6

Every Real Elliptic Curve Lives in a Donut

6.1 Complex Curves

Elliptic curves defined over $\mathbb{Q}$ — the rational elliptic curves — are standouts because they are so helpful in understanding and solving homogeneous third-degree number theory problems. But elliptic curves can just as well be defined over $\mathbb{R}$ or $\mathbb{C}$, and in these settings they reveal markedly different personalities. Studying them leads to surprises and deep connections. We now begin this journey. Here, the trio of fields

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

is an overall organizing principle that's as powerful as it is simple.

In this book we've so far worked mainly at the $\mathbb{Q}$ level, converting rational homogeneous number theory problems of degree at most 3 into lines, nondegenerate conics, and cubics. Geometry then helps us arrive at solutions. In $\mathbb{Q}^2$ we have seen that when they exist, rational points are everywhere dense in a nondegenerate conic, while for a cubic, in many cases there are only finitely many rational points. At the other extreme of $\mathbb{C}$, these degree-one, -two, and -three curves, respectively, become, topologically, a sphere, a sphere, and a torus. The rational points lying within any one of these surfaces constitute a very skinny subset and, amazingly, the genus of the full surfaces — a very global concept which Chapter 7 covers in more detail — tells us quite a bit about the nature of that skinny set of rational points. This remarkable linkage continues to hold even in higher degrees, where the genus is larger and number theory solutions are always sparse.

6.2 Complex Numbers Enlighten

The term "complex setting" can be taken in two different ways. One is in the affine sense as a direct analog of $\mathbb{R}^2$, namely $\mathbb{C}^2 = \mathbb{C} \times \mathbb{C}$. It can also be taken in the projective sense as a direct analog of $\mathbb{P}^2(\mathbb{R})$. In the first sense, $\mathbb{C}^2$ is just a vector space of complex dimension 2 — that is, the set of all ordered pairs (x, y) with x and y in $\mathbb{C}$ and supplied with scalar multiplication by elements of $\mathbb{C}$ as well as vector addition. As for the complex analog of $\mathbb{P}^2(\mathbb{R})$, we can proceed in essentially the same way as in the real case, which we looked at as the unit disk with antipodal points identified. This is depicted in Figure 2.2 on p. 40, where the map $r \to \frac{r}{|r|+1}$ compresses $\mathbb{R}$ to the open interval $(-1, 1)$, which is a "unit open 1-disk" — that is, an open 1-dimensional disk or interval of radius 1 (that is, diameter, or length, 2). We can do an analogous thing in the complex setting. If $|c|$ denotes the real distance from the origin to $c \in \mathbb{C}$, then the map $c \to \frac{c}{|c|+1}$ compresses $\mathbb{C}$ to the open 2-disk $|c| < 1$. Continuing our analogy to the real case, perform this shrinking on each complex 1-subspace of $\mathbb{C}^2$, the points of a complex 1-subspace being all $\mathbb{C}$-scalar multiples of some fixed nonorigin point in $\mathbb{C}^2$. The union of all these shrunken complex 1-subspaces is a subset of $\mathbb{C}^2$ and is analogous to the open real disk in $\mathbb{R}^2$. Now in the real case, we added boundary points of each 1-disk and identified antipodal points. That's equivalent to looking at any shrunken real line — an open interval bridging two endpoints — as a circle with a missing point, and adding that point to make a closed loop. The complex analog regards a shrunken $\mathbb{C}$ as a sphere with a missing point — essentially a Riemann sphere without its north pole. Now add that missing point. This collection of Riemann spheres is analogous to the set of circles in $\mathbb{P}^2(\mathbb{R})$. (Those nine lines in Figure 2.3 on p. 41, with antipodal points identified, are topologically nine circles.) We denote this complex analog as $\mathbb{P}^2(\mathbb{C})$. This is difficult to visualize, but we can informally think of $\mathbb{P}^2(\mathbb{C})$ as $\mathbb{C}^2$ to which we've added points at infinity. This will become clearer as we consider some examples.

Exercise 6.2.1. When the elliptic curve $y^2 = x^3 - x - 1$ is plotted in $\mathbb{C}^2$, most of its points are complex. $P = (0, i)$ is an example. The pseudocode in Appendix C styled for use by Maple or Mathematica can compute successive multiples of P and is algebraic in nature, so code working in the real rational setting can also calculate successive multiples of any complex P on the curve. Find $2P$, $3P$, and

$$4P = \left(\frac{223}{784}, \frac{24{,}655i}{21{,}952}\right).$$

Check that each of your points does in fact satisfy $y^2 = x^3 - x - 1$.

6.3 Plotting a Complex Circle

Example 6.3.1. Plotting a circle in the complex setting is illuminating and sets the stage for analogous plots of elliptic curves. Let's see what happens in the case of, say, the circle $x^2 + y^2 = 1$. We now assume x and y are complex, so we accordingly write $x = x_1 + ix_2$ and $y = y_1 + iy_2$. Substituting these into $x^2 + y^2 = 1$ and separating into real and imaginary parts gives

$$x_1^2 - x_2^2 + y_1^2 - y_2^2 = 1\,,$$

$$x_1x_2 + y_1y_2 = 0\,.$$

It turns out that each of these equations defines a real 3-dimensional surface in $\mathbb{R}^4$ and that their intersection is a real 2-dimensional surface in $\mathbb{R}^4$. Of course it's hard for most of us to visualize in 4-space, but one thing we can do is cut everything down by one dimension. This can be done by setting $x_2 = 0$, which defines a real 3D slice in $\mathbb{R}^4$. This 3D slice — (x_1, y_1, y_2)-space — intersects the real 3-dimensional surface in a real curve. By setting $x_2 = 0$, only x_1, y_1, and y_2 can assume nonzero values. With $x_2 = 0$, the two equations above reduce to

$$x_1^2 + y_1^2 - y_2^2 = 1\,,$$

$$y_1y_2 = 0\,.$$

Now $x_1^2 + y_1^2 - y_2^2 = 1$ defines a hyperboloid of one sheet, and $y_1y_2 = 0$ defines the union of the (x_1, y_2)-plane when $y_1 = 0$, together with the (x_1, y_1)-plane when $y_2 = 0$. The two equations mean that we don't see the whole hyperboloid, but only the part within these two planes. Figure 6.1 depicts this curve in (x_1, y_1, y_2)-space. $\diamond$

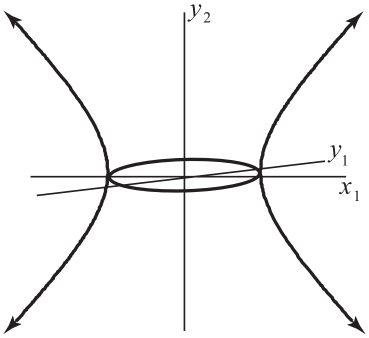

Figure 6.1. The intersection in $\mathbb{R}^4$ of the complex unit circle with (x_1, y_1, y_2)-space.

One of the things we've learned from the projective disk model $\mathbb{P}^2(\mathbb{R})$ is how points at infinity glue together the "ends" of branches to form topological loops. It turns out that the same thing happens in our complex analog, $\mathbb{P}^2(\mathbb{C})$. The hyperbola in the (x_1, y_2)-plane has two asymptotes, and at the "ends" of each asymptote, branches get glued together. For example, to go from Figure 6.1 to Figure 6.2, think of the hyperbola in Figure 6.1 as made of wire. Bend the two topmost wires in the (x_1, y_2)-plane — always keeping them in that plane — so that the two ends almost meet to make the top half of a circle. Repeat, using the bottommost two wires to make the bottom half of a circle. Then, just before soldering the ends together, twist the right half of the circle 180 degrees about the x_1-axis, and then do the soldering. The two soldered joints are the two points "∞" in the figure. Topologically, we get one big loop having two solder joints, with this big loop still touching the real circle at 1 and -1 in Figure 6.2.

Now our space $\mathbb{R}^3$ defined by $x_2 = 0$ is but one 3-dimensional slice in $\mathbb{R}^4$. More generally, any $x_2 = r =$ a constant defines a 3-space parallel to the slice $x_2 = 0$. As r varies through $\mathbb{R}$, these slices fill 4-space. As an example, let's vary x_2 a little — say, to $x_2 = r = 0.1$. With $x_2 = 0.1$, $x_1^2 - x_2^2 + y_1^2 - y_2^2 = 1$ then becomes $x_1^2 + y_1^2 - y_2^2 = 1.01$. This still defines a hyperboloid of one sheet, but instead of being intersected by the two planes defined by $y_1 y_2 = 0$, the intersection is with $y_1 y_2 + 0.1 x_1 = 0$, another hyperboloid of one sheet, but which looks quite a bit like two planes. This is depicted as the shaded surface in Figure 6.3, and the

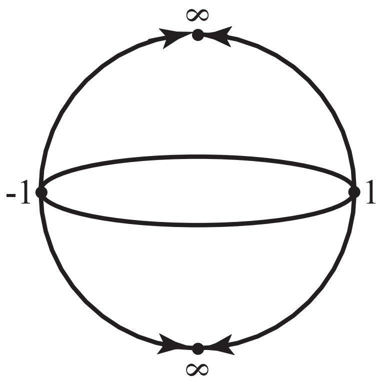

Figure 6.2. Here, the skeleton curve in Figure 6.1 has been topologically massaged to fit on a sphere.

previous circle-plus-hyperbola has morphed slightly into two curves, each running close to the circle-plus-hyperbola. For clarity, we have drawn just one of these.

If we topologically transform Figure 6.3 in the same way that Figure 6.1 got transformed into Figure 6.2, our curve now appears somewhat

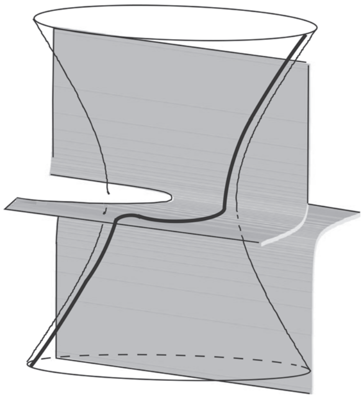

Figure 6.3. This shows one of the two disjoint curves comprising the intersection of the complex circle with the slice $x_2 = r = 0.1$.

like the largest loop in the top half of Figure 6.4, and the other curve (the one we didn't draw in Figure 6.3) appears dashed on the rear lower half. We see other loops corresponding to other choices of $r > 0$, and these loops fill out half the sphere, the other half being filled out as r runs through negative values. ◇

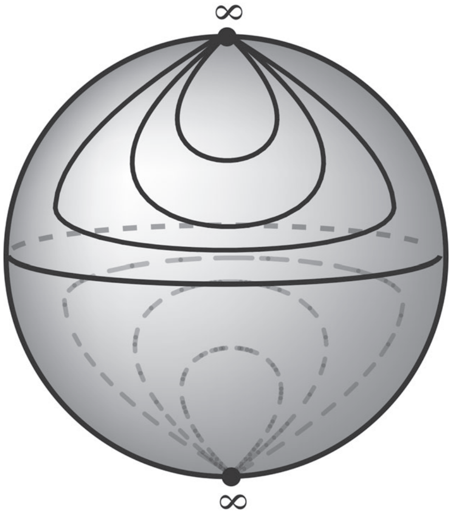

Figure 6.4. As r in Figure 6.3 takes on all real values, the sphere gets covered with disjoint curves. A few are shown here.

Exercise 6.3.2. Follow the approach used for plotting a complex circle to plot the (x_1, y_1, y_2)-slice of the complex parabola $y = x^2$. Topologically map this skeleton to a sphere. Then plot the slices $x_2 = r$ (r small) to see how nearby parallel slices map to the sphere. Finally, indicate how the remaining curves fit on the sphere by sketching in a few more curves.

Exercise 6.3.3. Redo Exercise 6.3.2 for the complex hyperbola $xy = 1$.

6.4 Plotting a Complex Elliptic Curve

Let's now explore plotting an elliptic curve in the complex setting. We can mimic the steps used for a unit circle in the last section.

Example 6.4.1. For simplicity, we choose the elliptic curve $y^2 = x^3 - x$, shown in Figure 6.5, where we see both its affine plot and its image in the projective disk. As with the circle, assume x and y are complex and set $x = x_1 + ix_2$ and $y = y_1 + iy_2$. Substituting these into $y^2 = x^3 - x$ and separating into real and imaginary parts gives

$$y_1^2 - y_2^2 = x_1^3 - 3x_1x_2^2 - x_1\,,$$
$$2y_1y_2 = 3x_1^2x_2 - x_2^3 - x_2\,.$$

Again, as with the circle, we first look in the 3-space $x_2 = 0$, and in that (x_1, y_1, y_2)-space the two equations lead to

$$y_1^2 - y_2^2 = x_1^3 - x_1 \quad \text{and} \quad y_1y_2 = 0\,.$$

As before, $y_1y_2 = 0$ defines the union of the (x_1, y_1)- and (x_1, y_2)-planes. In the (x_1, y_1)-plane defined by $y_2 = 0$, the equation $y_1^2 - y_2^2 = x_1^3 - x_1$ becomes $y_1^2 = x_1^3 - x_1$ and we see the real elliptic curve we started with. In the (x_1, y_2)-plane given by $y_1 = 0$, the equation $y_1^2 - y_2^2 = x_1^3 - x_1$ becomes $-y_2^2 = x_1^3 - x_1$, and we see essentially the reflection of the curve about the y_1-axis but drawn in the (x_1, y_2)-plane. These curves are depicted in Figure 6.6, where the more lightly drawn part is the reflected curve.

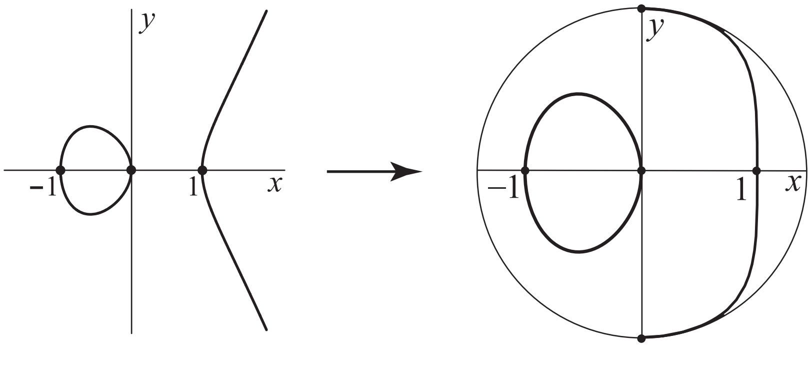

Figure 6.5. Image in the projective plane of a cubic.

In this figure, we see two branches. The one in the (x_1, y_1)-plane joins ends at the end of the y_1-axis, and the other joins its ends at the end of the y_2-axis. Since we're in the complex setting, the (y_1, y_2)-plane is the Riemann sphere minus its point at infinity, and the two branches both join up at this one point at infinity, $\{\infty\}$. So topologically, one loop goes from $x_1 = 1$ to $\{\infty\}$, with another one going from $x_1 = -1$ to $\{\infty\}$. Together with

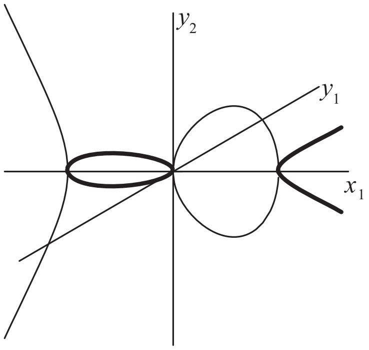

Figure 6.6. The heavily drawn curves are what we see in the usual (x_1, y_1)-plane. Looking in the (x_1, y_2)-plane reveals a “reflected” image of this.

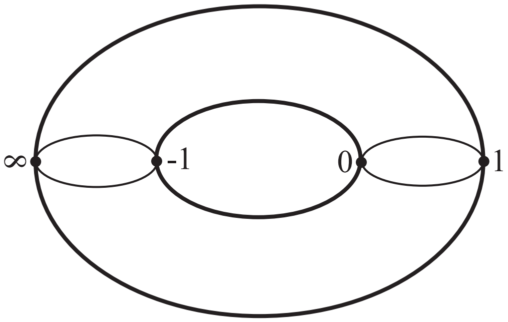

Figure 6.7. The ordinary and reflected curves in Figure 6.6 topologically lie on a torus in a natural way.

the other two loops touching the origin, there are four loops. In Figure 6.6, two of the loops can be regarded as made of thin wire, and the other two, out of thick wire. These wires can be bent to form the skeleton depicted in Figure 6.7. In analogy to the circle example, we can now let x_2 run through real values r, and the corresponding curves fill out the torus. The curves for $r > 0$ are depicted in Figure 6.8 and fill out half the torus — the upper-front and lower-rear quarters in the picture. The lower front and upper rear are filled in as r runs through negative values. ◇

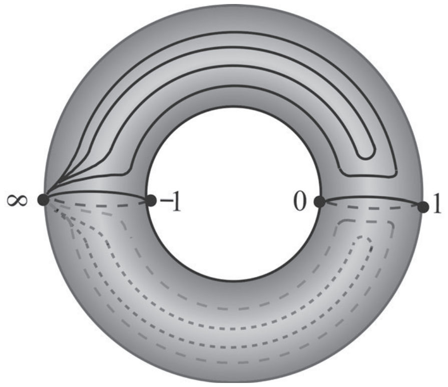

Figure 6.8. The curves corresponding to setting x_2 equal to real constants topologically cover the torus with curves, any two touching only at the point at infinity.

Exercise 6.4.2. For the complex curve $y^2 = x(x^2 - 1)(x^2 - 4)$, use the methods of this section to topologically map intersections of the curve with 3D slices of $\mathbb{R}^4$ to a double torus like the one in Figure 7.1 on p. 156.

6.5 Subgroups and Cosets

When we plotted an elliptic curve in the complex setting, curves over $\mathbb{R}$ played a pivotal role since the complex curve was built up as the union of real curves. In fact, the very suggestive torus skeleton depicted in Figure 6.7 is topologically the union of, for example, the real elliptic curve $y^2 = x^3 - x$ and its "brother" $y^2 = -x^3 + x$, obtained by replacing x everywhere by $-x$. There are two groups within this skeleton, one of them consisting of the heavily drawn curves in Figure 6.7 and the other one comprised of the lightly drawn curves. It's not hard to see this fundamental fact shared by elliptic curves over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$:

> The connect-the-dots addition algorithm works in all three settings. Therefore an elliptic curve in $\mathbb{P}^2(\mathbb{C})$ is an abelian group, as are the real and the rational subsets of the curve.

From now on, the term *group* will always mean *abelian group*.

To keep things simple, we continue to work with the familiar example of $y^2 = x^3 - x$ shown in the left picture of Figure 6.5, although our arguments are basically the same for any short-form elliptic curve with an oval. (If there's no oval in the usual real plane, it gets revealed by looking in a different plane, as we will see in the next section.) The curve $y^2 = x^3 - x$ is an abelian group in $\mathbb{P}^2(\mathbb{R})$ under our connect-the-dots algorithm for which a line cutting the curve is basic. From just looking at the curve, it seems that any line cuts the real curve in exactly one or three points. This is easy to check. On the one hand, when $m \neq 0$, we can substitute $mx + b$ for y in $y^2 - x^3 + x = 0$ and this results in a cubic in x which factors into three linear terms. If their roots are all real, the line cuts the curve in three real points. If the roots are not all real, then because complex roots come in conjugate pairs, there's just one real root and the line cuts the real curve in a single point. On the other hand, if the line is vertical, say $x = x_0$, then plugging x_0 in for x in $y^2 - x^3 + x = 0$ gives y^2 equal to some constant $k = x_0^3 - x_0$. There are three cases. If $k > 0$, then y has two distinct solutions which, with the point at infinity, give three points of intersection. If $k = 0$, then y^2 has two identical solutions, meaning the line $x = x_0$ is tangent to the oval and this double point plus the point at infinity again gives three intersections. $k < 0$ means the only point of intersection is the point at infinity.

This leads to two observations about the connect-the-dots algorithm.

- The first is that any line intersects the branch in at least one point. If the line is vertical, it intersects the branch at infinity. If it's not vertical, the reader can pin this case down in Exercise 6.5.8 on p. 148.

- The second observation is that any line intersects the oval in either zero or two points. To see this, suppose a line intersects the oval in a point P. As a second point Q travels once around the oval (starting, say, from P as a tangent line there), the angle the line makes with the horizontal changes by 180°, thus accounting for all slopes of the rotating line. So because every line has a slope (which can be infinity), a line intersecting the oval in some point P also intersects it in some second point Q. The line can't intersect the oval in three points since the line must intersect the branch, so any line intersects the oval in either zero or two points.

The above observations allow us to see this fundamental group-theoretic fact about the branch and the oval.

> The branch (always including the point at infinity) is a subgroup of the entire real projective elliptic curve.

This is because the points of the branch are closed under addition and subtraction. That's true because the sum of two points in the branch must remain in the branch. If not, that sum would be in the oval, an impossibility because the oval either contains zero or two points, never just one point. Ditto for the difference.

Notation 6.5.1. When a group and subgroup are clear from context, we let G denote the group and H the subgroup. ◊

Definition 6.5.2. A *coset* of a subgroup H of a group G is a translate by some $g_0 \in G$ of that subgroup — that is, a set $H + g_0 = \{h + g_0 \;:\; h \in H\}$. We typically assume $g_0 \notin H$. ◊

Comment 6.5.3. It is easy to show that a coset of H is either H itself or is disjoint from it. ◊

Example 6.5.4. $\mathbb{R}^2$ is a group G under vector addition, and any line H through the origin is a subgroup of G. Any (parallel) translate of H by a vector g_0 is a coset of H. Another line through the origin $H' \neq H$ intersects each coset in one point, and we may let that point in the subgroup H' serve as a representative of the coset. The set H' of all these representatives can be thought of as the quotient group G/H of G by H. ◊

Example 6.5.5. In the above example, G can be any $\mathbb{R}^n$ and H, any subspace of G. H' can be any vector space complement of H — for example, the orthogonal complement of H. If H has dimension d, then H' is G/H and has dimension $n - d$. In the expected sense, G is isomorphic to $H \times G/H$ — that is, each element of G can be identified with a pair of elements from the subgroups H and H'. ◊

Example 6.5.6. In Example 6.5.5, the field $\mathbb{R}$ can be replaced by $\mathbb{C}$ or in fact by any field. ◊

Example 6.5.7. Let $G = \mathbb{R}$ and $H = \mathbb{Z}$. Then $\mathbb{R}/\mathbb{Z}$ is isomorphic to a circle group $\bigcirc$. The points of the circle parameterize the cosets of $\mathbb{Z}$ in $\mathbb{R}$, as do the points of the interval $[0, 1) \subset \mathbb{R}$. $\mathbb{R}$ is 1-dimensional. An analogous 2-dimensional example would be $G = \mathbb{R}^2$ and H equal to the plane lattice $\mathbb{Z}^2$. Then $\mathbb{R}^2/\mathbb{Z}^2$ is isomorphic to the torus group $\bigcirc \times \bigcirc$. Its points (or equally well, those of $[0, 1) \times [0, 1) \subset \mathbb{R}^2$) parameterize the points of this quotient group. $\diamond$

Exercise 6.5.8. Establish that any nonvertical line intersects the branch of the cubic $y^2 = x^3 - x$.

Exercise 6.5.9. If $\mathbb{Z}^n$ is a cyclic group of order n and $\mathbb{Z}^m$ is a subgroup of $\mathbb{Z}^n$, when is $(\mathbb{Z}^n)/(\mathbb{Z}^m)$ isomorphic to $\mathbb{Z}^{n-m}$?

Returning to the real curve $y^2 = x^3 - x$ shown in Figure 6.5 on p. 143, we know that the branch is a subgroup H of the entire real elliptic curve G, but what can we say about what's left over, the oval? We've seen before that not only is the oval a loop, the branch is, too — its ends connect at infinity, so the branch forms a topological loop. This suggests that perhaps the oval is a coset of the branch! This is indeed so, and here's why: If g_0 is any point of the oval, we will show that the oval is the set $H + g_0$ — that is, the oval is a group-theoretic translate of the branch by an element of the oval. To see that if g is any point of the oval, then there's an $h \in H$ so that $g = h + g_0$, let $-g$ be the reflection of g about the x-axis. By our previous observations, the line through $-g$ and g_0 intersects H in a point h, so we have $g = h + g_0$, as desired. This is depicted in Figure 6.9.

We've learned that the real projective elliptic curve is a group G, as is its branch H. It's fair to ask what the quotient G/H is. Since H has just the oval as a coset, the quotient consists of two elements, and the only 2-element group is $\mathbb{Z}_2$. The observations above fit in with this, because with $\{0 \leftrightarrow \text{branch}\}$ and $\{1 \leftrightarrow \text{oval}\}$, $0 + 0 = 0$ corresponds to the sum of two branch points being a branch point, while $0+1 = 1+0 = 1$ corresponds to the sum of a branch point and oval point being in the oval. Also, $1+1 = 0$ corresponds to the sum of two points in the oval being in the branch.

We are not quite done with our story. Take a look at Figure 6.6 on p. 144 depicting the space curve obtained when our elliptic curve in the complex setting is sliced by the (x_1, y_1, y_2)-subspace of $\mathbb{R}^4$. The part of the space curve that's drawn more lightly is a reflected (and rotated) version of the darker curve we've been discussing. If we look at this lightly

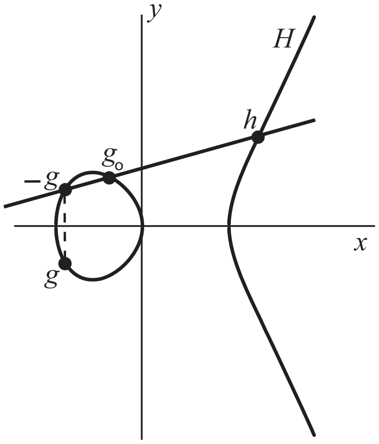

Figure 6.9. The oval is a coset of the real projective curve's branch group. Given points g_0, g in the oval, the line through g_0 and $-g$ intersects the branch in h. This picture can be summed up as $h + g_0 = g$.

drawn curve as simply a plane curve in the usual (x, y)-plane, its equation is obtained from $y^2 = x^3 - x$ by replacing x everywhere with $-x$, giving $y^2 = -x^3+x$. Just as the branch in $\mathbb{R}^2$ of any curve $y^2 = x^3+ax+b$ is rightmost and opens to the right, so the branch of any curve $y^2 = -x^3 - ax + b$ is leftmost and opens to the left. Since our definition of addition in elliptic curves uses lines and is purely geometric, the same goes for our reflected version, which tells us that it, too, is an abelian group, say, G'. All our other observations similarly hold, meaning that its branch H' is a subgroup, its oval is a coset of that subgroup, and the quotient G'/H' is the group $\mathbb{Z}_2$.

Figure 6.7 on p. 144 puts these groups, subgroups, and cosets together in a suggestive way as the skeleton of a torus. The two heavily drawn loops together form G, while the other two loops form G'. The outer, heavily drawn loop is the branch subgroup H of G, while the more lightly drawn loop on the left is H' and touches H at infinity, their mutual 0-element. The other two loops are the other two ovals in Figure 6.6 on p. 144 and are cosets of H and H'. To see this more easily, Figure 6.7 is redrawn as Figure 6.10 with additional labeling. Notice that the two heavy loops are disjoint, as are the two lightly drawn loops. In each case, the coset is disjoint from its subgroup.

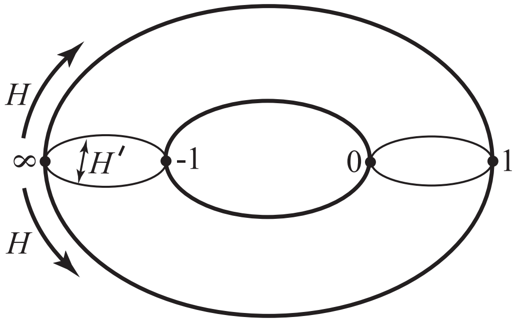

Figure 6.10. The two heavy loops form the group G. The outer heavy loop is the branch subgroup H of G. That branch's 0-element is ∞ in the drawing. The two more lightly drawn loops form the group G'. The left light loop is the branch subgroup H' of G', and that branch's 0-element is again ∞ in the drawing.

Yet another way of relating Figure 6.7 to a skeleton torus is depicted in Figure 6.11. In it, the coordinate pairs appearing in the top and bottom drawings are the endpoints of curves in the top drawing. The curves in the top drawing correspond to the line segments in the bottom drawing in an obvious way. In the bottom drawing, all four corners of the square are labeled $(0, 0)$, meaning they're identified to a single point. The two points $(1, 0)$ are likewise identified, as are similarly the two points $(0, 1)$. The light shading stands for a thin rubber sheet, and the usual way of forming a topological torus from the square by identifying edges automatically makes these identifications. Without any shading, the bottom drawing depicts the real projective curve with its branch subgroup loops, coset loops, and how these touch each other, just as the top picture does. But filling in the square does much more — it corresponds to extending the real projective curve to a complex projective curve. (Although it's customarily called a curve or complex curve, visually it's a real, 2-dimensional surface. This difference is an artifact of nomenclature history.) The locus of $y^2 = x^3 - x$ in $\mathbb{C}^2$, plus the point at infinity, is a closed surface in $\mathbb{P}^2(\mathbb{C})$ — a topological torus. The filled-in square in the bottom drawing has a natural group structure via vector addition in which coordinatewise addition is taken modulo 2, and there is a corresponding group structure in the surface extending that of the real curve: Addition via the connect-the-dots algorithm is defined geometrically just as in the real case, except now the lines are complex. So, for example, two distinct points in the vector space $\mathbb{C}^2$ determine a unique complex line through them.

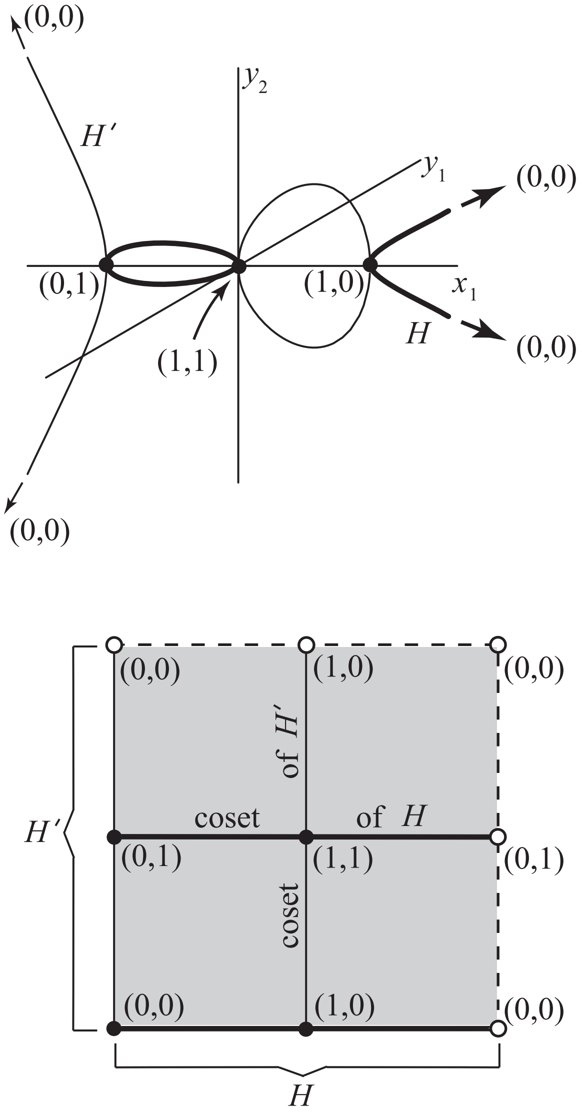

Figure 6.11. Another way our 3D slice relates to a torus skeleton.

6.6 Elliptic Curves with No Oval

Example 6.6.1. The branch and loop in Figure 6.5 are topologically just the loops we tend to draw when making an everyday sketch of a torus. But this chapter's title is "Every Real Elliptic Curve Lives in a Donut," and lots of elliptic curves in the real plane have only a branch and no oval. So, for example, what's the story about an elliptic curve such as $y^2 = x^3 + x$? Its plot in Figure 6.12 reveals no oval.

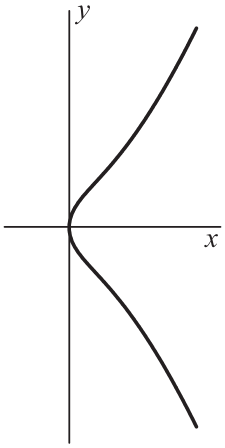

Figure 6.12. This curve $y^2 = x^3 + x = (x + i)x(x - i)$ is analogous to $y^2 = x^3 - x = (x + 1)x(x - 1)$. Every elliptic curve has a loop, but where is it?

If we play the same game as in the last two examples, we get a curve in the 3-dimensional slice $x_2 = 0$ consisting of the branch shown in Figure 6.12, plus its reflection (rotated by 90°) analogous to those examples. That is, the space curve we get is the union of two branches touching at $(0, 0, 0)$, one in the (x_1, y_1)-plane opening to the right, the other in the (x_1, y_2)-plane opening to the left. Now all of 4-space is covered by the parallel 3-slices $x_2 = r$ as r runs through $\mathbb{R}$, but seeing a torus by visualizing the union of the curves in these 3-slices doesn't sound very doable. It would be nice to have a single 3-slice showing loops the way Figure 6.6 does.

How can we get such a slice? A clue is revealed if we work by analogy. Factor the right-hand side of $y^2 = x^3 - x$ into $x^3 - x = (x+1)x(x-1)$. These three factors are 0 when $x = 0$ and ± 1 — the three points where all loops cross the x_1-axis. The loops/branches themselves are formed using x_1 as a parameter. Let's do an analogous thing for $y^2 = x^3 + x$: The right-hand side factors into $(x+i)x(x-i)$, and this is 0 when $x = 0$ and $\pm i$. The loops now cross the x_2-axis (rather than the x_1-axis) at three points, and the loops/branches are formed using x_2 as a parameter. As one example, at $x_2 = \frac{i}{2}$, y^2 is $x^3 + x$ evaluated at $\frac{i}{2}$, which is $\frac{3i}{8}$, and y itself works out to be

$$\pm\frac{\sqrt{3}(1+i)}{4}.$$

These y-values lie in the plane $y_1 = y_2$ in (x_2, y_1, y_2)-space. Similarly, at $x_2 = -\frac{i}{2}$, y^2 is $-\frac{3i}{8}$, and the y-values are

$$\pm\frac{\sqrt{3}(1-i)}{4}.$$

These lie in the plane $y_1 = -y_2$ in (x_2, y_1, y_2)-space. These two planes $y_1 = y_2$ and $y_1 = -y_2$ are perpendicular to each other, in the same sense that $y_1 = 0$ and $y_2 = 0$ are perpendicular in the previous two examples. It turns out that we get a picture looking much like Figure 6.6, just oriented differently in 4-space. ◇

In the previous section we obtained various group-theoretic results for elliptic curves having an oval. Using the ideas just above, we can replicate the geometric arguments of the previous section to get analogous group-theoretic results when there's no oval in (x_1, y_1)-space by using (x_2, y_1, y_2)-space as a 3D slice instead of (x_1, y_1, y_2)-space.