Lines and Conic Sections in Different Norms

Sandra Davis, LeMoyne College

OSU Summer REU Program, 1991

We wish to determine what conic sections look like as we change the norm. We are familiar with conic sections in the Euclidean norm, also called the 2-norm. A parabola is the set of points equidistant from a fixed point F and a fixed line D, the directrix. An ellipse is the set of points P such that for a fixed point, F and a fixed line, D and a real number, e such that 0<e<1, the distance from F to P is e times the distance from P to D.

Definition: The function N is a norm if the following properties hold for vectors v and w and scaler k:

    1.   N(v) >= 0 for all v and N(v) = 0 iff v = 0
    2.   N(k*v) = |k|*N(v)
    3.   N(v+w) = N(v) + N(w)

All of the functions such that $N(x,y) = (|x|^p + |y|^p)^{1/p}$ are norms for all p > 0. Norms of this form are called p-norms. As p approaches infinity the norm can be represented as max{|x|,|y|}. In this paper we will be discussing norms where p >= 1.
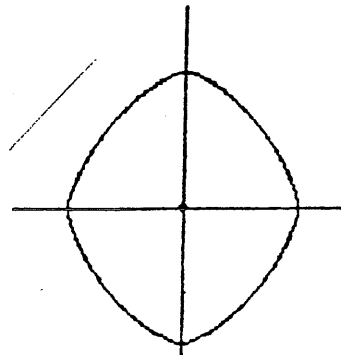
The circle has a formula of $(|x|^p + |y|^p)^{1/p} = c$ where c is a constant and the origin is its center. The unit circle looks like it "puffs out" as p increases from one to infinity. The fact that all circles where p >= 1 are convex is important for later proofs.

Four Unit Circles
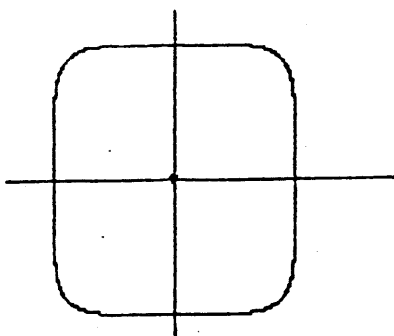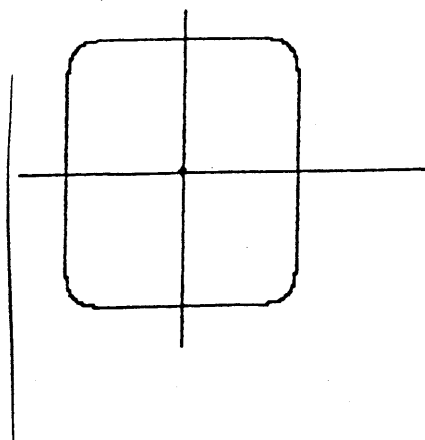
P = 1
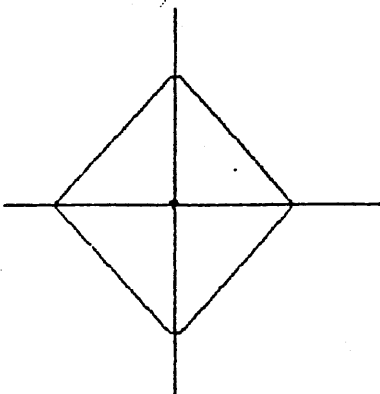
P = 1.5

P = 6

P = 9

Before we can determine what the conic sections other than the circle look like, we must know what a line is in the different norms. Geometrically, a line is a collection of segments whose sum totals the calculated distance between two common end points, say (x_0, y_0) and (x_1, y_1). In a p-norm, the distance, D is

$$D = (|x_0 - x_1|^p + |y_0 - y_1|^p)^{1/p}.$$

We then say for all (a,b) satisfying equation 1 that follows

$$(|x_0 - a|^p + |y_0 - b|^p)^{1/p} + (|x_1 - a|^p + |y_1 - b|^p)^{1/p} = D$$

that there is a segment from (x_0, y_0) to (a,b) combined with the segment (a,b) to (x_1, y_1) that together form one of the many lines that may be in the set of lines that have distance D.

Theorem 1: A line contains all segments in any p-norm such that the straight line is in the collection. A straight line is considered a line of the form $y = mx + b$.

Proof: The equation of a line is $b = m(a - x_0) - y_0$ where m is the slope and $m = (y_0 - y_1)/(x_0 - x_1)$. Substitute the line into equation 1 and we have

$$[|a-x_0|^p + |m(a-x_0)|^p]^{1/p} + [|a-x_1|^p + |m(a-x_0) + y_0 - y_1|^p]^{1/p} = D$$

Factor the first part and substitute for m and we have

$$|a-x_0|[(|x_1-x_0|^p + |y_1-y_0|^p)/|x_1-x_0|]^{1/p}$$
$$+ [|a-x_1|^p + |y_1-y_0|^p/|x_1-x_0|^p]^{1/p} = D$$

which is equivalent to

$$|(a-x_0)/(x_1-x_0)|D + |a-x_1|(1+|y_1-y_0|^p/|x_1-x_0|^p)^{1/p} = D.$$

Through similar manipulations we have

$$|a-x_0|/|x_1-x_0|D + |a-x_1|/|x_1-x_0|D = D.$$

Divide by D and it is easy to see that the equation holds for all a where a is a value between x_0 and x_1 and thus every (a,b) on the straight line forms a line of distance D.::

Theorem 2: A line is unique except in the one and infinite norms.

Proof: Unit circles are convex and curved in all but the one and infinite norms when $p \geq 1$. Therefore, two circles may intersect in 0,1,2, or infinite locations. Look at a line from (a,b) to (c,d) with distance D. Construct a unit circle with radius r around (a,b) and radius D-r around (c,d).

Suppose the circles intersect in 2 points. Then the intersection of the circles is non-empty and must contain a point say (x,y) that is not on either circle's graph. Since (x,y) is in the circle of radius r the distance (a,b) to (x,y) < r and similarly the distance from (x,y) to (c,d) would be < D-r. Summing them the distance (a,b) to (c,d) < D, which contradicts the triangle inequality defined in the norm. If they intersect in an infinite number of places the circles must be on top of each other. This can only happen when they have the same center or the circles contain straight lines as happens in the one and infinite norms. Therefore, circles may only intersect in one or zero points. From Theorem 1 we know that for all a there is a b such that equation 1 is satisfied. Therefore, the circles must intersect and they intersect in only one point. As r increases from 0 to D each a has exactly one b corresponding since the unit circles only intersect once for each r. Therefore, the line is unique for each p where 1 < p < infinity. (This argument can easily be extended to the case when 0<p<1.)::

Note that in the one and infinite norms there are an infinite number of different lines that have distance D. However, if the change in x is equal to the change in y, a unique straight line is produced.

Theorem 3: The shortest distance between a point and a line in the 1-norm is the vertical distance when the slope of the line is less than 1 and the horizontal distance when the slope of the line is greater than or equal to 1.

Proof: We begin by parameterizing our line; p=(0,b)+k(1,m) and chose point (0,0). We then want to minimize the function |k|+|b+km|. Look at k>0 and b>=-km. We then have k+b+km. The first derivative is m+1 and never = 0 for |m|<1 so the minimum must be at an endpoint. If |m|>=1 then m+1=0 when m=-1. So when m=-1 all points are equidistant. M=-1 is there exists 2 endpoints. The minimum distance between (0,0) and p is b. This occurs when k=0 or k=b, creating the two end points, (0,b) and (b,0). When |m| > 1, we know from definition that the change in y is greater than the change in x. Therefore, it follows that the shortest distance is to (b,0), the horizontal distance. Also, for |m| < 1 the shortest distance is to (0,b), the vertical distance. The other cases involving the other values for the absolute values follow in the same manner.::

We continue by trying to determine the shortest distance between a point (x,y) and a line in any given norm where p > 1. We will start by parameterizing the line so its formula is p=(0,b)+t(1,m). This is equivalent to saying p(t)=(t,b+tm).

Case I: We will look at x-t<0 and tm+b>0. This case is equivalent to x-t>0 and tm+b<0. (In later equations we may multiply through by -1 thus reversing the conditions.) We wish to minimize the function

$$((x-t)^p + (y-(b+tm))^p)^{1/p}.$$

If the inside of the function is minimized then the 1/p root will be so it is only necessary to look at $(x-t)^p + (y-(b+tm))^p$. Take the first derivative with respect to t. Using the first derivative test, we have

$$-p((x-t)^{p-1} + m(y-(b+tm))^{p-1} = 0.$$

This equation is only defined in the real number system for $m>0$. (If $m<0$ we would then be taking a negative root as well as having a positive expression equaling a negative expression.) Through algebraic manipulations we have the following series of equations:

$$x-t = (-y+b)m^{1/(p-1)}+tm^{1+1/(p-1)}$$

$$t(1+m^{p/(p-1)}) = (y-b)m^{1/(p-1)} +x$$

$$t = \frac{(y-b)m^{1/(p-1)}+x}{1 + m^{p/(p-1)}}.$$

Case II: We look at $t-x>0$ and $tm+b-y>0$ which is equivalent to the case $t-x<0$ and $tm+b-y<0$. The procedure for Case I follows and we have

$$t = \frac{(b-y)(-m)^{1/(p-1)}+x}{1 -m(-m)^{p/(p-1)}}.$$

Other Cases: If $m=b=y=0$ then the point is on the line so this case is trivial. If $t=x$ then the minimum distance is the horizontal distance. If $t=(y-b)/m$ then the vertical distance is the minimum distance between the point and the line.

In cases I and II, the point $(0,b)+t(1,m)$ gives a minimum distance when t equals the given expression in each. For example, the distance between (x,y) and $(((y-b)m^{1/(p-1)}+x)/(1+m^{p/(p-1)}, ((y-b)m^{p/(p-1)}+mx)/(1+m^{p/(p-1)})+b)$ is the shortest distance in Case I.

We then wish to see what the slope of the shortest distance is when $p>1$. We look at

$$\text{Slope} = \frac{tm+b-y}{t-x}$$

We then chose case I or II and substitute for t. This simplifies to

$$\text{Slope} = -m^{-1/(p-1)}$$

in both cases. This result can also be found by minimizing

$$(|t|^p+|tm+b|^p)^{1/p}.$$

We are now ready to begin looking at different conic sections. We
will begin by looking at parabolas with focus (0,1) and directrix,
y=mx+b.   The formula in the 1-norm is

$$|x|+|y-1|=\min\{|x-(y-b)/m|, |y-mx-b|\}.$$

Note that in the min statement we know which is the minimum by
looking to see if |m| is greater than or less than 1.   It is easy
to observe that if b>1 then the parabola will be concave down for
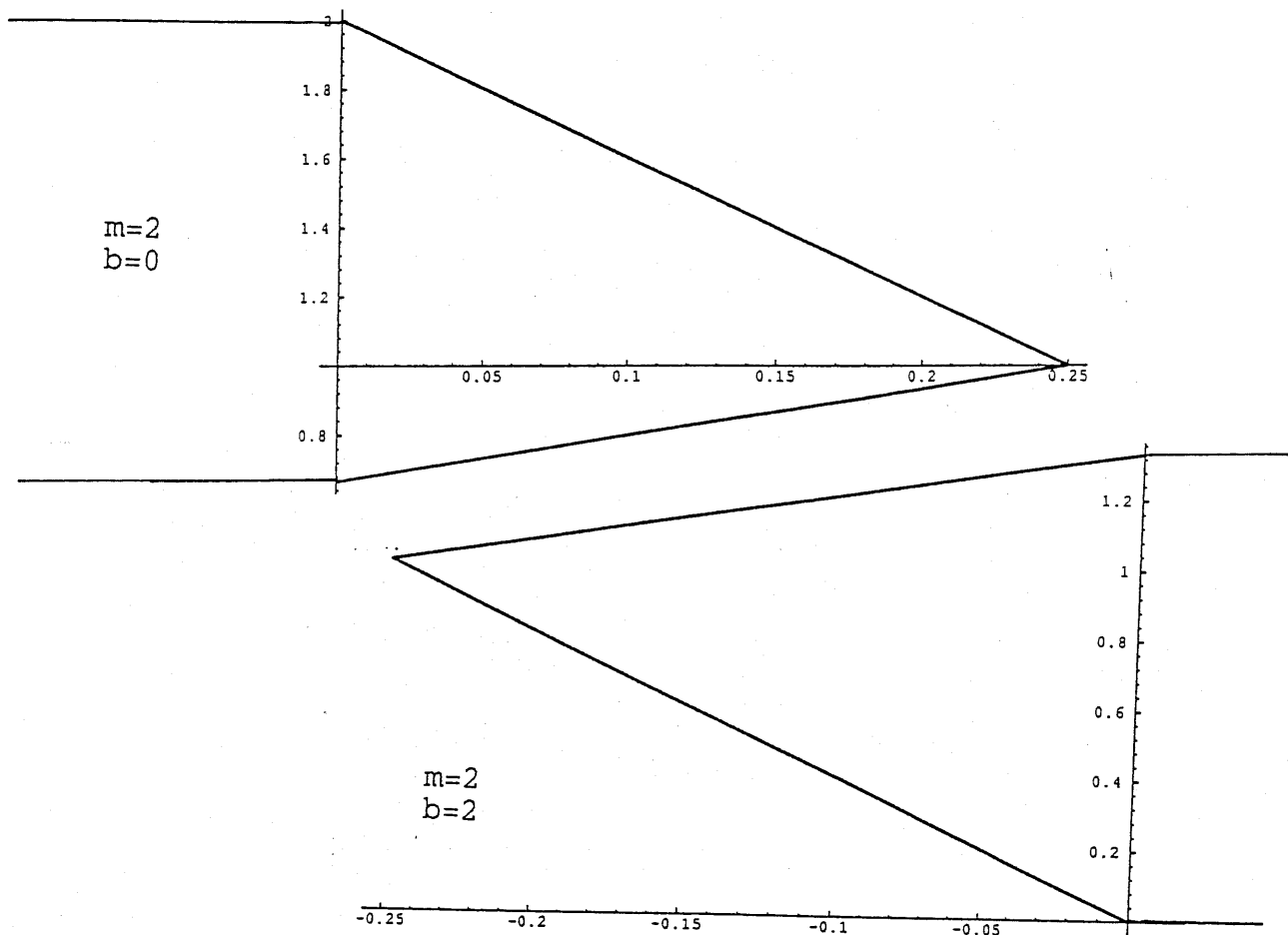m<1 and if b<1 and m<1 then it will be concave up.

m=5
b=0

m=.5
b=3
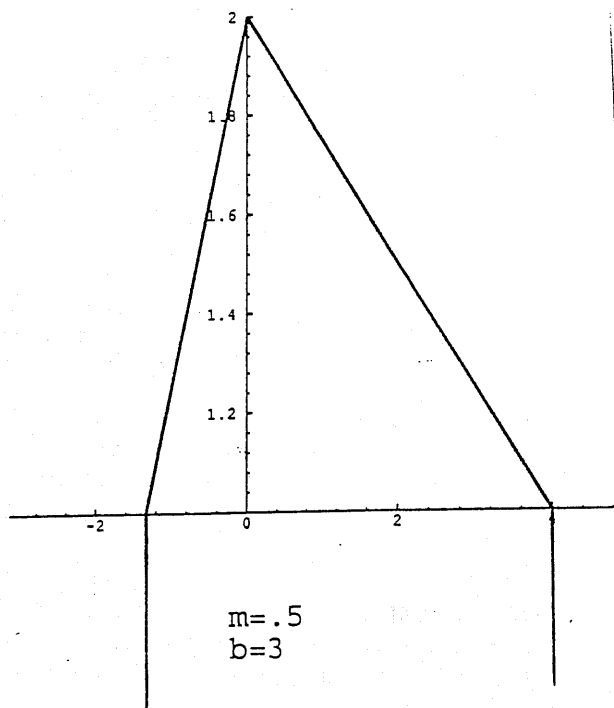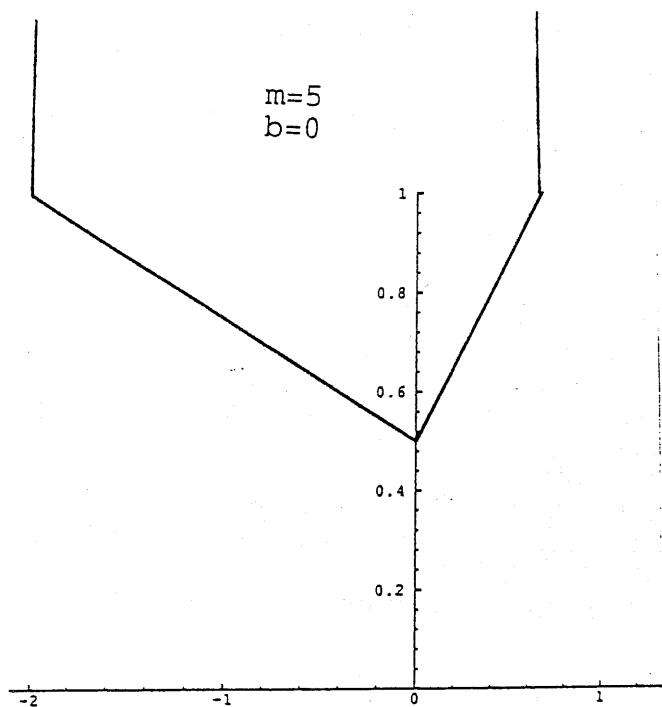
If b>1 is raised the graph also rises until it flips.

m=.5
b=0

b=-1
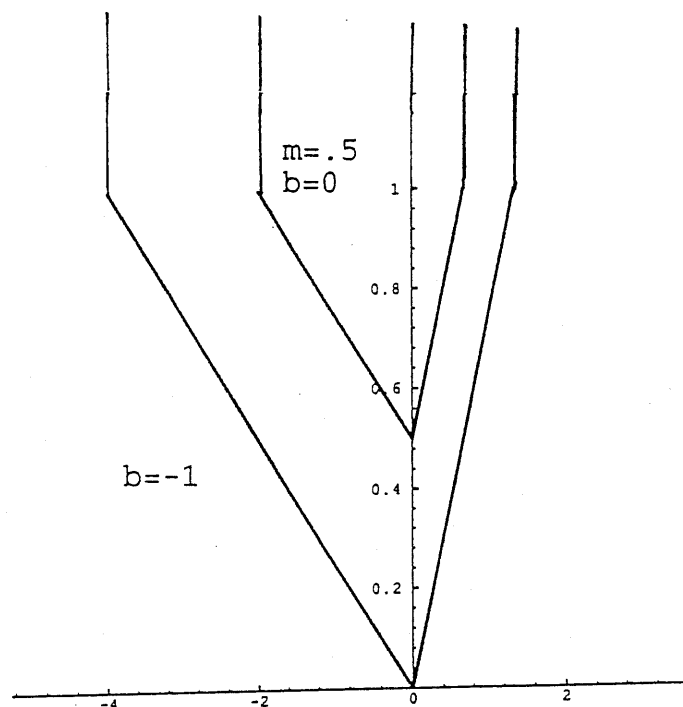
As m goes from -1 to one the parallel lines extending to infinity shift from right to left.

b=0
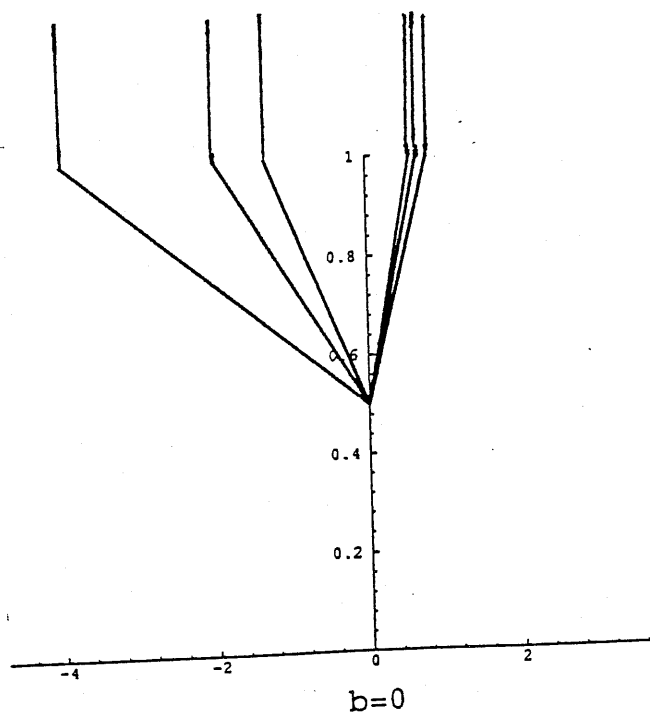
If m>1 and b<1 or m<-1 and b>1, then the parabola is concave left.
If m<-1 and b<1 or m>1 and b>1, the parabola is concave right.

m=2
b=0

m=2
b=2

Lowering b widens the space between the parallel lines extending to
infinity and lifts the graph. Also, changing b changes how far
left the graph extends.

m = 5
b = -60

b = -35
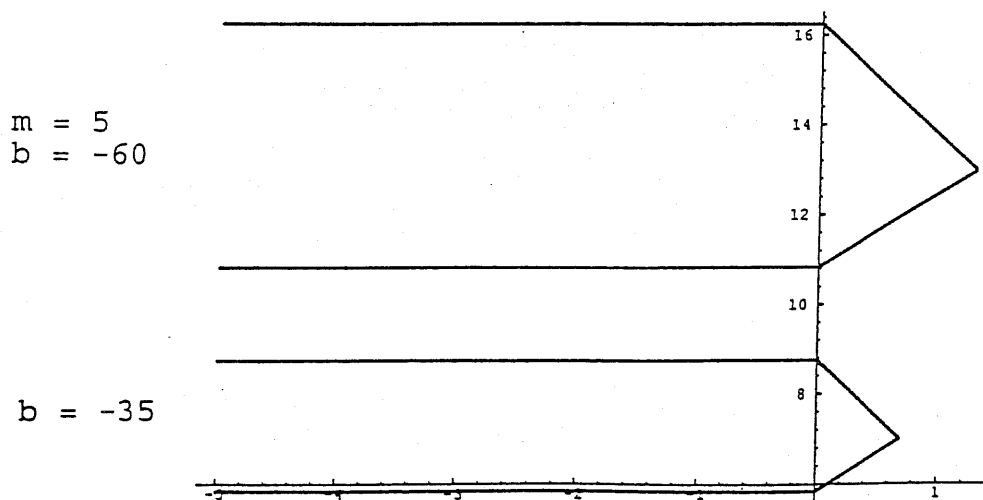
Changing the slope shrinks the graph.

b=0
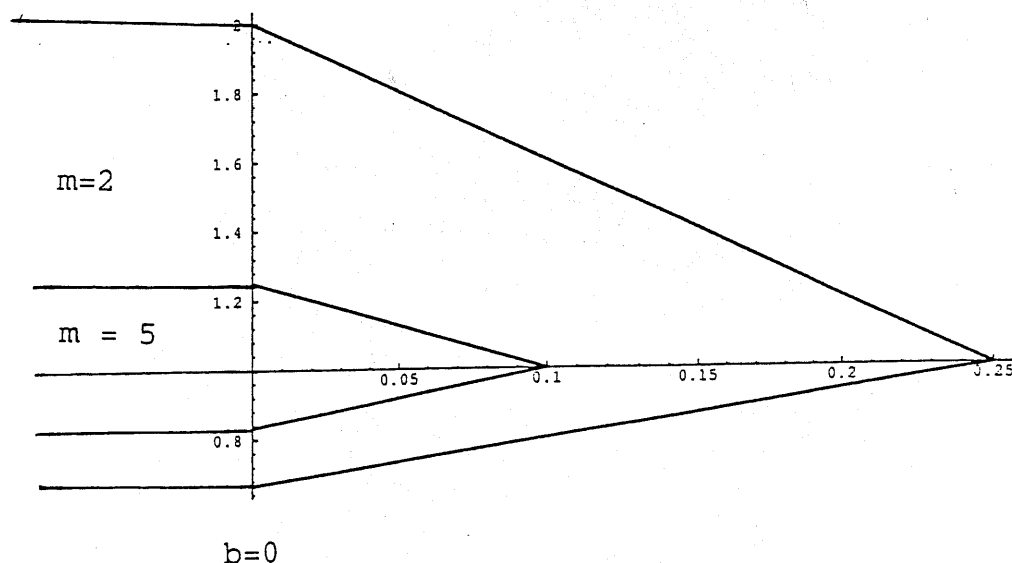
In larger norms, the parabolas move and the center axis tips as it does in the two norm. The only observed difference in movement is that the 1-norm graphs don't tip. As p increases, the graphs of parabolas fill out very much like the unit circles and approach the graph of the infinite norm very quickly. The infinite graph has sides appear to be parallel to the center axis. The bottom of the parabola square off and appear to become flat. The question of whether or not there is a consistent relation among cases as to where the graph squares off remains. The formula for a parabola in norm > 1 is
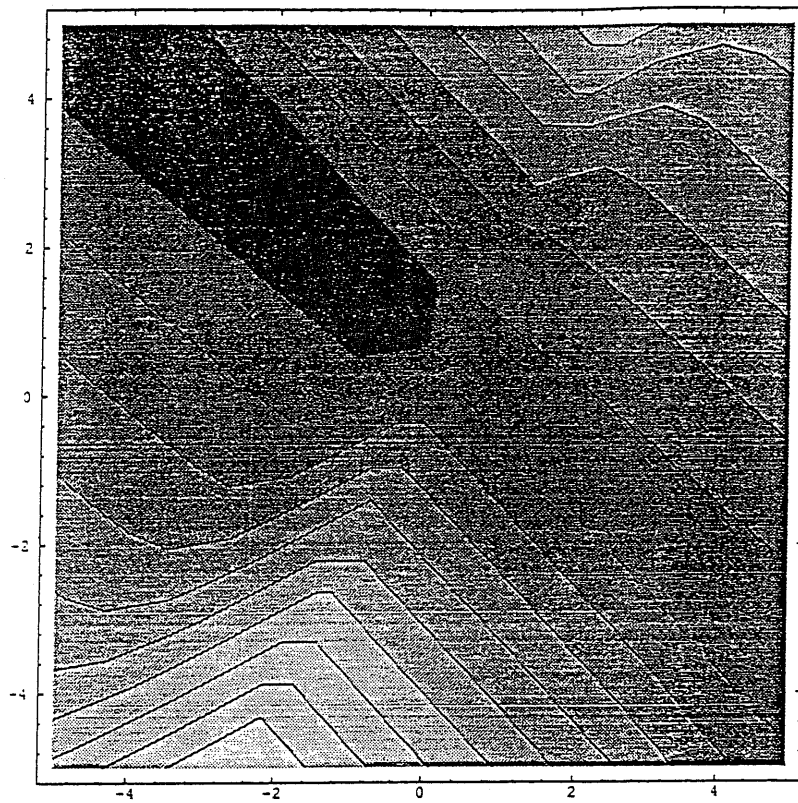
$$(|x|^p + |y-1|^p)^{1/p} = (|x-t|^p + |y-tm-b|^p)^{1/p}$$

where t is as found above. The following graphs contain the same parabola in different norms. Observe the boundaries of the black region.
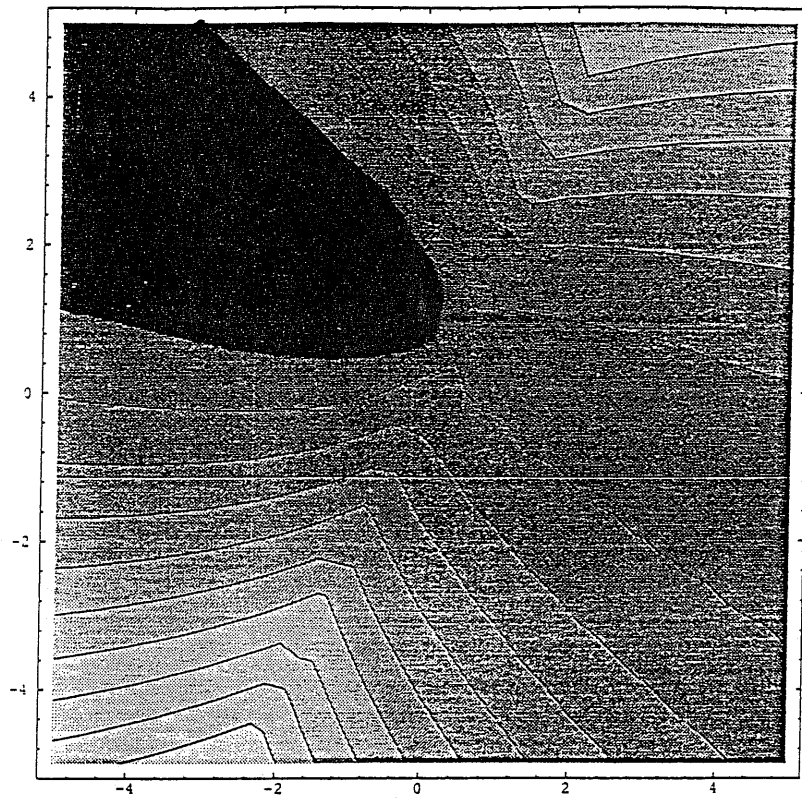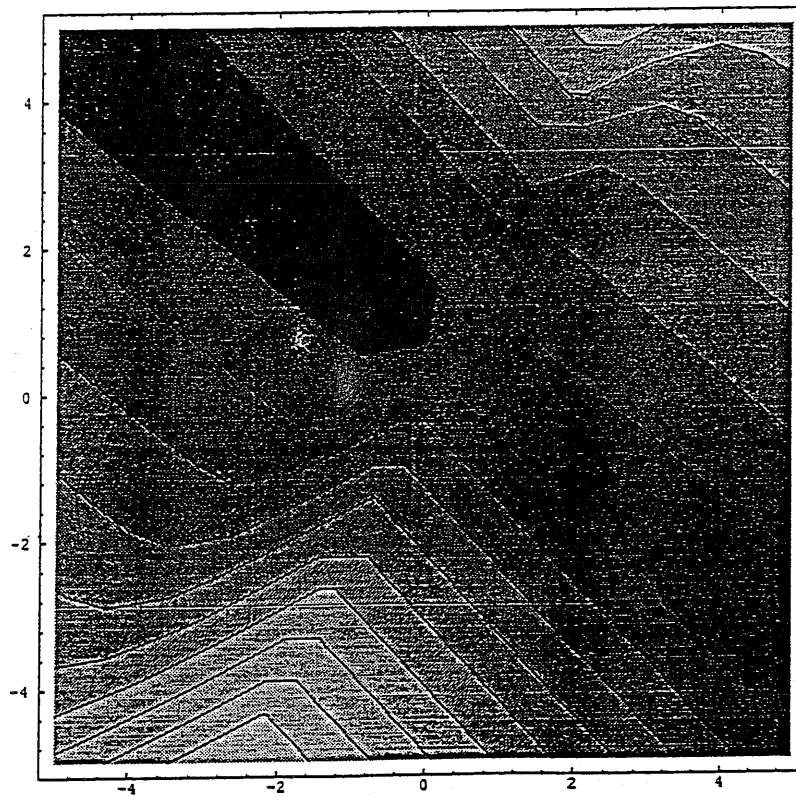
p=2
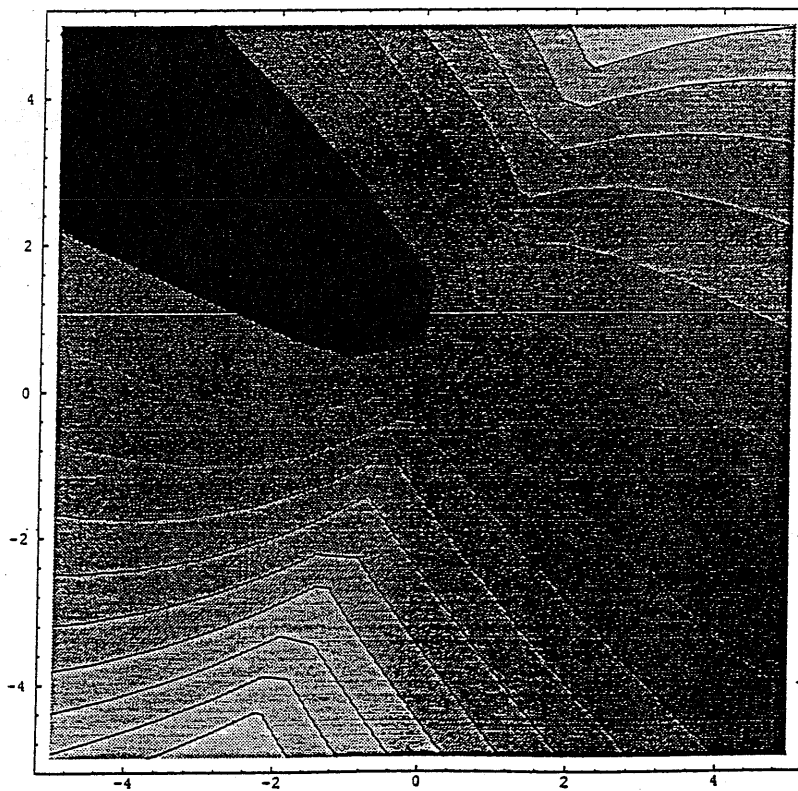
p=3

p=10

p=20

Ellipses can be determined in two ways in the 2-norm. One description uses two focus such that the distance from one to a point plus the distance from the other to that point equals some constant. The other way to describe an ellipse is with the focus directrix definition. Which is a better description outside the two norm. If you think of a cone being a set of circles of increasing size stacked on one another, I believe that the focus directrix definition is better. We will look at some examples in the one norm using the two focus description.

We will start by observing an ellipse with one focus at (-1,0) and the other at (1,0). Here we obtain an ellipse with six sides.
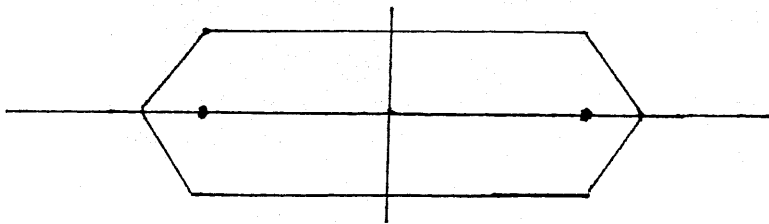
Conic sections are cross sections of a cone. The circle in the one norm has only 4 sides. How is it possible to make 4 sides into six. Further, if we look at the ellipse with one focus at (1,1) and the other at (-1,-1) we have an 8 sided figure.
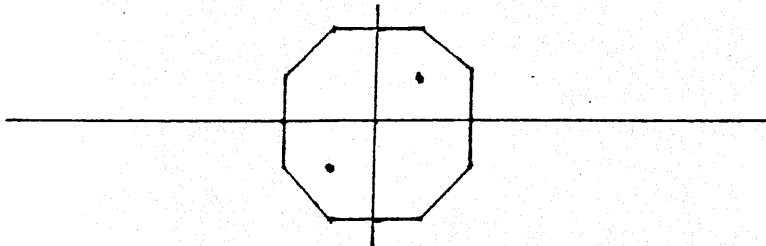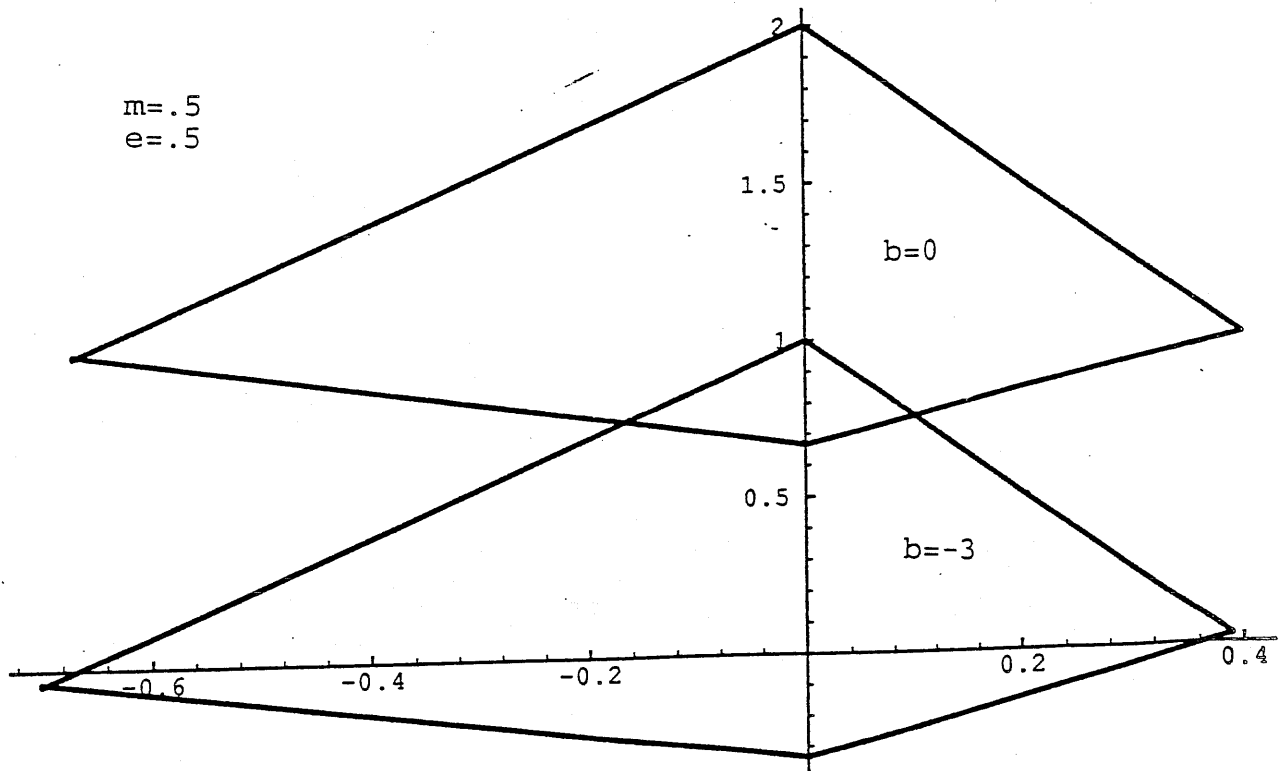
It seems like the directrix definition may be a better description. All ellipses here have 4 sides so it is possible to imagine it being a slice of the cone. These ellipses also behave in a manner similar to the parabolas. The values of m and b determine which direction the longer side of the ellipse faces in the same way they did for the parabola. Using the focus, (0,1) and the directrix y=mx+b, in the one norm the formula for an ellipse is

$$|x|+|y-1| = e * \min\{|x-(y-b)/m|, |y-mx-b|\}.$$

If b<1 and m increases the ellipse is shrunk.

e=.5
b=0

m=.5

m=2

2

1.8

1.6

1.4

1.2

0.2

0.4

-0.2

-0.6

0.4

0.8
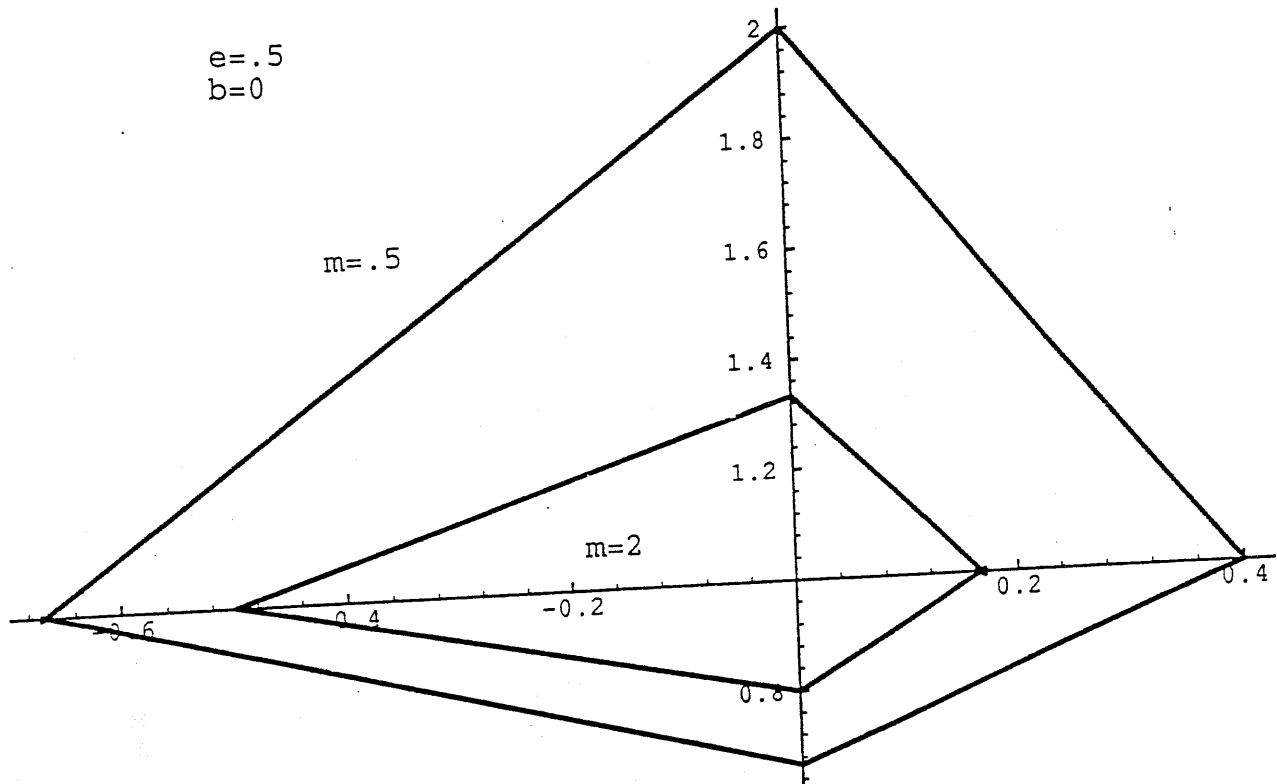
If b<1 and b is decreased the graph is lowered.

m=.5
e=.5

b=0

b=-3

2

1.5

1

0.5

-0.6

-0.4

-0.2

0.2

0.4

As e decreases the figure is smaller. Depending on the value of b
in relation to the focus if m is decreased either the width or the
height will decrease. Notice in the following figure that if b<1
then the width decreases.

b=0

m=.5
e=.5

m=.25
e=.5

m=.5
e=.25
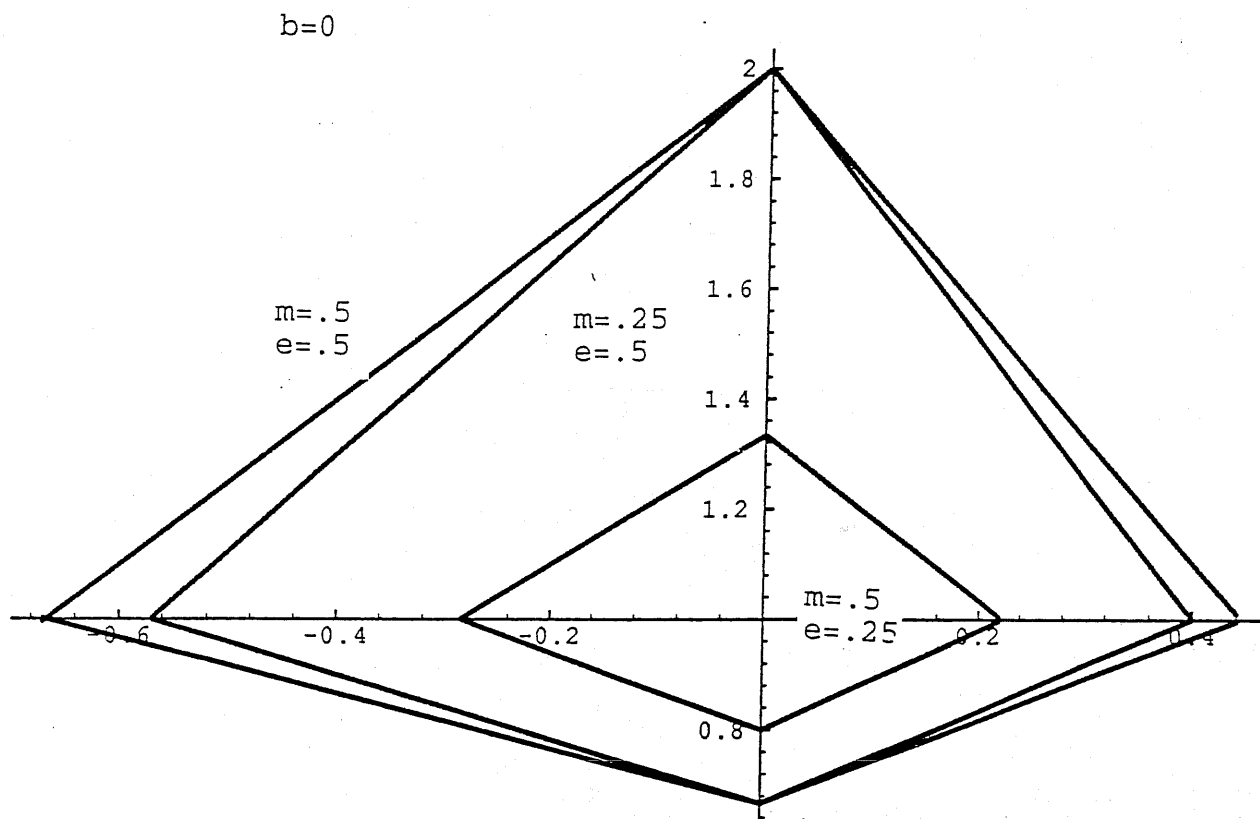
In norms lager than one again the ellipses fill out to the infinite
norm at a rapid rate. Here, they tip as in the two norm the main
difference in their movement. The formula for an ellipse in norm
> 1 is
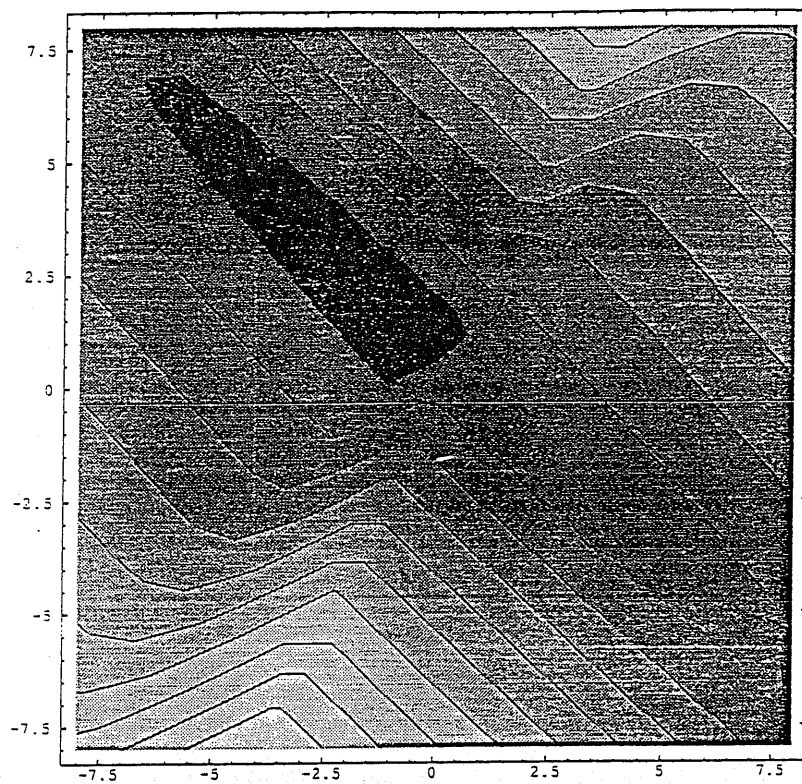
$$(|x|^p+|y-1|^p)^{1/p} = e*(|x-t|^p+|y-tm-b|^p)^{1/p}$$

where t is as found above and 0<e<1. The following graphs are all
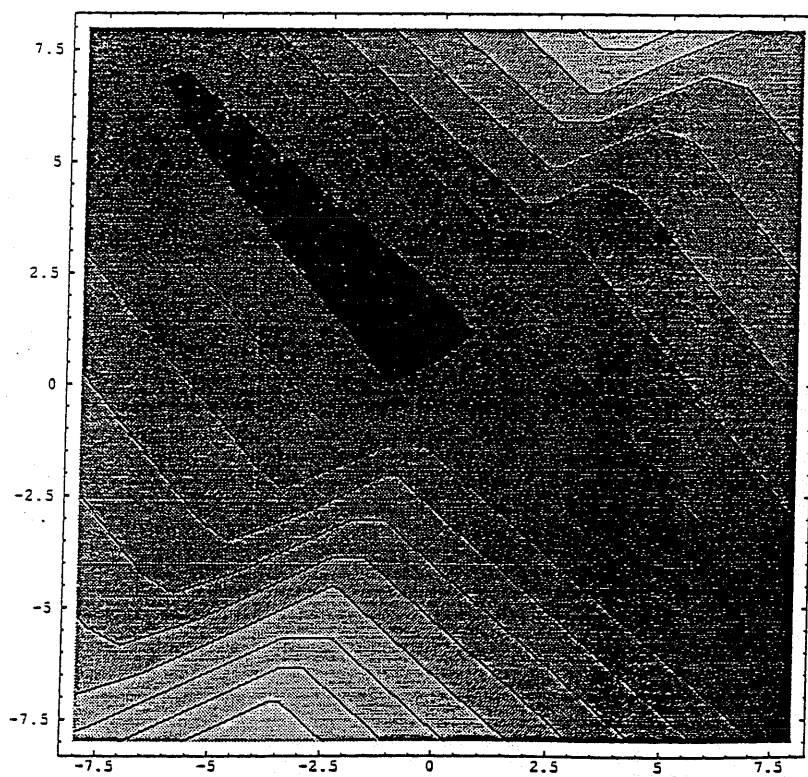of the same ellipse in different norms. Observe the boundary of
the black region.
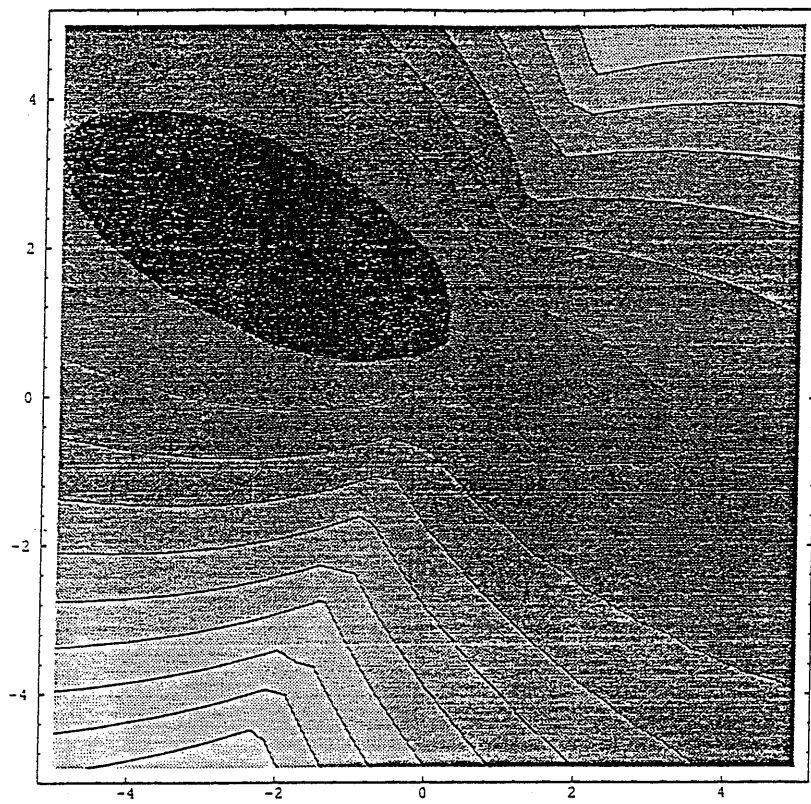
p=2
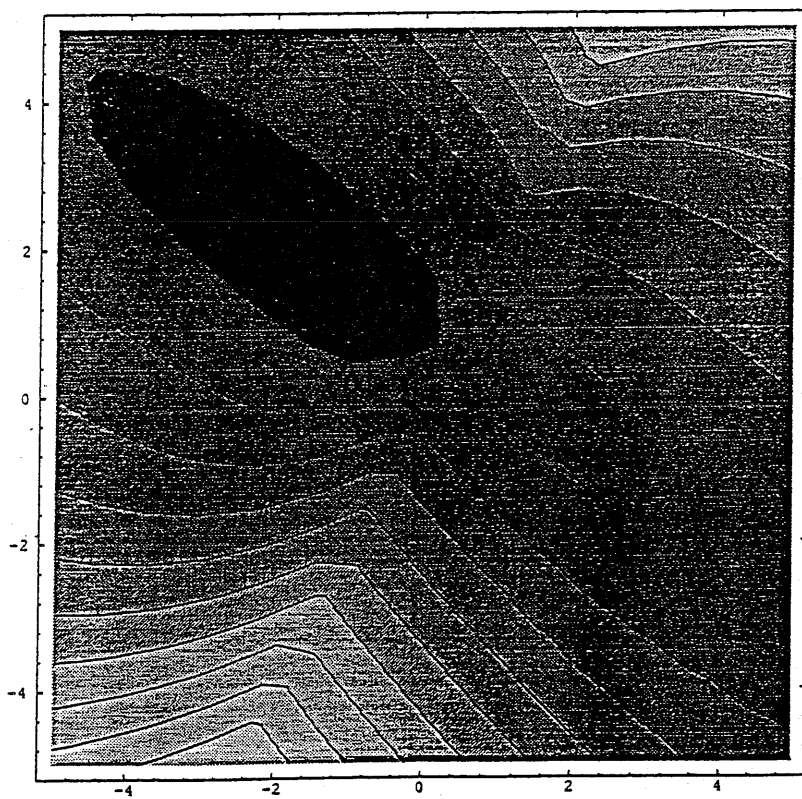
p=3

p=10

p=20

Summary:
While a line is not unique in the one and infinite norms, it is in the rest of the norms. Parabolas and ellipses are unique in norms greater than or equal to one. They have the convenience that the shortest distance between a point and a line is at a slope of $-m^{-1/(p-1)}$ in relation to the directrix's slope, m. The parabolas and ellipses appear to act as has been studied in the two norm.

BIBLIOGRAPHY

Effects of Changing the Measurement of Distance in the Plane on the
    Shapes of Conic Sections, Kimberly A. Linsley, Senior
    Thesis, 1987, unpublished.