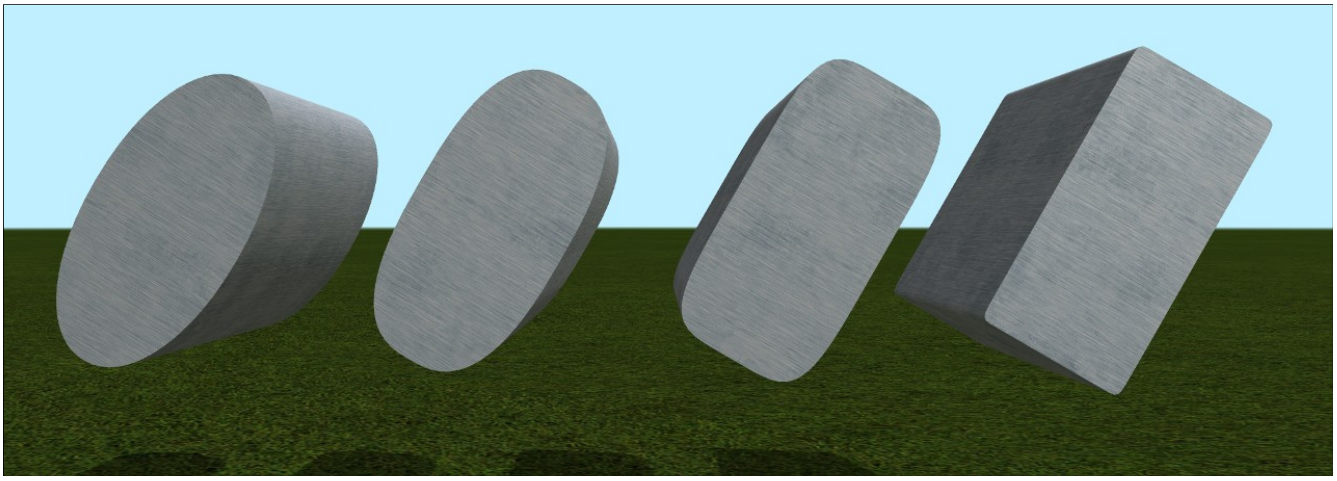
Distance Hyperellipse to Point

Hyperellipse is a curve defined by the equation:

$$\left|\frac{x}{a}\right|^n + \left|\frac{y}{b}\right|^n = 1\,,\quad n \geqslant 2\,. \tag{1}$$



Hyperelliptic cylinders of length 1, a=0.15, b=0.3. Left to right: n=2.1, 2.8, 5, 30.

Parametric equation of ¼ curve ($x, y \geqslant 0$):

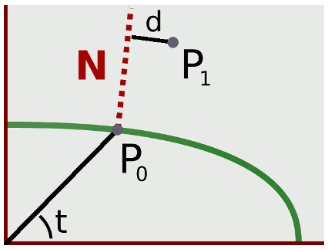
$$x = \cos(t)\left(\frac{\cos^n t}{a^n} + \frac{\sin^n t}{b^n}\right)^{-1/n},\quad y = \sin(t)\left(\frac{\cos^n t}{a^n} + \frac{\sin^n t}{b^n}\right)^{-1/n},\quad n \geqslant 2,\quad t \in [0, \pi/2]. \tag{2}$$

After taking partial derivatives from (1), plugging x, y from (2), then simplification it results in vector-valued parametric equation of normal vector to the curve:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right);\quad \vec{N} = \left(\frac{\cos^{n-1} t}{a^n}, \frac{\sin^{n-1} t}{b^n}\right). \tag{3}$$

“Corner” on the interval $t \in [0, \pi/2]$ is the point $\left(a \cdot 2^{-1/n},\, b \cdot 2^{-1/n}\right)$ where $t = \arctan(b/a)$, $\vec{N} = (1/a,\ 1/b)$.

<u>Find the point P_0 on hyperellipse closest to given point P_1.</u>



Using formula – distance d from point P_1 to the line defined by point P_0 on the line and direction vector N:

$$\mathrm{d} = \frac{(\vec{P}_1 - \vec{P}_0) \cdot \vec{N}^{\perp}}{\|\vec{N}\|} = \frac{(P_{1x} - P_{0x}) N_y - (P_{1y} - P_{0y}) N_x}{\sqrt{N_x^2 + N_y^2}}\,. \tag{4}$$

Plugging x, y from equation (2), N from equation (3), after simplification it is obtained a formula for distance from P_1 to the line:

$$\mathrm{d} = \left[P_{1x}\frac{\sin^{n-1}}{b^n} - P_{1y}\frac{\cos^{n-1}}{a^n} - \left(\frac{\cos^n t}{a^n} + \frac{\sin^n t}{b^n}\right)^{-1/n}\left(\frac{\cos t\, \sin^{n-1} t}{b^n} - \frac{\sin t\, \cos^{n-1} t}{a^n}\right)\right] \Big/ \sqrt{\frac{\cos^{2n-2} t}{a^{2n}} + \frac{\sin^{2n-2} t}{b^{2n}}}\,. \tag{5}$$

This is a transcendental function of one variable. Distance d is set equal to zero then a root-finding algorithm is applied.

Function properties

- Signs of components of P_1 are equal to signs of corresponding components of P_0. Parametrization (2) is defined only on interval $t \in [0, \pi/2]$ where $P_{1x} \geqslant 0, P_{1y} \geqslant 0$. Before computation, P_1 is moved to the quadrant where equation (2) is defined, and after P_0 is found its components are assigned original signs from P_1.
- The function is continuous. At the endpoints of the interval $[0, \pi/2]$ function values take different signs, thus function has root for every $P_{1x} > 0, P_{1y} > 0$.
- If P_1 is located inside the hyperellipse then it is possible for function (5) to have more than 1 root. On one of the intervals $t \in [0, \arctan(b/a)] \vee t \in [\arctan(b/a), \pi/2]$ it exists exactly 1 root and it corresponds to minimal distance between P_1 and P_0. Proof is given in appendix [].
- The function is continuously differentiable on the interval $t \in [0, \pi/2]$. Derivative:

$$
\begin{aligned}
\frac{d}{dt} \;=\;& \operatorname{len}^{-3}(n-1)\left(\frac{\sin t \cos^{2n-3} t}{a^{2n}} - \frac{\cos t \sin^{2n-3} t}{b^{2n}}\right)\left(\left(P_{1x} - r_0^{-1/n}\cos t\right)\frac{\sin^{n-1} t}{b^n} - \left(P_{1y} - r_0^{-1/n}\sin t\right)\frac{\cos^{n-1} t}{a^n}\right) \\
&+ \operatorname{len}^{-1}(n-1)\left(P_{1x}\cos t\,\frac{\sin^{n-2} t}{b^n} + P_{1y}\sin t\,\frac{\cos^{n-2} t}{a^n}\right) \\
&- r_0^{-1/n-1}\left(\left((n-1)\sin^2 t - 1\right)\frac{\cos^{2n-2} t}{a^{2n}} + \left((n-1)\cos^2 t - 1\right)\frac{\sin^{2n-2} t}{b^{2n}} + (n-1)(\cos^4 t + \sin^4 t)\frac{\cos^{n-2} t \sin^{n-2} t}{a^n b^n}\right), \text{ where} \\
r_0 \;=\;& \left(\frac{\cos^n t}{a^n} + \frac{\sin^n t}{b^n}\right), \quad \operatorname{len} = \sqrt{\frac{\cos^{2n-2} t}{a^{2n}} + \frac{\sin^{2n-2} t}{b^{2n}}}\;.
\end{aligned}
\tag{6}
$$

Application of root-finding algorithm

Utilized is Newton-Raphson method combined with bisection. It always returns the result if it is given an interval such that interval contains root and function values take different signs at interval endpoints. If on some Newton-Raphson iteration i any of the following conditions arise:

- Derivative equals zero;
- Next approximation x_i appears outside of the defined interval;
- Non-convergency or divergency $|f(x_i)| \geqslant |f(x_{i-1})|$ occurs;

then bisection iteration is performed. For that it performs 0 – 2 additional function evaluations to determine signs at interval endpoints.

Root-finding algorithm requires initial guess t_0 and bounding interval. First it computes dC – signed distance P_1 to the line defined by the “corner” point and normal vector in “corner” point. Then t_0 and bounding interval are found using the following formula:

$$
\begin{cases}
t_0 = P_{1y} \cdot \arctan(b/a) / (P_{1y} + |dC|), & t \in [0, \arctan(b/a)] \quad \text{if } P_1 \text{ is to the right of the line,} \\
t_0 = \arctan(b/a) + (\pi/2 - \arctan(b/a)) \cdot |dC| / (P_{1x} + |dC|), & t \in [\arctan(b/a), \pi/2] \quad \text{if other case.}
\end{cases}
$$

Termination criterion for root-finding algorithm

Algorithm terminates when it reaches given absolute error bound $\varepsilon \geqslant |f(x_i)|$. For formula (5) the meaning of that number is distance d. If distance P_1 to hyperellipse is large enough then given ε may be unreachable because of limited precision. Also for the application it would be more adequate if absolute error bound is expressed in angle t. To address the above issues, if distance P_1 to P_0 is greater than 1 then computed function value and derivative are divided by the distance P_1 to P_0.

Measurements

It was measured average number of function calls fnCalls required to reach absolute error $\varepsilon = 10^{-12}$ or below. Pseudorandom number generator was used. In every measurement 10^3 random hyperellipses were formed and for each one distance to 10^3 random points was computed.

Parameters	fnCalls
n=2..4, a,b=1..2, $\lvert P_{1x}\rvert$, $\lvert P_{1y}\rvert$ < 3	4.35
n=2..4, a,b=1..2, $\lvert P_{1x}\rvert$, $\lvert P_{1y}\rvert$ < 10^8	5.39
n=2..4, a=1, b=15..20, $\lvert P_{1x}\rvert$ < 2, $\lvert P_{1y}\rvert$ < 2b	7.07
n=2..4, a=1, b=15..20, $\lvert P_{1x}\rvert$, $\lvert P_{1y}\rvert$ < 10^8	7.18
n=10, a,b=1..2, $\lvert P_{1x}\rvert$, $\lvert P_{1y}\rvert$ < 3	5.49
n=10, a=1, b=15..20, $\lvert P_{1x}\rvert$ < 2, $\lvert P_{1y}\rvert$ < 2b	7.97
n=50, a,b=1..2, $\lvert P_{1x}\rvert$, $\lvert P_{1y}\rvert$ < 3	6.17