigenvalues are n-fold degenerate, then the system is said to be degenerate and the eigenvectors are not linearly independent. In such cases, the additional constraint that the eigenvectors be orthogonal,

$$\mathbf{X}_i \cdot \mathbf{X}_j = |\mathbf{X}_i|\,|\mathbf{X}_j|\,\delta_{ij}, \tag{9}$$

where δ_{ij} is the Kronecker delta, can be applied to yield n additional constraints, thus allowing solution for the eigenvectors.

Eigenvalues may be computed in *Mathematica* using Eigenvalues[*matrix*]. Eigenvectors and eigenvalues can be returned together using the command Eigensystem[*matrix*].

Assume we know the eigenvalue for

$$\mathsf{A}\,\mathbf{X} = \lambda\,\mathbf{X}. \tag{10}$$

Adding a constant times the identity matrix to A,

$$(\mathsf{A} + c\,\mathsf{I})\,\mathbf{X} = (\lambda + c)\,\mathbf{X} \equiv \lambda'\,\mathbf{X}, \tag{11}$$

so the new eigenvalues equal the old plus c. Multiplying A by a constant c

$$(c\,\mathbf{A})\,\mathbf{X} = c\,(\lambda\,\mathbf{X}) \equiv \lambda'\,\mathbf{X}, \tag{12}$$

so the new eigenvalues are the old multiplied by c.

Now consider a similarity transformation of $\mathbf{A}$. Let $|\mathbf{A}|$ be the determinant of $\mathbf{A}$, then

$$|\mathbf{Z}^{-1}\,\mathbf{A}\,\mathbf{Z} - \lambda\,\mathbf{I}| = |\mathbf{Z}^{-1}\,(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{Z}| \tag{13}$$

$$= |\mathbf{Z}|\,|\mathbf{A} - \lambda\,\mathbf{I}|\,|\mathbf{Z}^{-1}| \tag{14}$$

$$= |\mathbf{A} - \lambda\,\mathbf{I}|, \tag{15}$$

so the eigenvalues are the same as for $\mathbf{A}$.

SEE ALSO:

Brauer's Theorem, Characteristic Equation, Characteristic Polynomial, Complex Matrix, Condition Number, Eigen Decomposition, Eigen Decomposition Theorem, Eigenfunction, Eigenvector, Frobenius Theorem, Gershgorin Circle Theorem, Lanczos Algorithm, Lyapunov's First Theorem, Lyapunov's Second Theorem, Matrix Diagonalization, Ostrowski's Theorem, Perron's Theorem, Perron-Frobenius Theorem, Poincaré Separation Theorem, Random Matrix, Real Matrix, Schur's Inequalities, Similarity Transformation, Sturmian Separation Theorem, Sylvester's Inertia Law, Wielandt's Theorem

REFERENCES

Arfken, G. "Eigenvectors, Eigenvalues." §4.7 in *Mathematical Methods for Physicists, 3rd ed.* Orlando, FL: Academic Press, pp. 229-237, 1985.

Hoffman, K. and Kunze, R. "Characteristic Values." §6.2 in *Linear Algebra, 2nd ed.* Englewood Cliffs, NJ: Prentice-Hall, p. 182, 1971.

Kaltofen, E. "Challenges of Symbolic Computation: My Favorite Open Problems." *J. Symb. Comput.* **29**, 891-919, 2000.

Marcus, M. and Minc, H. *Introduction to Linear Algebra.* New York: Dover, p. 145, 1988.

Nash, J. C. "The Algebraic Eigenvalue Problem." Ch. 9 in *Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation, 2nd ed.* Bristol, England: Adam Hilger, pp. 102-118, 1990.

Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; and Vetterling, W. T. "Eigensystems." Ch. 11 in *Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed.* Cambridge, England: Cambridge University Press, pp. 449-489, 1992.

Referenced on Wolfram|Alpha: Eigenvalue

CITE APPENDIX 3 AS:

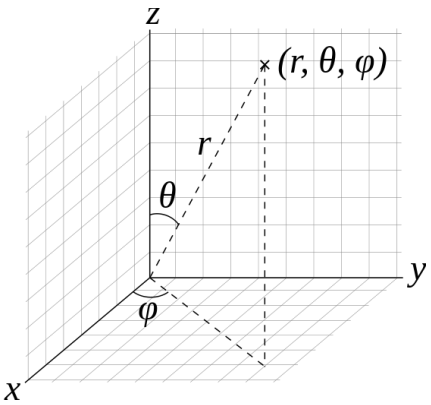
Weisstein, Eric W. "Eigenvalue." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/Eigenvalue.html

APPENDIX 4

The Spherical Coordinate system: a description from Wikipedia (http://en.wikipedia.org/wiki/Spherical_coordinate_system)

In mathematics, a spherical coordinate system is a coordinate system for three-dimensional space where the position of a point is specified by three numbers: the radial distance of that point from a fixed origin, its polar angle measured from a fixed zenith direction, and the azimuth angle of its orthogonal projection on a reference plane that passes through the origin and is orthogonal to the zenith, measured from a fixed reference direction on that plane.

The radial distance is also called the **radius** or **radial coordinate**. The polar angle may be called **colatitude**, **zenith angle**, **normal angle**, or **inclination angle**.



Spherical coordinates (*r*, *θ*, *φ*) as commonly used in *physics*: radial distance *r*, polar angle *θ* (theta), and azimuthal angle *φ* (phi). The symbol *ρ* (rho) is often used instead of *r*.

The use of symbols and the order of the coordinates differs between sources. In one system frequently encountered in physics (*r, θ, φ*) gives the radial distance, polar angle, and azimuthal angle, whereas in another

system used in many mathematics books (r, θ, φ) gives the radial distance, azimuthal angle, and polar angle. In both systems ρ is often used instead of r. Other conventions are also used so great care needs to be taken to check which one is being used.

A number of different spherical coordinate systems following other conventions are used outside mathematics. In a geographical coordinate system positions are measured in latitude, longitude and height or altitude. There are a number of different celestial coordinate systems based on different fundamental planes and with different terms for the various coordinates. The spherical coordinate systems used in mathematics normally use radians rather than degrees and measure the azimuthal angle counter-clockwise rather than clockwise. The inclination angle is often replaced by the **elevation angle** measured from the reference plane. Elevation angle of zero is at the horizon.

The concept of spherical coordinates can be extended to higher dimensional spaces and are then referred to as hyperspherical coordinates.

Definition

To define a spherical coordinate system, one must choose two orthogonal directions, the *zenith* and the *azimuth reference*, and an *origin* point in space. These choices determine a reference plane that contains the origin and is perpendicular to the zenith. The spherical coordinates of a point *P* are then defined as follows:

- the *radius* or *radial distance* is the Euclidean distance from the origin *O* to *P*.
- the *inclination* (or *polar angle*) is the angle between the zenith direction and the line segment *OP*.
- the *azimuth* (or *azimuthal angle*) is the signed angle measured from the azimuth reference direction to the orthogonal projection of the line segment *OP* on the reference plane.

The sign of the azimuth is determined by choosing what is a *positive* sense of turning about the zenith. This choice is arbitrary, and is part of the coordinate system's definition.

The *elevation* angle is 90 degrees ($\pi/2$ radians) minus the inclination angle.

If the inclination is zero or 180 degrees (π radians), the azimuth is arbitrary. If the radius is zero, both azimuth and inclination are arbitrary.

In linear algebra, the vector from the origin *O* to the point *P* is often called the position vector of *P*.

Conventions

Several different conventions exist for representing the three coordinates, and for the order in which they should be written. The use of (r, θ, φ) to denote, respectively, radial distance, inclination (or elevation), and azimuth, is common practice in physics, and is specified by **ISO standard 31-11**.

However, some authors (including mathematicians) use φ for inclination (or elevation) and θ for azimuth, which "provides a logical extension of the usual polar coordinates notation" [1]. Some authors may also list the azimuth before the inclination (or elevation), and/or use ρ instead of r for radial distance. Some combinations of these choices result in a left-handed coordinate system. The standard convention (r, θ, φ) conflicts with the usual notation for the two-dimensional polar coordinates, where θ is often used for the azimuth. It may also conflict with the notation used for three-dimensional cylindrical coordinates [1].

The angles are typically measured in degrees (°) or radians (rad), where 360° = 2π rad. Degrees are most common in geography, astronomy, and engineering, whereas radians are commonly used in mathematics and theoretical physics. The unit for radial distance is usually determined by the context.

REFERENCES

1. Eric W. Weisstein (2005-10-26). "Spherical Coordinates". MathWorld. (http://mathworld.wolfram.com/SphericalCoordinates.html)