Distance Between a Circle and a Disk in 3D

David Eberly
Geometric Tools, LLC
http://www.geometrictools.com/

Created: May 4, 2005
Last Modified: March 1, 2008

Contents

1 Introduction **2**

2 Distance from Point to Disk **2**

3 Distance from Curve to Disk **3**

4 Distance from Circle to Disk **3**

4.1 Partitioning the Interval 4

4.2 Minimization Case 1 4

4.3 Minimation Case 2 5

1 Introduction

A circle in 3D is represented by a center $\mathbf{C}$, a radius r, and a plane containing the circle, $\mathbf{N} \cdot (\mathbf{X} - \mathbf{C}) = 0$, where $\mathbf{N}$ is a unit length normal to the plane. If $\mathbf{U}$ and $\mathbf{V}$ are also unit length vectors so that $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{N}$ form a right-handed orthonormal coordinate system (the matrix with these vectors as columns is orthonormal with determinant 1), then the circle is parameterized by

$$\mathbf{X} = \mathbf{C} + r(\cos(\theta)\mathbf{U} + \sin(\theta)\mathbf{V})$$

for angles $\theta \in [0, 2\pi)$. A disk in 3D is the set of points

$$\mathbf{X} = \mathbf{C} + s(\cos(\theta)\mathbf{U} + \sin(\theta)\mathbf{V})$$

where $0 \le s \le r$.

2 Distance from Point to Disk

Any point $\mathbf{P}$ can be represented in terms of the coordinate system $\{\mathbf{C}; \mathbf{U}, \mathbf{V}, \mathbf{N}\}$ by

$$\mathbf{P} = \mathbf{C} + x\mathbf{U} + y\mathbf{V} + z\mathbf{N}$$

where $x = \mathbf{U} \cdot (\mathbf{P} - \mathbf{C})$, $y = \mathbf{V} \cdot (\mathbf{P} - \mathbf{C})$, and $z = \mathbf{N} \cdot (\mathbf{P} - \mathbf{C})$. The distance from $\mathbf{P}$ to the disk involves finding the closest point, call it $\mathbf{K}$, on the disk to $\mathbf{P}$. This point can be determined by looking at the projection of $\mathbf{P}$ onto the plane of the disk,

$$\mathbf{Q} = \mathbf{C} + x\mathbf{U} + y\mathbf{V}$$

If $\mathbf{Q}$ is inside the disk, then it is the closest point to $\mathbf{P}$ and the distance δ is $|z|$. The condition for being inside the disk is $x^2 + y^2 \le r^2$. The closest point and squared distance are

$$\mathbf{K} = \mathbf{Q}, \quad \delta^2 = z^2, \quad \text{for } x^2 + y^2 \le r^2$$

If $\mathbf{Q}$ is outside the disk, then the closest point to $\mathbf{P}$ is on the disk's circular boundary along the ray from $\mathbf{C}$ to $\mathbf{Q}$. The closest point and squared distance are

$$\mathbf{K} = \mathbf{C} + r\,\frac{\mathbf{Q} - \mathbf{C}}{|\mathbf{Q} - \mathbf{C}|}, \quad \delta^2 = |\mathbf{P} - \mathbf{K}|^2, \quad \text{for } x^2 + y^2 > r^2$$

Notice that

$$\begin{aligned} \mathbf{P} - \mathbf{K} &= (\mathbf{P} - \mathbf{C}) - (\mathbf{K} - \mathbf{C}) \\ &= (x\mathbf{U} + y\mathbf{V} + z\mathbf{N}) - r\left(\tfrac{x\mathrm{U} + y\mathrm{V}}{|x\mathrm{U} + y\mathrm{V}|}\right) \\ &= x\left(1 - \tfrac{r}{\sqrt{x^2+y^2}}\right)\mathbf{U} + y\left(1 - \tfrac{r}{\sqrt{x^2+y^2}}\right)\mathbf{V} + z\mathbf{W} \end{aligned}$$

which implies

$$|\mathbf{P} - \mathbf{K}|^2 = (x^2 + y^2)\left(1 - \frac{r}{\sqrt{x^2 + y^2}}\right)^2 + z^2 = x^2 + y^2 + z^2 + r^2 - 2r\sqrt{x^2 + y^2}$$

In summary, the squared distance from $\mathbf{P}$ to the disk is

$$\delta^2 = \left\{ \begin{array}{lcl} z^2 & , & x^2+y^2 \le r^2 \\ x^2+y^2+z^2+r^2-2r\sqrt{x^2+y^2} & , & x^2+y^2 > r^2 \end{array} \right\}$$

3 Distance from Curve to Disk

A parametric curve is of the form $\mathbf{P}(t)$ for $t \in [t_0, t_1]$. Let us assume that the curve is continuously differentiable. The squared distance from the curve to the disk is

$$\bar{\delta}^2 = \min_{t\in[t_0,t_1]} \delta^2(t) = \min_{t\in[t_0,t_1]} \left\{ \begin{array}{lcl} z(t)^2 & , & x(t)^2+y(t)^2 \le r^2 \\ x(t)^2+y(t)^2+z(t)^2+r^2-2r\sqrt{x(t)^2+y(t)^2} & , & x(t)^2+y(t)^2 > r^2 \end{array} \right\}$$

The condition $x(t)^2 + y(t)^2 \le r^2$ partitions the interval $I = [t_0, t_1]$ into two sets of subintervals. One set of subintervals satisfies the condition; the other set does not.

If $I_0 \subseteq I$ is a subinterval that satisfies $x(t)^2 + y(t)^2 \le r^2$, then we need to minimize $z(t)^2$ for $t \in I_0 = [a, b]$. This is a calculus problem. The minimum occurs either at the end points $t = a$ or $t = b$ or at a point where the derivative is zero. The following equation sets the derivative to zero; a factor 2 has already been cancelled from the equation:

$$z(t)z'(t) = 0 \tag{1}$$

Naturally, this equation is solved in two parts: $z(t) = 0$ or $z'(t) = 0$.

If $I_1 \subseteq I$ is a subinterval that satisfies $x(t)^2 + y(t)^2 > r^2$, then we need to minimize $x(t)^2 + y(t)^2 + z(t)^2 + r^2 - 2r\sqrt{x(t)^2 + y(t)^2}$ for $t \in I_1 = [a, b]$. The minimum occurs either at $t = a$ or $t = b$ or at a point where the derivatives is zero. The following equation sets the derivative to zero; a factor 2 has already been cancelled from the equation:

$$x(t)x'(t) + y(t)y'(t) + z(t)z'(t) - r\,\frac{x(t)x'(t) + y(t)y'(t)}{\sqrt{x(t)^2 + y(t)^2}} = 0$$

Formally, the square root may be eliminated, but you need to take care in identifying extraneous solutions introduced by the squaring:

$$\left(x(t)^2 + y(t)^2\right)\left(x(t)x'(t) + y(t)y'(t) + z(t)z'(t)\right)^2 - r^2\left(x(t)x'(t) + y(t)y'(t)\right)^2 = 0 \tag{2}$$

The reason for the reformulation is that if the parametric curve is polynomial in t, then the zero-derivative equations are polynomials in t, in which case standard numerical packages for polynomial root finding may be used to solve the problem.

After all subintervals are processed for their minima, the minimum of the minima is chosen for $\bar{\delta}^2$.

4 Distance from Circle to Disk

Consider the parametric circle that has center $\mathbf{T}$, radius ρ, and lies on the plane $\mathbf{M} \cdot (\mathbf{X} - \mathbf{T}) = 0$ for a unit-length normal $\mathbf{M}$. Let $\mathbf{A}$ and $\mathbf{B}$ be unit-length vectors so that $\{\mathbf{A}, \mathbf{B}, \mathbf{M}\}$ is a right-handed orthonormal set. The parametric form for the circle is

$$\mathbf{P}(t) = \mathbf{T} + \rho\left(\cos(t)\mathbf{A} + \sin(t)\mathbf{B}\right) = \mathbf{C} + x(t)\mathbf{U} + y(t)\mathbf{V} + z(t)\mathbf{N}$$

for $t \in [0, 2\pi)$. The middle expression is the parametric form. The right-hand side specifies the representation with respect to the coordinate system of the disk. Thus,

$$\begin{aligned}
x(t) &= \mathbf{U}\cdot(\mathbf{T}-\mathbf{C}) + (\rho\cos(t))\mathbf{U}\cdot\mathbf{A} + (\rho\sin(t))\mathbf{U}\cdot\mathbf{B} = a_0 + a_1\gamma + a_2\sigma \\
y(t) &= \mathbf{V}\cdot(\mathbf{T}-\mathbf{C}) + (\rho\cos(t))\mathbf{V}\cdot\mathbf{A} + (\rho\sin(t))\mathbf{V}\cdot\mathbf{B} = b_0 + b_1\gamma + b_2\sigma \\
z(t) &= \mathbf{N}\cdot(\mathbf{T}-\mathbf{C}) + (\rho\cos(t))\mathbf{N}\cdot\mathbf{A} + (\rho\sin(t))\mathbf{N}\cdot\mathbf{B} = c_0 + c_1\gamma + c_2\sigma
\end{aligned}$$

where

$$\begin{array}{lll}
a_0 = \mathbf{U}\cdot(\mathbf{T}-\mathbf{C}), & b_0 = \mathbf{V}\cdot(\mathbf{T}-\mathbf{C}), & c_0 = \mathbf{N}\cdot(\mathbf{T}-\mathbf{C}), \\
a_1 = \rho\mathbf{U}\cdot\mathbf{A}, & b_1 = \rho\mathbf{V}\cdot\mathbf{A}, & c_1 = \rho\mathbf{N}\cdot\mathbf{A}, \\
a_2 = \rho\mathbf{U}\cdot\mathbf{B}, & b_2 = \rho\mathbf{V}\cdot\mathbf{B}, & c_2 = \rho\mathbf{N}\cdot\mathbf{B},
\end{array}$$

and $\gamma = \cos(t)$ and $\sigma = \sin(t)$. Naturally, $\gamma^2 + \sigma^2 = 1$.

4.1 Partitioning the Interval

The interval $[0, 2\pi)$ is partitioned by setting $x(t)^2 + y(t)^2 - r^2 = 0$,

$$\begin{aligned}
0 &= x(t)^2 + y(t)^2 - r^2 \\
&= (a_0 + a_1\gamma + a_2\sigma)^2 + (b_0 + b_1\gamma + b_2\sigma)^2 - r^2 \\
&= (a_1^2 + b_1^2)\gamma^2 + 2(a_1a_2 + b_1b_2)\gamma\sigma + (a_2^2 + b_2^2)\sigma^2 + 2(a_0a_1 + b_0b_1)\gamma + 2(a_0a_2 + b_0b_2)\sigma + (a_0^2 + b_0^2 - r^2)
\end{aligned}$$

This is a quadratic equation in γ and σ. Using elimination theory for this equation and for $\gamma^2 + \sigma^2 = 1$, you can obtain a quartic equation in γ:

$$d_0 + d_1\gamma + d_2\gamma^2 + d_3\gamma^3 + d_4\gamma^4 = 0$$

If $\bar{\gamma}$ is a real-valued root of the equation, then the corresponding t values are solutions to $\cos(\bar{t}) = \bar{\gamma}$; in particular, the $\bar{t}$ are chosen in $[0, 2\pi)$. If $\bar{t}_0$ and $\bar{t}_1$ are two consecutive values in the partition, then $x(t)^2 + y(t)^2 - r^2 \leq 0$ on $[\bar{t}_0, \bar{t}_1]$ or $x(t)^2 + y(t)^2 - r^2 \geq 0$ on $[\bar{t}_0, \bar{t}_1]$. If the first case, then Equation (1) is solved during the minimization phase. If the second case, then Equation (2) is solved during the minimization phase.

4.2 Minimization Case 1

Replacing $z(t)$ into Equation (1) leads to

$$z(t) = c_0 + c_1\gamma + c_2\sigma = 0 \ \text{ or } \ z'(t) = -c_1\sigma + c_2\gamma = 0$$

For the first condition, we have $c_0 + c_1\gamma + c_2\sigma = 0$ and $\gamma^2 + \sigma^2 = 1$. Substituting the first (linear) equation into the second (quadratic) equation leads to a quadratic equation in either γ or σ, the choice depending on whether c_2 is zero or not. The second condition is handled similarly. For example, suppose that $c_2 \neq 0$; then

$$\sigma = -(c_0 + c_1\gamma)/c_2$$

and

$$1 = \gamma^2 + \left(\frac{c_0 + c_1\gamma}{c_2}\right)^2$$

which is equivalent to

$$(c_1^2 + c_2^2)\gamma^2 + (2c_0c_1)\gamma + (c_0^2 - c_2^2) = 0$$

If $\bar{\gamma}$ is a real-valued root to this equation, then $\bar{\sigma} = -(c_0 + c_1\bar{\gamma})/c_2$. The squared distance for this specific case can be calculated using the formula derived earlier.

4.3 Minimation Case 2

Replacing $x(t)$, $y(t)$, and $z(t)$ into Equation (2) leads to a formal polynomial equation of degree 6. This is clear from the expression

$$\begin{aligned}
x(t)x'(t) + y(t)y'(t) + z(t)z'(t) &= (a_0 + a_1\gamma + a_2\sigma)(-a_1\sigma + a_2\gamma) + (b_0 + b_1\gamma + b_2\sigma)(-b_1\sigma + b_2\gamma) \\
&\quad + (c_0 + c_1\gamma + c_2\sigma)(-c_1\sigma + c_2\gamma) \\
&= a_0(-a_1\sigma + a_2\gamma) + b_0(-b_1\sigma + b_2\gamma) + c_0(-c_1\sigma + c_2\gamma) \\
&\quad + (a_1a_2 + b_1b_2 + c_1c_2)(\gamma^2 - \sigma^2) \\
&\quad + ((a_2^2 + b_2^2 + c_2^2) - (a_1^2 + b_1^2 + c_1^2))\gamma\sigma
\end{aligned} \tag{3}$$

This is quadratic in γ and σ. When you square it, you get degree 4, and then when you multiply by the quadratic $x^2 + y^2$, you wind up with degree 6. However, notice that $(a_1, b_1, c_1)/\rho$ are the coordinates of the unit-length vector $\mathbf{A}$ in the disk's coordinate system and $(a_2, b_2, c_2)/\rho$ are the coordinates of the unit-length vector $\mathbf{B}$ in the disk's coordinate system. This means two things. First, $|(a_1, b_1, c_1)/\rho| = |(a_2, b_2, c_2)/\rho| = 1$ since $\mathbf{A}$ and $\mathbf{B}$ both have length 1, in which case $a_1^2+b_1^2+c_1^2 = a_2^2+b_2^2+c_2^2$. Second, $(a_1, b_1, c_1)/\rho \cdot (a_2, b_2, c_2)/\rho = 0$ since $\mathbf{A}$ and $\mathbf{B}$ are perpendicular, in which case $a_1a_2 + b_1b_2 + c_1c_2 = 0$. Consequently, the quadratic term in Equation (3) is zero and

$$x(t)x'(t) + y(t)y'(t) + z(t)z'(t) = a_0(-a_1\sigma + a_2\gamma) + b_0(-b_1\sigma + b_2\gamma) + c_0(-c_1\sigma + c_2\gamma)$$

Equation (2) is therefore degree 4 in γ and σ. Using elimination theory with this equation and with $\gamma^2+\sigma^2 = 1$ leads to a degree 8 equation in γ. The roots are computed numerically, each root $\bar{\gamma}$ is used to generate candidates $\bar{\sigma}$, and pairs $(\bar{\gamma}, \bar{\sigma})$ are used to compute candidate minimum distances.