

# Interactive Geometry and Critical Points

*Thomas F. Banchoff*

[Thomas\\_Banchoff@brown.edu](mailto:Thomas_Banchoff@brown.edu)

Mathematics Department

Brown University

Providence, RI 02912

U.S.A.

**Abstract:** Interactive geometry programs make it possible to treat a broad subject like critical points of functions in a unified way, across a wide range of courses at different levels. In this article, we show how color-coding of graphs of families of functions of one and more variables makes it possible to convey geometric information that is usually missing from introductory presentations of this material.

## Introduction

How does an interactive geometry program make it possible to treat a broad subject like critical points of functions in a unified way, across a wide range of courses at different levels? In this article, we show how color-coding of graphs of families of functions of one and more variables makes it possible to convey geometric information that is usually missing from introductory presentations of this material. We introduce progressive elaborations of the fundamental ideas of critical point theory to motivate and illustrate more and more complicated geometric phenomena.

## 1. Graphs of Functions of a Single Variable.

By the simple device of coloring red all segments of positive slope in a polygonal approximation of a function, we can immediately see where the function has local maxima and minima by observing the color changes that occur as we go from the left-hand endpoint of an interval to the right-hand endpoint. Specifically if the end segments have the same color, the number of critical points in the interval is even and if the end segments have opposite colors, the number of critical points in the interval is odd.

For the graph of a differentiable function of one variable  $f(x)$ , this means that the number of points  $c$  where  $f'(c) = 0$  in the interval  $[a,b]$  is even if  $f'(a)$  and  $f'(b)$  have the same sign and this number is odd if the signs are different.

Since the normal line is perpendicular to the tangent line to the function graph, we color the graph red near any point with upward unit normal having angle lying in the quarter-circle from  $\pi/2$  to  $\pi$ .

How does interactive graphics change the way we deal with this kind of presentation? With interactive graphics, we can carry out an online investigation of a family of functions depending on one or more parameters.

As examples of such families, we can see what happens as we change one coefficient of a polynomial. For linear functions,  $f(x) = mx + b$ , changing  $m$  will just rotate the line about its  $y$ -intercept, and changing  $b$  will alter the  $y$ -intercept while preserving the slope. For quadratic functions  $f(x) = ax^2 + mx + b$ , transforming the constant term moves the graph vertically, and

changing the coefficient of the squared term from positive to negative will change the graph from a parabola that opens upward to a parabola that opens downward. The effect of changing the coefficient of the linear term is more complicated, and it is a good exercise for students to experiment and to record their observations. When we keep  $a$  and  $b$  fixed and change  $m$ , the basic shape of the curve will not be altered, but the position of the critical point will change.

## 2. The Family of Cubic Equations

What about a cubic equation  $f(x) = x^3 + ax^2 + mx + b$ ? Once again, changing the constant term  $b$  only alters the  $y$ -intercept. In this case, changing the coefficient  $a$  will not alter the shape of the graph, but changing the slope  $m$  makes a big difference. Students can investigate the functions  $f(x) = x^3 + mx$  and observe the numbers of critical points of the graph. It is immediate to conjecture that when  $m$  is positive, there are no critical points for the function and when  $m$  is negative, the number of critical points is two. This conjecture can then be proven by showing that the derivative  $f'(x) = 3x^2 + m$  will never be zero if  $m$  is positive and that there are two values of  $x$  for which  $f'(x) = 0$  when  $m$  is negative. The intermediate case, with  $m = 0$ , has one “degenerate” critical point at the origin, where  $f'(0) = 0$  so the tangent line is horizontal.

We can color the points on the graph by a darker color when the slope of the tangent line is negative, so we have a visual record of the places where the slope changes from positive to negative (the local maxima of the curve) and where it changes from negative to positive (at the local minima).

Note that  $f''(x) = 6x$  no matter what  $m$  is, so the concavity of the graph changes from convex downward to convex upward as  $x$  goes from negative values to positive values in the domain. By making the graph thick when the second derivative is positive and thin where it is negative, we can observe the inflections points of the graph, where the second derivative changes sign and the concavity of the graph changes.

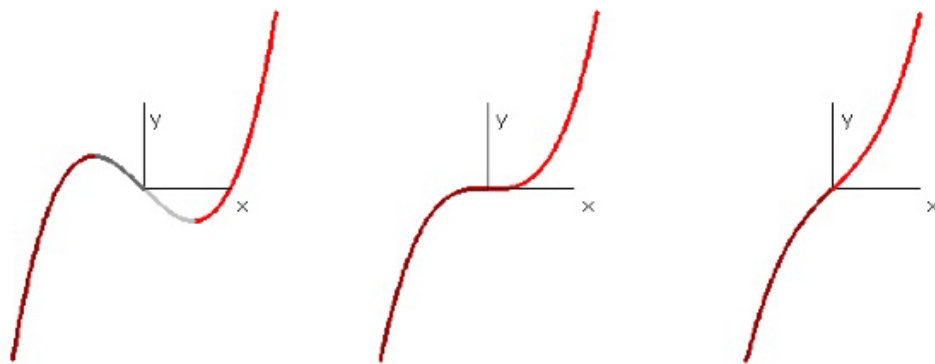


Figure 1: The family of cubic equations

## 3. The Family of Quartic Equations

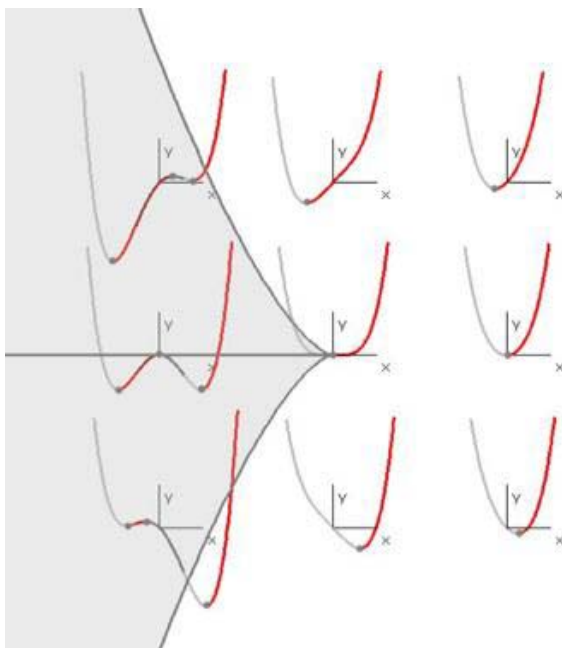
More interesting is the case of a quartic, or fourth degree polynomial  $f(x) = x^4 + cx^3 + ax^2 + mx + b$ . As in the case of the quadratic and the cubic, changing  $b$  only shifts the  $y$ -intercept up and

down without changing the shape of the curve. Also, changing the coefficient of the second term only moves the graph of the function from side to side without changing the shape of the graph, and in particular the number of critical points of the function. We are left with two coefficients to consider,  $f(x) = x^4 + ax^2 + mx$ , and changing the coefficients  $a$  and  $m$  can result in very different kinds of graphs.

Once again, we can invite students to explore the “two-dimensional control space” by choosing various values of  $a$  and  $m$  and recording the shapes of the graphs that result. Sometimes the number of critical points is 1, for example when  $m = 0$  and  $a$  is positive, and other times there are 3 critical points, for example when  $m = 0$  and  $a$  is negative. What happens if  $m$  is not zero? When will we get 3 critical points and when will we get 1? Will we ever get exactly 2 critical points? If  $m = 0$ , then we observe that for any negative number  $a$ , the graph will have three critical points, two minima and one local maxima. We can show this algebraically by computing  $f'(x) = 4x^3 + 2ax = 2x(2x^2+a)$ , which will equal zero for at  $x = 0$  and at two other values of  $x$  if  $a$  is negative, and only at  $x = 0$  if  $a$  is positive.

If we choose  $a = -1$  so that  $f(x) = x^4 - x^2 + mx$ , then for some values of  $m$  near zero, the number of critical points is still 3, but for some value of  $m$ , the number of critical points changes to 1. We can observe that this change occurs at the value of  $m$  for which the function has a horizontal inflection point, where  $f'(x) = 0$  and  $f''(x) = 0$ . Thus we must have both  $f'(x) = 4x^3 - 2x + m = 0$  and  $f''(x) = 12x^2 - 2 = 0$ . From the second equation,  $x$  is plus or minus the square root of  $1/6$ , so  $m = \pm(4/3)\sqrt{1/6}$ . For general  $a$ , we obtain the equations  $f'(x) = 4x^3 - 2ax + m = 0$  and  $f''(x) = 12x^2 + 2a = 0$  so  $m = \sqrt{-1/6}$  and  $27m^2 = -8a^3$ . This curve in the “control space” of all possible choices of  $m$  and  $a$  will separate the functions with 3 critical points from those with 1 critical point.

We can use the interactive graphics system to construct a collection of function graphs that exhibit the different configurations of critical points for different choices of  $a$  and  $m$ ,



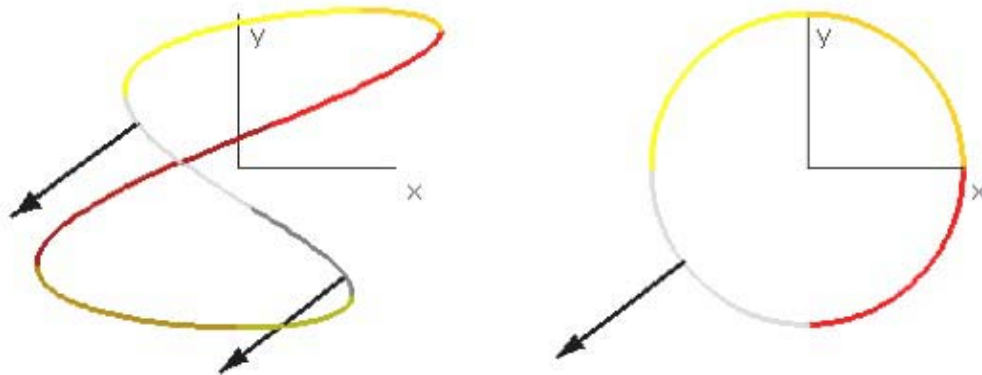
**Figure 2:** the family of quartic equations

Mathematicians familiar with catastrophe theory will recognize that these two families of functions, the cubic curves and the quartic curves, are the most important starting examples in that powerful theory. Originally arising in the sciences of optics and structural design, catastrophe theory has significant applications not only in physics and engineering, but also in biological and social sciences as well as in differential geometry and other parts of mathematics. It is noteworthy that interactive computer graphics makes it possible to introduce this subject to students just beginning the study of calculus and analytic geometry for functions of one variable.

#### 4. Critical Points for Parametric Curves

There are two natural ways to extend this construction. Instead of considering only the graphs of functions, we may consider parametric curves in the plane, and then we can consider graphs of functions of two variables, in three-dimensional space, and eventually parametric surfaces.

For a closed curve in the plane given in parametric form by  $(x(t), y(t))$  with  $t$  going from  $a$  to  $b$ , we color each segment according to the position of its normal vector: red for the second quadrant, orange for the third and yellow for the fourth. We can then read off the critical points of the horizontal coordinate function  $x(t)$  by finding points where the color changes from yellow to white or from red to orange, or conversely. The critical points of other coordinate function  $y(t)$  on the other hand occur where the color changes from white to red or from orange to yellow or conversely. The number of color changes for each of the two coordinate functions will be even and the total number will be even. This is true whether or not the curve intersects itself.



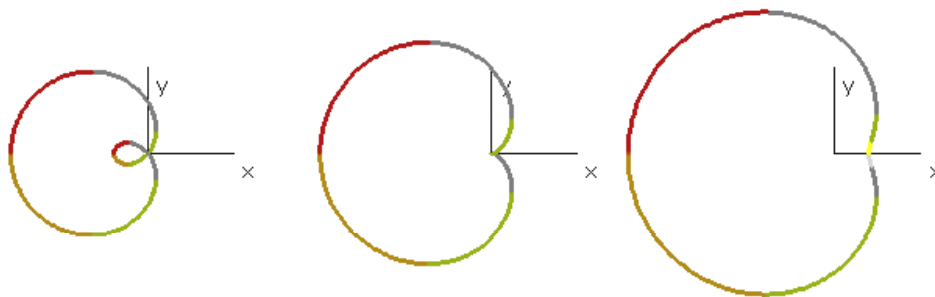
**Figure 3:** Color-labeled parametric curves in the plane

Using an interactive graphics program, we can display a parametric curve in the plane and compare the segments of the graph to the segments of a counterclockwise circle. If the curve is given by a pair of differentiable functions  $x(t)$  and  $y(t)$ , then the velocity vector at each point has coordinates  $(x'(t), y'(t))$  and rotating this vector by one-quarter turn in the counterclockwise direction gives the normal vector  $(-y'(t), x'(t))$ . The color of the curve is given by the color of the point on the circle where the outward normal vector points in the same direction as  $(-y'(t), x'(t))$ .

As the point  $(x(t), y(t))$  moves along the curve, the corresponding unit normal vector on the unit circle is either moving counterclockwise or clockwise, and we can indicate this by making the curve thin in the first case and thick in the second case. The points of the curve where the direction changes are the inflection points of the curve. Students can investigate various curves, and observe the number of points on the curve that correspond to any fixed point on the circle. How does the number of corresponding points change as we move the position on the circle? What happens as we pass a point on the circle that corresponds to an inflection point on the curve? The investigation of curves in the plane leads to the theory of tangential degree of a curve, an important concept in the differential geometry of curves in the plane.

## 5. Families of Parametric Curves

Just as we considered one- and two-parameter families of function graphs, we can explore families of closed parametric plane curves such as the cardioid family  $((-u+\cos(t))(-\cos(t)), (u+\cos(t))\sin(t))$ . Students can discover that the shapes of these curves depend on  $u$  and in particular, that the shape changes dramatically near  $u = 1$  and  $-1$ . As we watch an animation of this family as  $u$  runs from  $u = .8$  to  $u = 1.2$ , we can observe that the curve has a loop for  $u$  less than 1, then a cusp at  $u = 1$  and then a pair of inflection points. Students can also record what happens near other significant point, when  $u = -2, -1, 0, 1,$  and  $2$ . Similar behavior will occur for other families of curves in the plane, a subject of importance in differential geometry and in the topology of plane curves.



**Figure 4:** The cardioid family, color-coded

## 6. Critical Points and Functions of Two Variables

We can generalize the study of curves that are graphs of functions of one variable in the plane to the study of surfaces that are the graphs of functions of two variables in three-space. As in the case of functions of one variable, we want to find the range of a function, namely the smallest cylindrical box over the domain that will contain all of the function values for points in the domain. The top and bottom planes of such a cylinder will be horizontal tangent planes to the function graph in the case where the function is differentiable, or it might be that the highest or lowest point occurs at a boundary point of the domain of definition, and in that case the tangent line to the boundary curve will be horizontal at the local or global maxima or minima.

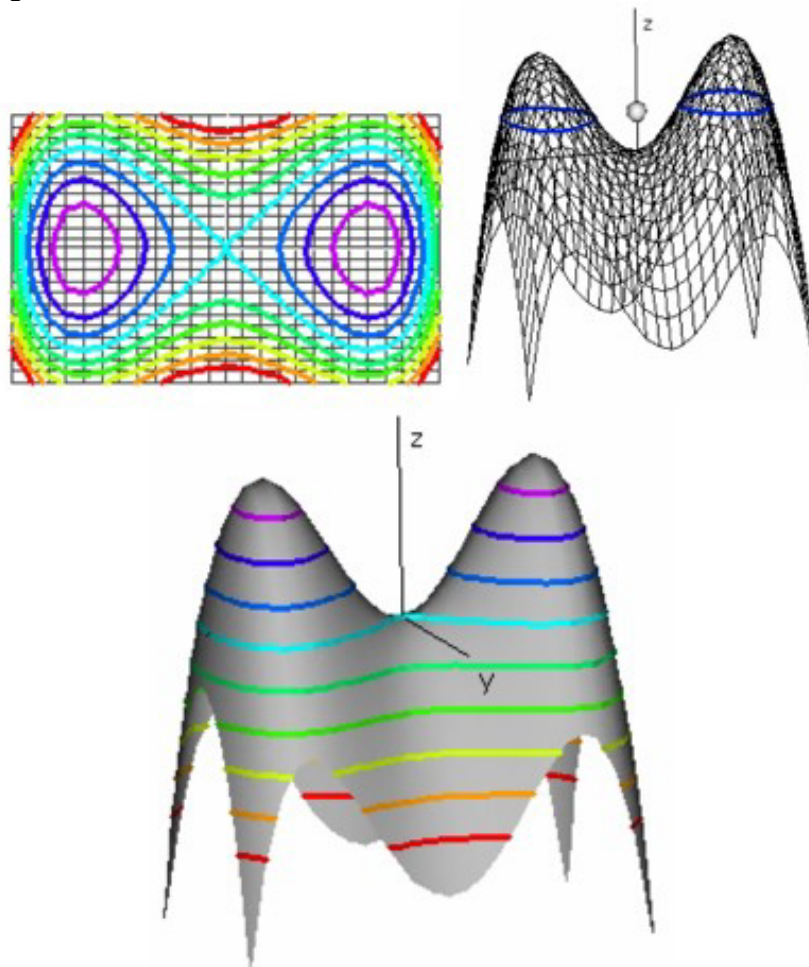
In addition to the global and local maxima and minima of functions of two variables, there are other points where the tangent plane is horizontal, for example the origin in the saddle-shaped graph of  $f(x,y) = x^2 - y^2$ , defined either over a square domain or a circular disc domain centered at the origin. This is a typical "saddle point" for the graph of the function. Finding the critical points

of a function includes finding all local maxima, local minima, and saddle points. An important theorem of Marston Morse states that, for almost all functions, the only critical points of a function of two variables are of this form.

Two particular examples give a good idea of the critical point behavior for functions of two variables. Both are named for geographical features: Twin Peaks and Crater Lake. Both of these surfaces have been described in the author's Scientific American Library volume "Beyond the Third Dimension" [1].

## 7. Twin Peaks the Geometry of Peaks and Passes

Twin Peaks is the graph of the function  $f(x,y) = -x^4 + 2x^2 - y^2$  over the domain  $-1.5 \leq x \leq 1.5$ ,  $-1 < y < 1$ . It is straightforward to show algebraically that  $f(x,y) \leq 1$  and that equality occurs at two "peaks"  $(1,0)$  and  $(-1,0)$ . Between the two peaks there is a saddle point at  $(0,0)$ , and there are no other critical points for the function. The level set for  $z = 1$  consists of the two points  $(1,0)$  and  $(-1,0)$ . For  $0 < z < 1$ , the level set consists of two closed curves. The level set for  $z = 0$  is a self-intersecting "figure eight" curve, and for  $z < 0$ , the level set the portion of a single closed curve inside the rectangular domain.



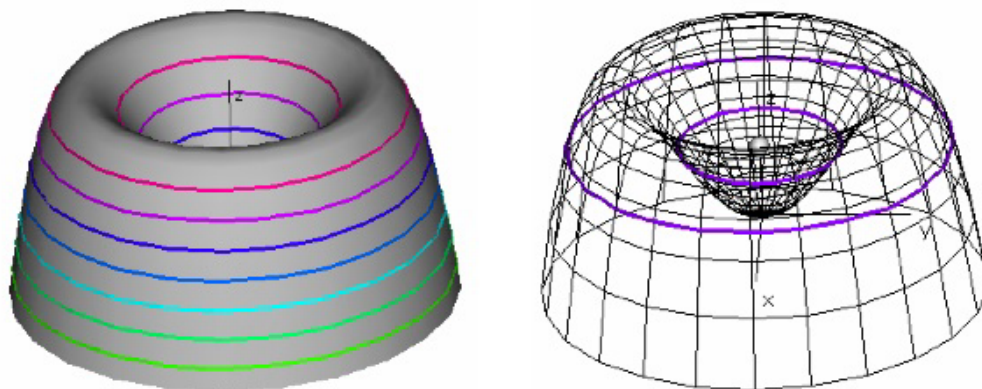
**Figure 5:** Graph and slices of Twin Peaks

This kind of "level set analysis" is crucial in an interactive graphics approach to graphs of functions of one or more variables. Students can investigate function graphs and report their observations on worksheets or by online messages. They can make conjectures about the reasons various configurations arise in graphs of certain kinds of equations, and then go on to find algebraic reasons for these observed phenomena.

For Twin Peaks, we can observe that the fact that there are two different global maxima is a sort of accident that can be removed by a slight perturbation. Just as adding a small linear term to the function  $f(x) = -x + 2x^2$  to get  $f(x) = -x^4 + 2x^2 + mx$  changed the heights of the global maxima, the same thing happens here. If there is an "earthquake" that shears the configuration by  $mx$ , we have the graph of  $f(x,y) = -x^4 + 2x^2 + mx - y^2$ , with two peaks at different heights, at least if  $m$  is small enough in absolute value. For an extreme earthquake, the number of peaks goes from 2 to 1 and the saddle point disappears. Students can determine by observation where that crucial changeover occurs.

## 8. Crater Lake and the Geometry of Pits and Passes

Crater Lake exhibits a different kind of unstable situation. The graph of  $g(x,y) = -(x^2 + y^2)^2 + 2(x^2 + y^2)$  is a surface of revolution about the  $z$ -axis with profile curve  $-x^4 + 2x^2$ . The maximum value of this function is 1, and there are infinitely many global maxima, lying over the circle  $x^2 + y^2 = 1$  in the domain. There is one local minimum at the origin. Once again an earthquake shear will perturb this situation to give the graph of  $g(x,y) = -(x^2 + y^2)^2 + 2(x^2 + y^2) + mx$ , which will still have one local minimum, but the circle of maxima will break up into one global maximum and one saddle point. The level set at the local minimum will be an isolated point surrounded by a single level curve. The level set at the global maximum will be a single point, and the level set at the saddle point will be a "double loop with a single crossing point". Near the saddle point, the level set looks like a pair of intersecting lines, as in the case of the figure-eight curve for Twin Peaks. Between the level of the maximum and the saddle, the level set is a single curve. Between the saddle and the local minimum, the level set consists of two curves, one inside the other (as opposed to next to each other as in the case of the figure-eight). Below the local minimum, the level set is again the part of a single curve lying in the domain.



**Figure 6:** Graph and slices of Crater Lake

This description of the critical points configuration assumes that the earthquake has not been too severe. If the number  $m$  becomes large enough then we obtain a function graph with one global maximum and no other critical points. The local minimum and the saddle point have come together and disappeared. The water in Crater Lake has spilled out. Students can find this point by observation (and compare the value with the  $m$  for the crucial earthquake in the case of Twin Peaks).

## 9. The Critical Point Theorem for Graphs of Functions of Two Variables

Both of these topographical examples suggest that it will be possible to obtain other configurations of critical points, say with  $n+1$  local maxima and  $n$  saddle points in between, then  $m$  local minima at the bottom of craters, each with one saddle point. This will produce a function with  $n+1$  maxima,  $n + m$  saddles and  $m$  minima, so in particular the number of maxima plus the number of minima is one greater than the number of saddles, and the total number of critical points will be an odd number.

We can collect these observations in a conjecture: For an "island" with one level shoreline curve and everything on the island above sea level, there will be at least one global maximum, and if all critical points are local maxima, local minima, and ordinary saddles, then the number of critical points is odd and moreover  $\# \text{maxima} - \# \text{saddles} + \# \text{minima} = 1$ . This conjecture epitomizes fundamental results in Critical Point Theory, of immense importance in global geometry and analysis over the last 80 years. It is also of fundamental importance for the geometry of surfaces, leading to a modern proof of the crucial Gauss-Bonnet Theorem as well as properties of knotted "strings" in contemporary molecular biology and theoretical physics.

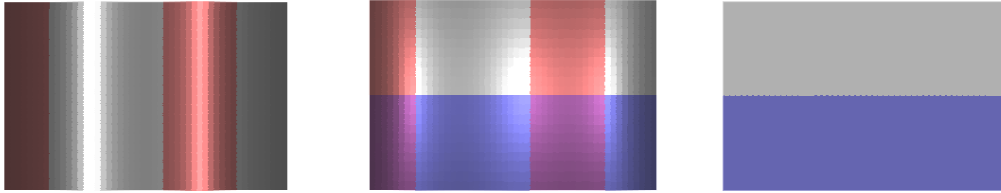
The mathematician who popularized critical point theory was Marston Morse, who developed the subject eighty years ago. He enjoyed giving popular lectures on the topic for students of all levels, and two of his previously unpublished presentations have just been printed in the November 2007 issue of *The American Mathematical Monthly* [2]. On a visit to Providence RI in the 1970's, he came to the computer graphics laboratory at Brown University where my computer scientist colleague Charles Strauss and I were developing a program for analyzing the geometry of surfaces in three and four dimensions. Although the graphics programs were primitive by today's standards, and very slow, Marston Morse immediately appreciated the potential of such approaches for teaching, research, and public exposition of geometric ideas arising in critical point theory.

## 10. Color-Coding Graphs of Functions and Partial Derivatives

How does interactive computer graphics help to illustrate the geometry of critical points for graphs of functions of two variables, or families of such functions? For a differentiable function there is a well-defined tangent plane at each point of the graph and an upward-pointing normal perpendicular to that plane. If  $f_x(x,y)$  and  $f_y(x,y)$  denote the first partial derivatives of  $f$  with respect to  $x$  and  $y$  respectively, then  $(-f_x(x,y), -f_y(x,y), 1)$  will be a normal vector. We color the surface white if  $f_x(x,y)$  is positive and red if it is negative, and we color it white if  $f_y(x,y)$  is

positive and blue if it is negative. We can use both colors additively to color the surface purple where it is both red and blue. Thus we color a point white, red, purple, or blue as the point  $(-f_x(x,y), -f_y(x,y))$  lies in the first, second, third, or fourth quadrant in the plane.

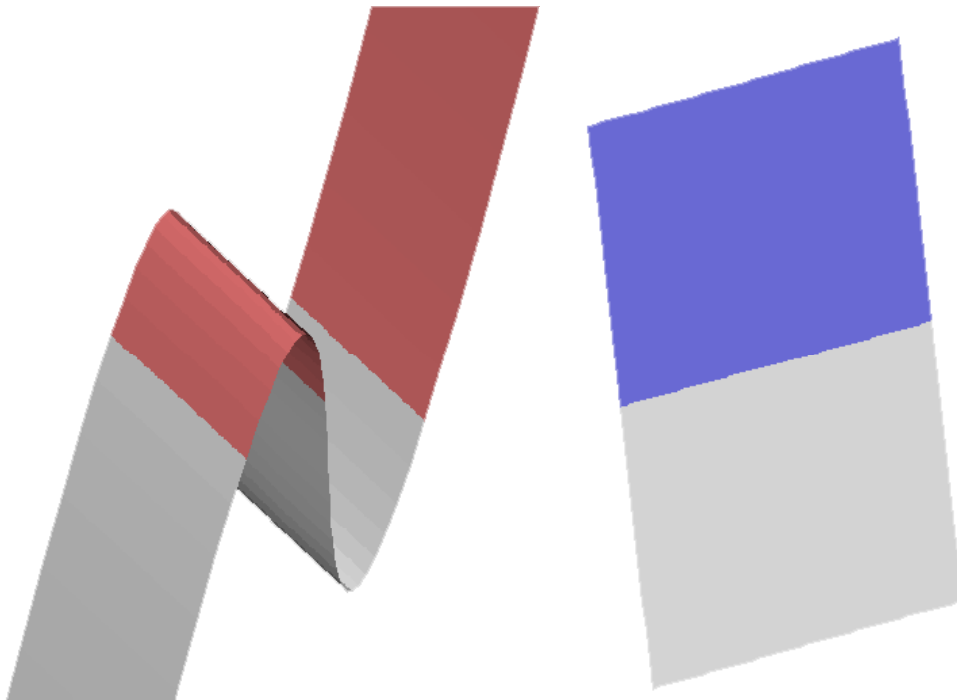
A point will be a critical point if all four colors meet at the point. We can also express this condition by saying that the point is in the intersection of the locus  $f_x(x,y) = 0$  and the locus  $f_y(x,y) = 0$ . Usually those conditions are expressed and dealt with only algebraically, but interactive computer graphics provides a direct visual display of the geometric properties of these partial derivative functions.



**Figure 7:** Perturbed Twin Peaks, color-coded domains

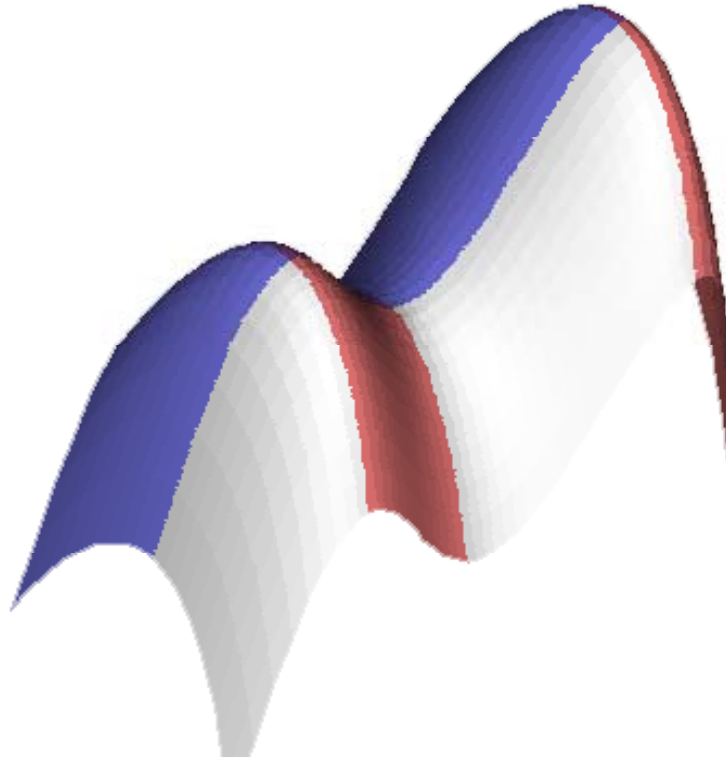
We obtain a bonus from this method of coloring because we can tell whether a critical point represents a maximum or minimum on one hand or a saddle point on the other. In the first case, the four regions at a point are red, purple, blue and white in counterclockwise order, while in the second case, the cyclic order is red, white, blue, and purple. The difference between these two orderings is the basis of the sign of the Gauss spherical image mapping defined on the surface.

This method of visualizing critical point configurations is especially helpful when we are exploring one-parameter families of functions, for example the perturbations of Twin Peaks.



**Figure 8:** Color-coded partial derivative graphs for Twin Peaks

The picture on the left shows the graph of the first partial derivative of  $f(x,y)$  with respect to  $x$ , colored to indicate where the value of that function is positive. The colored region is separated from the uncolored region by curves indicating where the  $f_x(x,y) = 0$ . The picture on the right shows the corresponding graph for the first partial derivative with respect to  $y$ . The colored region indicates where  $f_y(x,y)$  is positive, and that region is bounded by the curve where  $f_y(x,y) = 0$ .

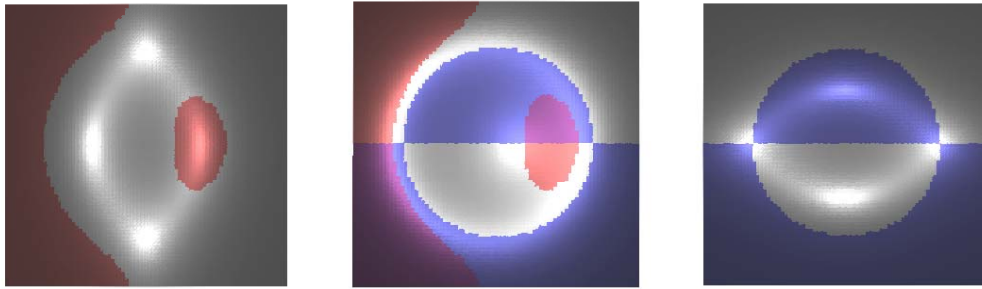


**Figure 9:** Color-coded graph of perturbed Twin Peaks

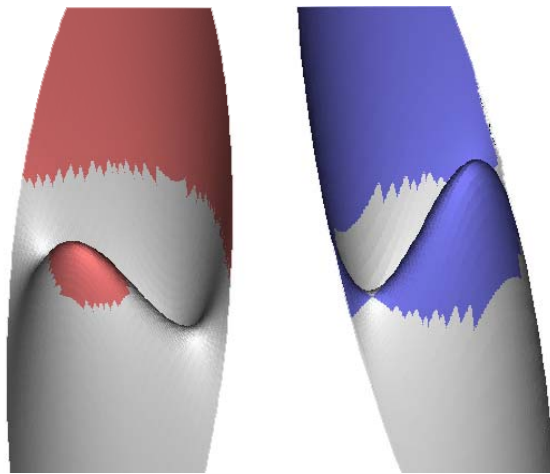
The graph of the perturbed Twin Peaks function colored according to the signs of the two first partial derivatives has four different colors near any critical point of the function, where the curves corresponding to  $f_x(x,y) = 0$  and  $f_y(x,y) = 0$  intersect.

## 11. Perturbations of the Graph of Crater Lake

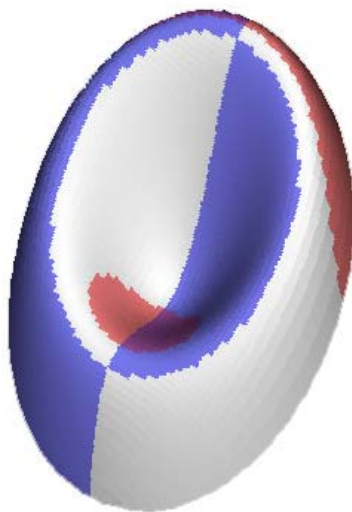
We can carry out the same analysis for the Crater Lake function to show the graphs of the two first partial derivatives. In the domain, we have a pair of curves where  $f_x(x,y) = 0$  and a pair of intersecting curves where  $f_y(x,y) = 0$ . The intersections of these two pairs of curves yield three critical points, one global maximum, one ordinary saddle, and one local minimum.



**Figure 10:** Perturbed Crater Lake, color-coded domain



**Figure 11:** Color-coded partial derivative graphs for Crater Lake

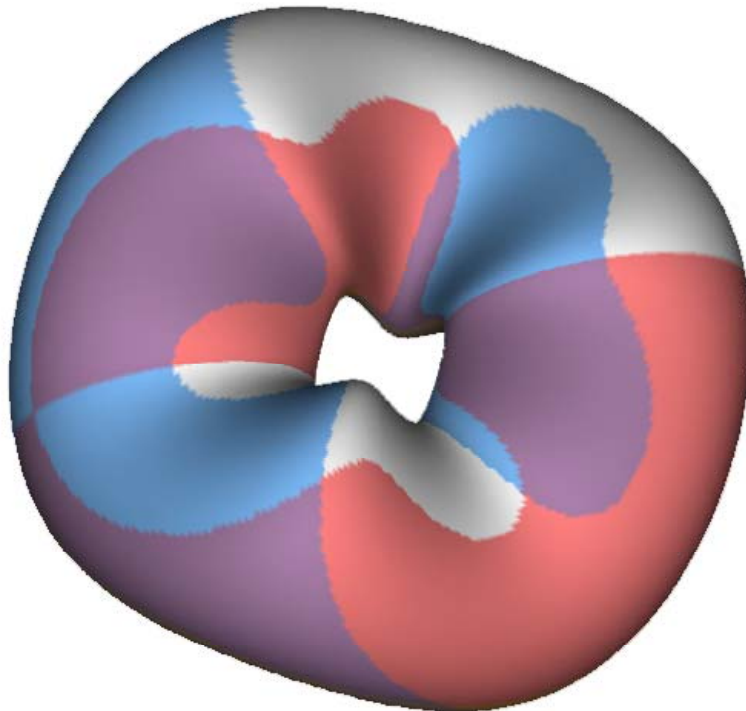


**Figure 12:** Color-coded graph of perturbed Crater Lake

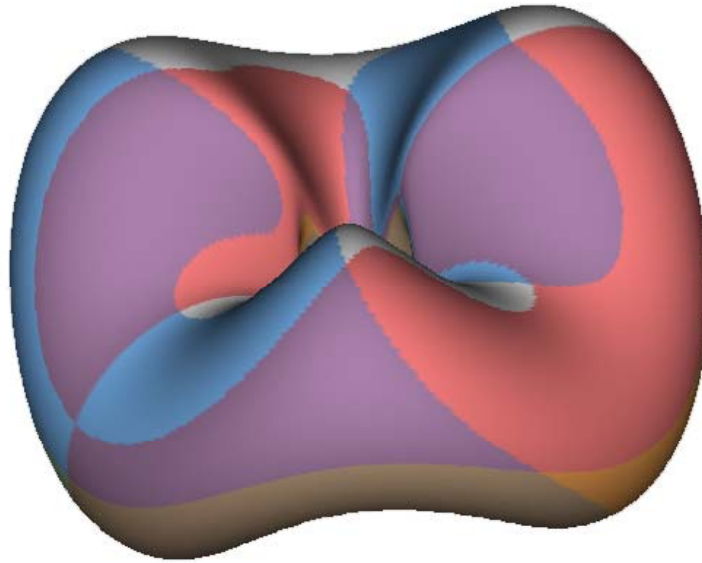
The graph of perturbed Creator Lakes function colored according to the signs of the two first partial derivatives has four different colors near any critical point of the function, where the curves corresponding to  $f_x(x,y) = 0$  and  $f_y(x,y) = 0$  intersect.

## 11. Parametric Surfaces in Three-Dimensional Space

Just as we went from graphs of functions of one variable to closed parametric curves in the plane, we can go from graphs of functions of two variables to parametric surfaces in space. For such surfaces, the outer normal vector can point into the upper hemisphere where the coloring is white, red, purple, and blue, or into the lower hemisphere, for which directions the color is yellow, overlaid as appropriate with red to make orange, or blue to make green, or purple to make brown. The eight quadrants of the unit sphere are then color-coded in such a way that a small polygonal region on a surface receives the color of the octant within which its outward normal vector lies.



**Figure 13:** Color-coded parametric warped torus, top view



**Figure 14:** Color-coded parametric warped torus, side view

We can exploit this color visualization to give an interpretation of the Hopf degree theorem for embedded (non-self-intersecting) surfaces and a modern proof of the extrinsic form of the Gauss-Bonnet Theorem, the most important theorem in the differential geometry of curves and surfaces. This will be the subject of a subsequent paper on interactive differential geometry.

## **Conclusion**

Interactive computer graphics makes it possible to investigate phenomena connected with critical points of functions in the plane and in three-dimensional space, starting with elementary calculus and proceeding to theorems in differential geometry and topology. These techniques have great potential for engaging students and general audiences as well as providing fruitful areas for research, in pedagogy as well as geometry and topology.

## **Acknowledgements**

Special thanks are due to Michael Schwarz, who wrote the Java demonstrations used for the illustrations in this article. The software that produced the Java applets was created at Brown University by David Eigen.

## **References**

- [1] Banchoff, Thomas “Beyond the Third Dimension” (1990) Scientific American Library,
- [2] Morse, Marston, “Topology and Equilibria” (November 2007) The American Mathematical Monthly, p. 819-834.