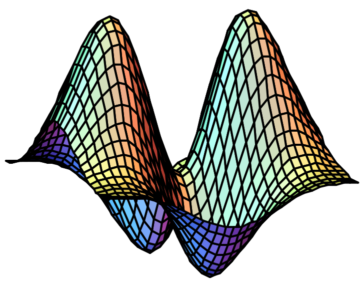
Critical Points of a Real-Valued Function of Two Variables

by Tom Davis[1]

At a critical or stationary point on a surface, the tangent plane is horizontal, analogous to a horizontal tangent line for critical points on a curve.



Click to Animate (click again to stop)

Animate this graph for $\theta = 0 \ldots 2\pi$ in steps of $\frac{1}{15}\pi$ for a total of 30 frames in a cycle at 10 frames/second.

Given a twice continuously differentiable or C^2, real-valued function z of a vector variable (x,y), denote the partial derivative of z with respect to x as z_x and the partial derivative of z with respect to y as z_y. Then the gradient of z is the vector grad z = (z_x,z_y).

Analogous to to the necessary condition that the first derivative vanish for critical points on a curve, a point (x_0,y_0) is a critical point if, and only if, grad $z(x_0,y_0) = (0,0)$.

A critical point is a local or relative extremum if it is either a local minimum or a local maximum. A critical point that is not a local extremum is called a saddlepoint, analogous to an inflection point on a curve.

Denote the partial derivatives of z_x and z_y as z_{xx}, $z_{xy} = z_{yx}$, and z_{yy}. Then the Hessian matrix[2] of z is the 2 x 2 symmetric matrix $H(z) = [h_{ij}(x,y)]$ where $h_{11} = z_{xx}$, $h_{12} = z_{xy} = h_{21}$, and $h_{22} = z_{yy}$. The Hessian determinant of z is the discriminant $D = |H| = h_{11} h_{22} - h_{12}^2$.

Analogous to the second derivative test for critical points on a curve, a critical point (x_0,y_0) is a

- strict or isolated local minimum if H is positive definite ($h_{11} > 0$, $D > 0$) at (x_0,y_0)
- strict local maximum if H is negative definite ($h_{11} < 0$, $D > 0$) at (x_0,y_0)
- saddlepoint if H is indefinite ($D < 0$) at (x_0,y_0)

If $D = 0$ at (x_0,y_0), the critical point is called degenerate, and its classification requires further analysis.

The example presented in this lecture is solved using analytical techniques. For a numerical example, see Newton's optimization method.

[1]Dept. of Civ. and Envir. Engrg., Univ. of South Florida, Tampa, FL 33620.

[2]After L. O. Hesse (1811-74), a student of C. G. J. Jacobi (1804-51).

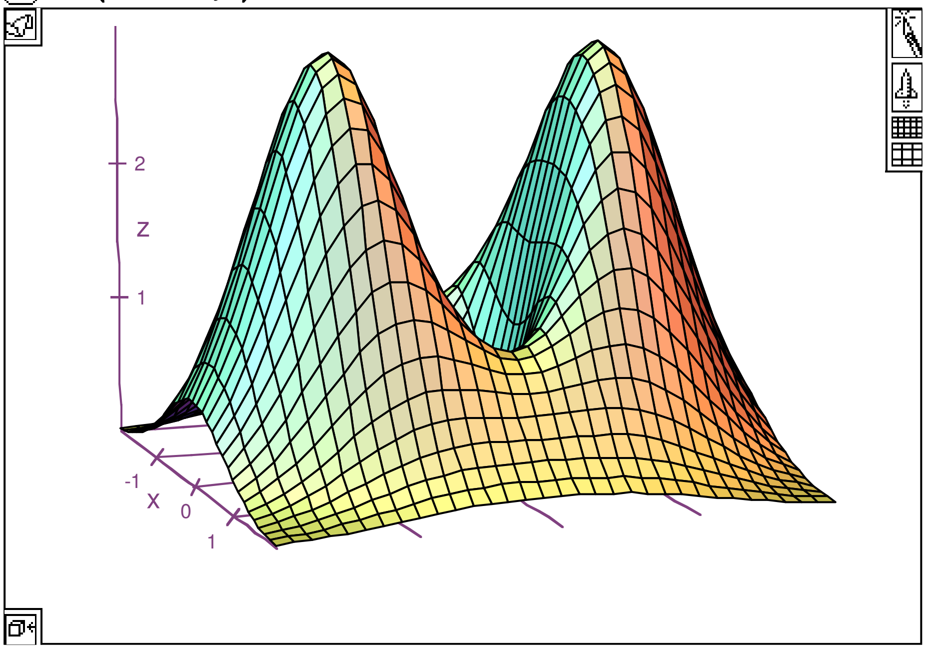
Real-Valued Function z(x, y)

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Function

$a = 1$ $b = 3$ $c = 1$ $d = 1$

$z = \left(a\,x^2 + b\,y^2\right) e^{1 - c\,x^2 - d\,y^2}$



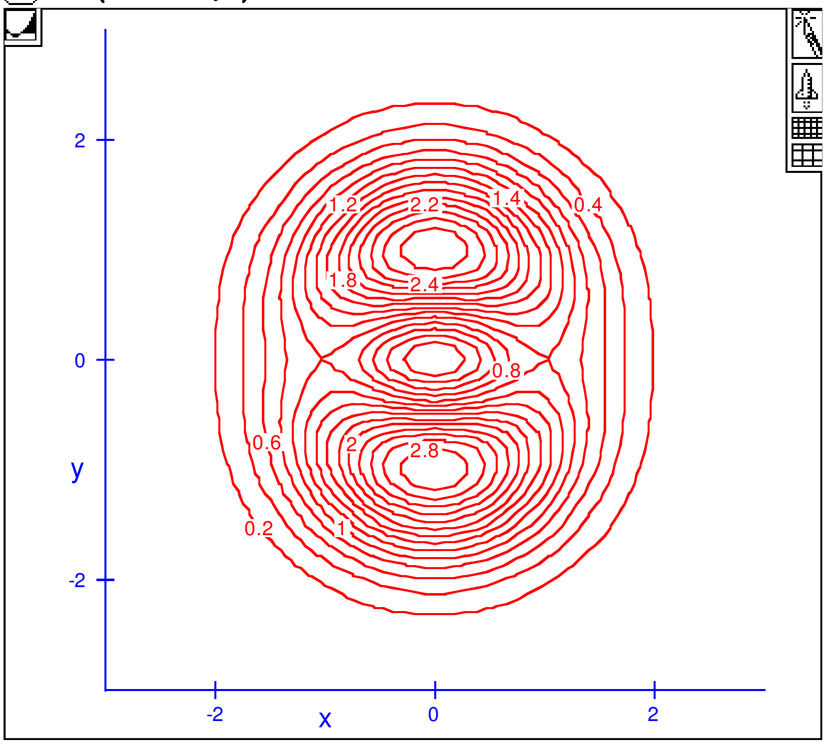
See contour plot for another view of z.
Explore: Rotate the graph and identify minima, maxima, and saddlepoints of z.
Experiment: Change the coefficient a =1 to a = -1. Comment on your observations.

Contours of z

Contours

$a = 1$ $b = 3$ $c = 1$ $d = 1$

$z = \left(a x^2 + b y^2\right) e^{1 - c x^2 - d y^2}$



Explore: Local extrema appear as "blank" spots among concentric contours, the tops of elevations or bottoms of depressions. No contour line may cross another. Saddlepoints are located where contour lines *appear* to intersect. Use the knife and rocket tools to investigate points of interest.
Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Partial Derivative z_x

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Derivative

$a = 1$ $b = 3$ $c = 1$ $d = 1$

The variables (x, y) are independent of each other.

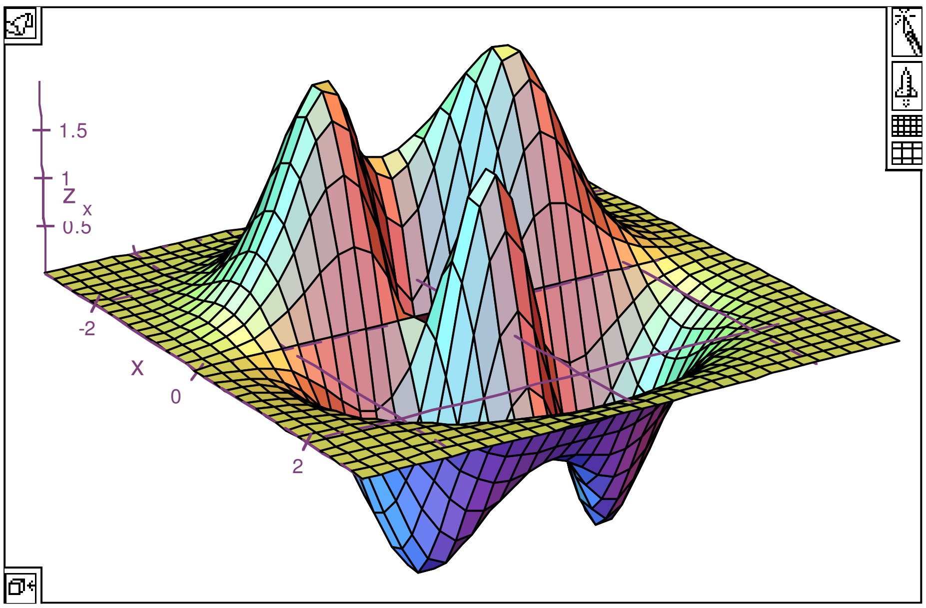
$$z = \left(a x^2 + b y^2\right) e^{1 - c x^2 - d y^2}$$

$$z_x = \frac{\partial}{\partial x} z$$

$$z_x = -2c\left(a x^2 + b y^2\right) e^{-c x^2 - d y^2 + 1} x + 2 a e^{-c x^2 - d y^2 + 1} x$$ Substitute

$$z_x = 2\left(-a c x^2 - b c y^2 + a\right) e^{-c x^2 - d y^2 + 1} x$$ Collect

$$z_x = 2 x\left(-a c x^2 - b c y^2 + a\right) e^{-c x^2 - d y^2 + 1}$$ Commute



See contour plot for another view of z_x.

Explore: Rotate the graph and identify the zeros of z_x.

Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

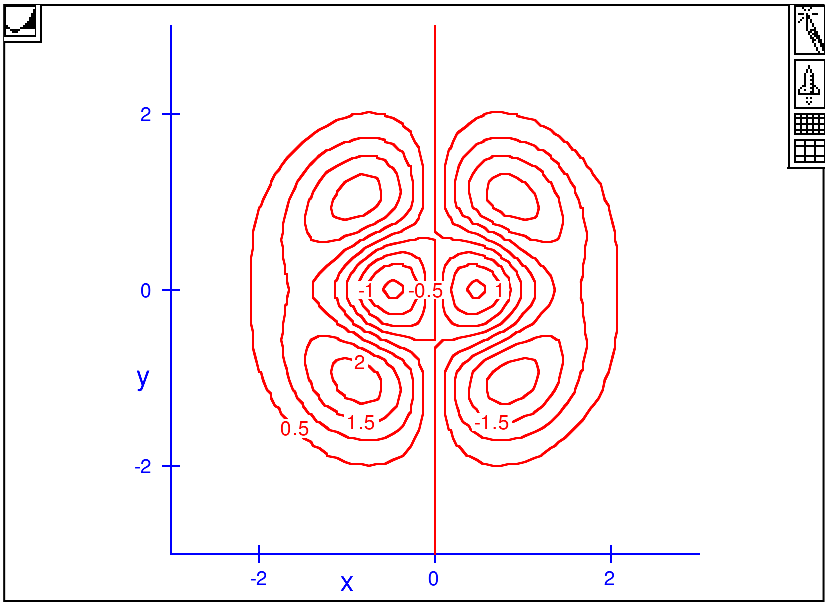
Contours of z_x

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Contours

$a = 1$ $b = 3$ $c = 1$ $d = 1$

$z_x = 2x(-acx^2 - bcy^2 + a)e^{-cx^2 - dy^2 + 1}$



$-3 \ldots 3$ = left...right True Proportions

$-3 \ldots 3$ = bottom...top cropping Moderately

Graph Declarations

Contours at (x, y) where x = left ... right and y = bottom ... top ; normally spaced contours of z_x computed with matching resolution, shown in normal line style, colored Red with labels

Explore: Click on the "normally" button in Graph Declarations and change to "at zero only".
Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Zeros of z_x

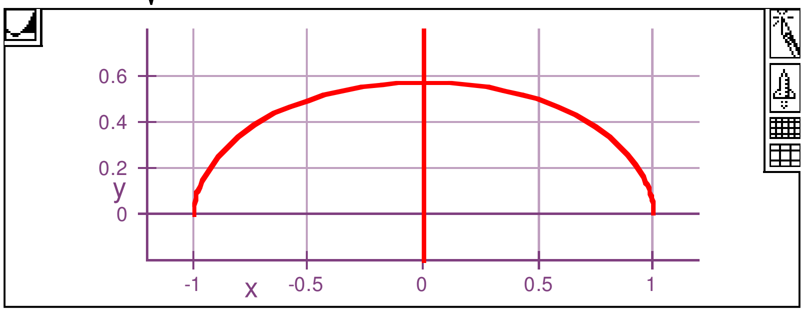
Zeros

$z_x = 0$ iff ($x = 0$) or ($- a c x^2 - b c y^2 + a = 0$)

$a = 1$ $b = 3$ $c = 1$ $d = 1$

$- a c x^2 - b c y^2 + a = 0$

$y = \sqrt{\frac{- a c x^2 + a}{b c}}$ Isolate



$y = -\sqrt{\frac{- a c x^2 + a}{b c}}$ Isolate

Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Partial Derivative z_y

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Derivative

$a = 1$ $b = 3$ $c = 1$ $d = 1$

The variables (x, y) are independent of each other.

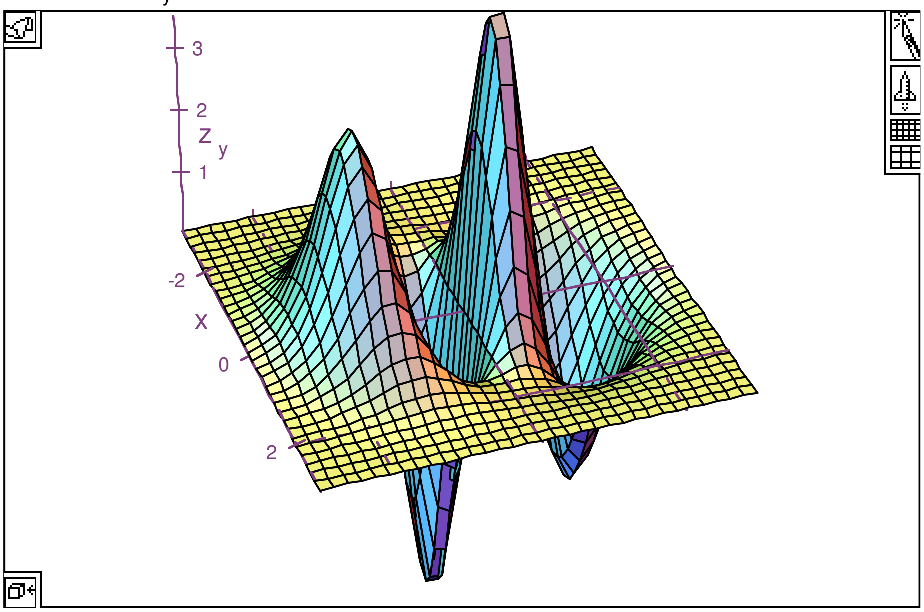
$z = (a x^2 + b y^2) e^{1 - c x^2 - d y^2}$

$z_y = \frac{\partial}{\partial y} z$

$z_y = -2d(a x^2 + b y^2) e^{-c x^2 - d y^2 + 1} y + 2b e^{-c x^2 - d y^2 + 1} y$ Substitute

$z_y = 2(-a d x^2 - b d y^2 + b) e^{-c x^2 - d y^2 + 1} y$ Collect

$z_y = 2y(-a d x^2 - b d y^2 + b) e^{-c x^2 - d y^2 + 1}$ Commute



See contour plot for another view of z_y.

Explore: Rotate the graph and identify the zeros of z_y.

Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

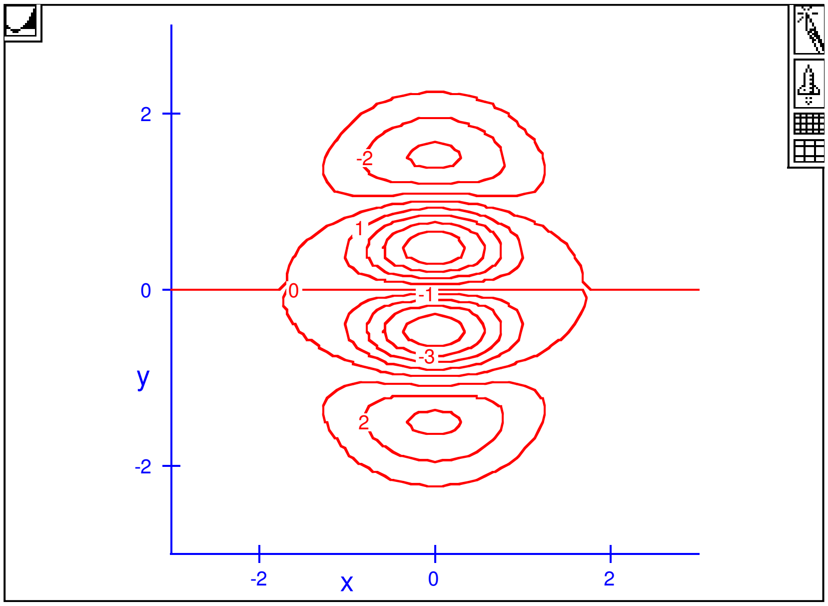
Contours of z_y

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Contours

$a = 1$ $b = 3$ $c = 1$ $d = 1$

$z_y = 2y\left(-a\,d\,x^2 - b\,d\,y^2 + b\right)e^{-c\,x^2 - d\,y^2 + 1}$



$-3 \ldots 3$ = left...right True Proportions

$-3 \ldots 3$ = bottom...top cropping Moderately

Graph Declarations

Contours at (x, y) where x = left ... right and
y = bottom ... top ; normally spaced contours of z_y
computed with matching resolution, shown in normal
line style, colored Red with labels

Explore: Click on the "normally" button in Graph Declarations and change to "at zero only".
Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Zeros of z_y

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

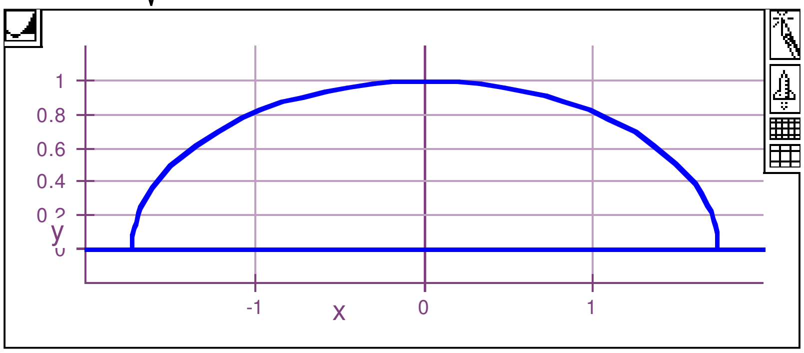
Zeros

$z_y = 0$ iff ($y = 0$) or ($- a d x^2 - b d y^2 + b = 0$)

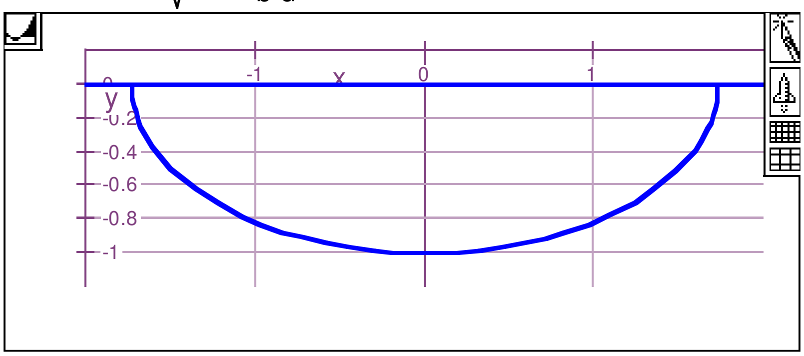
$a = 1$ $b = 3$ $c = 1$ $d = 1$

$- a d x^2 - b d y^2 + b = 0$

$$y = \sqrt{\frac{- a d x^2 + b}{b d}}$$ Isolate



$$y = -\sqrt{\frac{- a d x^2 + b}{b d}}$$ Isolate



Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Critical Points ($z_x = z_y = 0$)

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Solution

(x, y) is a critical point iff [($x = 0$) or ($-acx^2 - bcy^2 + a = 0$)]

and [($y = 0$) or ($-adx^2 - bdy^2 + b = 0$)]

$a = 1$ $b = 3$ $c = 1$ $d = 1$

$-acx^2 - bcy^2 + a = 0$

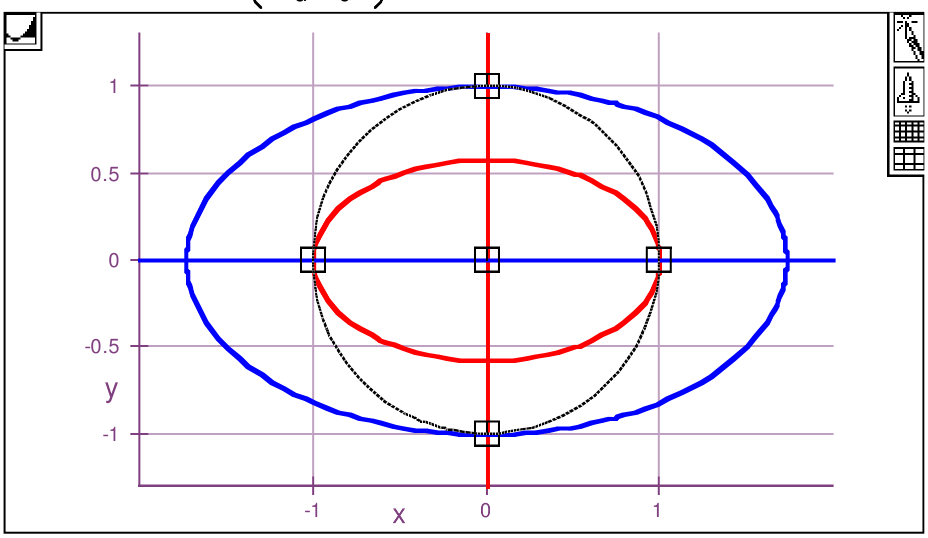
$u = \sqrt{\frac{1}{c}}$ $v = \sqrt{\frac{a}{bc}}$ x and y intercepts

$-adx^2 - bdy^2 + b = 0$

$\alpha = \sqrt{\frac{b}{ad}}$ $\beta = \sqrt{\frac{1}{d}}$ x and y intercepts

$$\text{CriticalPoints} = \begin{pmatrix} 0 & 0 \\ 0 & \beta \\ 0 & -\beta \\ u & 0 \\ -u & 0 \end{pmatrix} \qquad \text{CriticalPoints} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 0 \\ -1 & 0 \end{pmatrix}$$

Calculate



Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Hessian Matrix H and Determinant D

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Hessian

$a = 1$ $b = 3$ $c = 1$ $d = 1$

The variables (x, y) are independent of each other.

$z = (a\,x^2 + b\,y^2)\,e^{1 - c\,x^2 - d\,y^2}$

$h_{11}(x, y) = \left(\frac{\partial}{\partial x}\right)^2 z$

$h_{11}(x, y) = 2\left(a\left[2c^2x^4 - 5c\,x^2 + 1\right] - b\,c\,y^2 + 2b\,c^2x^2y^2\right)e^{-c\,x^2 - d\,y^2 + 1}$ Collect

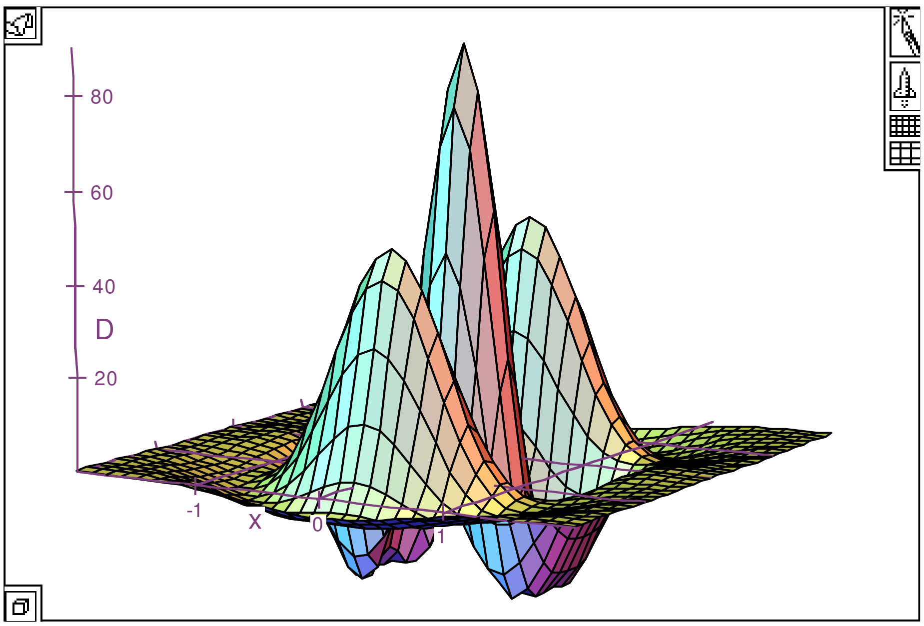
$h_{22}(x, y) = \left(\frac{\partial}{\partial y}\right)^2 z$

$h_{22}(x, y) = 2\left(b\left[2d^2y^4 - 5d\,y^2 + 1\right] - a\,d\,x^2 + 2a\,d^2x^2y^2\right)e^{-c\,x^2 - d\,y^2 + 1}$ Collect

$h_{12}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x} z\right)$

$h_{12}(x, y) = 4\left(a\,c\,d\,x^2 + b\,c\,d\,y^2 - b\,c - a\,d\right)e^{-c\,x^2 - d\,y^2 + 1}\,x\,y$ Collect

$H(x, y) = \begin{pmatrix} h_{11}[x, y] & h_{12}[x, y] \\ h_{12}[x, y] & h_{22}[x, y] \end{pmatrix}$ $D = \left|H(x, y)\right|$



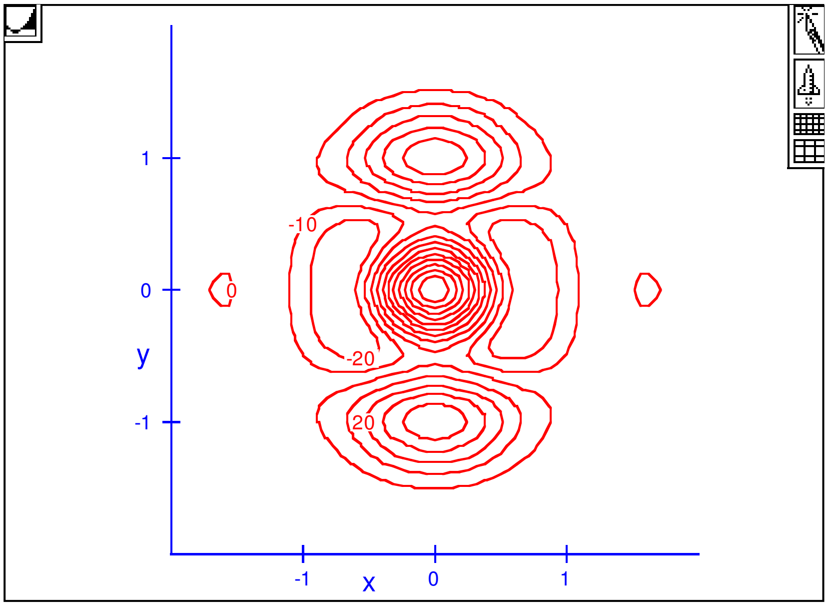
See contour plot for another view of D.
Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Contours of D

Contours

a = 1 b = 3 c = 1 d = 1

$D = |H(x, y)|$



Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Classification of Critical Points

LIVEMATH Maker PLUG-IN INPUT MATH HELP FILE

Classification

$a = 1$ $b = 3$ $c = 1$ $d = 1$ $u = \sqrt{\frac{1}{c}}$ $\beta = \sqrt{\frac{1}{d}}$

$h_{11}(x, y) = 2\left(a\left[2c^2x^4 - 5cx^2 + 1\right] - bcy^2 + 2bc^2x^2y^2\right)e^{-cx^2 - dy^2 + 1}$

$h_{22}(x, y) = 2\left(b\left[2d^2y^4 - 5dy^2 + 1\right] - adx^2 + 2ad^2x^2y^2\right)e^{-cx^2 - dy^2 + 1}$

$h_{12}(x, y) = 4\left(acdx^2 + bcdy^2 - bc - ad\right)e^{-cx^2 - dy^2 + 1}xy$

$H(x, y) = \begin{pmatrix} h_{11}[x, y] & h_{12}[x, y] \\ h_{12}[x, y] & h_{22}[x, y] \end{pmatrix}$ $D(x, y) = |H(x, y)|$

$H(0,0)$ $D(0,0)$ $D(0,0) = 88.669$ Calculate

$H(0,0) = \begin{pmatrix} 5.4366 & 0 \\ 0 & 16.31 \end{pmatrix}$ Calculate $H(0,0) = \begin{pmatrix} 2e & 0 \\ 0 & 6e \end{pmatrix}$ UnCalculate

$Class(0,0)$ $Class(0,0) =$ StrictLocalMinimum Calculate

$H(0, \beta)$ $D(0, \beta)$ $D(0, \beta) = 48$ Calculate

$H(0, \beta) = \begin{pmatrix} -4 & 0 \\ 0 & -12 \end{pmatrix}$ Calculate

$Class(0, \beta)$ $Class(0, 1) =$ StrictLocalMaximum Calculate

$H(0, -\beta)$ $D(0, -\beta)$ $D(0, -\beta) = 48$ Calculate

$H(0, -\beta) = \begin{pmatrix} -4 & 0 \\ 0 & -12 \end{pmatrix}$ Calculate

$Class(0, -\beta)$ $Class(0, -1) =$ StrictLocalMaximum Calculate

$H(u, 0)$ $D(u, 0)$ $D(u, 0) = -16$ Calculate

$H(u, 0) = \begin{pmatrix} -4 & 0 \\ 0 & 4 \end{pmatrix}$ Calculate

$Class(u, 0)$ $Class(1, 0) =$ Saddlepoint Calculate

$H(-u, 0)$ $D(-u, 0)$ $D(-u, 0) = -16$ Calculate

$H(-u, 0) = \begin{pmatrix} -4 & 0 \\ 0 & 4 \end{pmatrix}$ Calculate

$Class(-u, 0)$ $Class(-1, 0) =$ Saddlepoint Calculate

Experiment: Change the coefficient a = 1 to a = -1. Comment on your observations.

Appendix. Glossary

critical, or stationary, point

A point at which a function attains a critical value, either a local extremum or an n-dimensional saddlepoint. The term *stationary point* may be visualized by placing a perfectly balanced ball on a smooth curve or surface: at a stationary point the ball is in static equilibrium and will not roll.

C^2

Twice continuously differentiable. A scalar-valued function of a vector variable is said to be of class C^1 when the first partial derivatives of the function exist and are continuous. If these derivatives, in turn, have continuous partials, the function is said to be of class C^2. A related notion is that of *geometric continuity*: a function with continuous tangent is of class G^1, and a function with continuous curvature is of class G^2.

degenerate critical point

A critical point at which the Hessian determinant is zero. The behavior of a function near such points might be tested graphically with level sets (contours) and sections or by some other method.

discriminant

A relatively simple expression that determines some of the properties, as the nature of the roots, of a given equation or function.

gradient

The first derivative of a scalar-valued function of a vector variable.

Hessian determinant

The determinant of the Hessian matrix.

Hessian matrix

The second derivative of a scalar-valued function of a vector variable. The Hessian matrix is the Jacobian matrix of the gradient.

indefinite matrix

Let x be an n-element column vector and x^T its transpose. Then an n x n symmetric matrix M is indefinite if the quadratic form $x^T M x < 0$ for some x and > 0 for other x.

Jacobian matrix

The first derivative of a vector-valued function of a vector variable.

local maximum

A point at which a function attains a value that is a maximum relative to other points in the neighborhood. If the value at this point is the largest value of the function, the point is a *global*, or *absolute*, *maximum*, and the function is said to be *concave*.

local minimum

A point at which a function attains a value that is a minimum relative to other points in the neighborhood. If the value at this point is the smallest value of the function, the point is a *global*, or *absolute*, *minimum*, and the function is said to be *convex*.

local, or relative, extremum

A point at which a function attains a value that is a maximum or minimum relative to other points in the neighborhood. If the value at this point is the largest or smallest value of the function, the point is a *global*, or *absolute*, *extremum*. Also called *turning point*.

negative definite matrix

Let x be an n-element column vector and x^T its transpose. Then an n x n symmetric matrix M is negative definite if for all nonzero x, the quadratic form $x^T M x < 0$.

negative semidefinite matrix

Let x be an n-element column vector and x^T its transpose. Then an n x n symmetric matrix M is negative semidefinite if for all nonzero x, the quadratic form $x^T M x \leq 0$.

positive definite matrix

Let x be an n-element column vector and x^T its transpose. Then an n x n symmetric matrix M is positive definite if for all nonzero x, the quadratic form $x^T M x > 0$.

positive semidefinite matrix

Let x be an n-element column vector and x^T its transpose. Then an n x n symmetric matrix M is positive semidefinite if for all nonzero x, the quadratic form $x^T M x \geq 0$.

saddlepoint

A critical point that is not a local extremum. At an n-dimensional saddlepoint, the tangent hyperplane is horizontal, but the value of the function is neither a local minimum nor a local maximum.

strict, or isolated, local extremum

A local extremum at which the function value is strictly less than (for a strict local minimum) or strictly greater than (for a strict local maximum) the function value of nearby points. In 3-space, local extrema that are not isolated exhibit as level lines or contours.

Appendix. References

Buck, R. C. (1978). *Advanced calculus*. McGraw-Hill, New York.

Daintith, J., and Nelson, R. D. (1989). *The Penguin dictionary of mathematics*. Penguin, London, England.

Hildebrand, F. B. (1963). *Advanced calculus for applications*. Prentis-Hall, Englewood Cliffs, N.J.

Isaacson, E., and Keller, H. B. (1966). *Analysis of numerical methods*. Wiley, New York.

Marsden, J. E., and Tromba, A. J. (1981). *Vector calculus*. Freeman, San Francisco.

O'Connor, J. J., and Robertson, E. F. (2000). *MacTutor history of mathematics*. <http://www-history.mcs.st-andrews.ac.uk/history/>

Reklaitis, G. V., Ravindran, A., and Ragsdell, K. M. (1983). *Engineering optimization*. Wiley, New York.

Stewart, G. W. (1973). *Introduction to matrix computations*. Academic, New York.