CHAPTER III

ELLIPTIC CURVE ALGORITHMS

3.1 Terminology and notation

For reference in the following sections, we collect here the notation, terminology and formulae concerning elliptic curves which we will use throughout this chapter.

An elliptic curve E defined over $\mathbb{Q}$ has an equation or *model* of the form

$$E: \qquad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.1.1}$$

where the coefficients $a_i \in \mathbb{Q}$. We call such an equation a *Weierstrass equation* for E, and denote this model by $[a_1, a_2, a_3, a_4, a_6]$. We say that (3.1.1) is *integral* or *defined over* $\mathbb{Z}$ if all the a_i are in $\mathbb{Z}$. From these coefficients we derive the auxiliary quantities

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, \\ b_4 &= a_1a_3 + 2a_4, \\ b_6 &= a_3^2 + 4a_6, \\ b_8 &= a_1^2a_6 - a_1a_3a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2, \end{aligned}$$

the *invariants*

$$\begin{aligned} c_4 &= b_2^2 - 24b_4, \\ c_6 &= -b_2^3 + 36b_2b_4 - 216b_6, \end{aligned}$$

the *discriminant*

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6,$$

and the *j-invariant*

$$j = c_4^3/\Delta,$$

which are related by the identities

$$4b_8 = b_2b_6 - b_4^2 \qquad \text{and} \qquad 1728\Delta = c_4^3 - c_6^2.$$

The discriminant Δ must be non-zero for the curve defined by equation (3.1.1) to be non-singular and hence an elliptic curve. The j-invariant is (as its name suggests) invariant under isomorphism; elliptic curves with the same j are called *twists*: they are isomorphic over an algebraic extension, but not necessarily over $\mathbb{Q}$. The invariants c_4 and c_6 are sufficient to determine E up to isomorphism (over $\mathbb{Q}$) since E is isomorphic to

$$Y^2 = X^3 - 27c_4X - 54c_6.$$

The most general isomorphism from E to a second curve E' given by an equation of the form (3.1.1), which we usually think of as a change of coordinates on E itself, is $T(r, s, t, u)$, given by

$$\begin{aligned} x &= u^2x' + r \\ y &= u^3y' + su^2x' + t \end{aligned} \tag{3.1.2}$$

where $r, s, t \in \mathbb{Q}$ and $u \in \mathbb{Q}^*$. The effect of $T(r, s, t, u)$ on the coefficients a_i is given by

$$\begin{aligned} ua_1' &= a_1 + 2s \\ u^2a_2' &= a_2 - sa_1 + 3r - s^2 \\ u^3a_3' &= a_3 + ra_1 + 2t \\ u^4a_4' &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st \\ u^6a_6' &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1 \end{aligned} \tag{3.1.3}$$

so that

$$u^4c_4' = c_4, \qquad u^6c_6' = c_6, \qquad u^{12}\Delta' = \Delta \quad \text{and} \quad j' = j.$$

The transformations $T(0, 0, 0, u)$ we will refer to as scaling transformations; these have the effect of dividing each coefficient a_i by u^i, and similarly for each of the other quantities, according to its weight. Here a_i, b_i and c_i have weight i, while Δ has weight 12 and j has weight 0. By applying $T(0, 0, 0, u)$ for suitable u we can always transform to an integral model; all the invariants are then integral, except (possibly) for j. Among such integral models, those for which the positive integer $|\Delta|$ is minimal are called *global minimal models* for E. We will give in the next section a simple algorithm for finding such a model, given the invariants c_4 and c_6 of any model. Clearly, isomorphisms between minimal models must have $u = \pm 1$ and $r, s, t \in \mathbb{Z}$. We may normalize so that $a_1, a_3 \in \{0, 1\}$ and $a_2 \in \{-1, 0, 1\}$, by suitable choice of s, r and t (in that order), as may be seen from (3.1.3). Such an equation will be called *reduced*, and it is not hard to show that it is unique: the only transformation other than the identity $T(0, 0, 0, 1)$ from a reduced model to any another reduced model is the transformation $T(0, -a_1, -a_3, -1)$, which takes any model to itself; this is just the negation map $(x, y) \mapsto (x, y - a_1x - a_3)$ from the curve to itself. Thus every elliptic curve E defined over $\mathbb{Q}$ has a *unique* reduced minimal model. This fact makes it very easy to recognize curves: in Table 1 we give the coefficients of such a model for each of the curves there.

Given integers c_4 and c_6, two questions arise: is there a curve over $\mathbb{Q}$ with these invariants, and is it minimal? Clearly we must have $c_4^3 - c_6^2 = 1728\Delta$ with $\Delta \neq 0$. A solution to the first problem is given by Kraus in Proposition 2 of [30], which states the following.

PROPOSITION 3.1.1. *Let c_4, c_6 be integers such that $\Delta = (c_4^3 - c_6^2)/1728$ is a non-zero integer. In order for there to exist an elliptic curve E with a model* (3.1.1) *defined over $\mathbb{Z}$ having invariants c_4 and c_6, it is necessary and sufficient that*

(1) $c_6 \not\equiv \pm 9 \pmod{27}$;

(2) *either* $c_6 \equiv -1 \pmod 4$, *or* $c_4 \equiv 0 \pmod{16}$ *and* $c_6 \equiv 0, 8 \pmod{32}$.

The conditions of Proposition 3.1.1 will be referred to as *Kraus's conditions*. If we are given integers c_4 and c_6 satisfying these conditions, we can recover the coefficients a_i of the reduced model of the curve with c_4 and c_6 as invariants, using the formulae already given in Chapter 2, Section 14, which we repeat here for convenience:

$$\begin{aligned} b_2 &= -c_6 \mod 12 \in \{-5, \dots, 6\}; \\ b_4 &= (b_2^2 - c_4)/24; \\ b_6 &= (-b_2^3 + 36b_2b_4 - c_6)/216; \\ a_1 &= b_2 \mod 2 \in \{0, 1\}; \\ a_3 &= b_6 \mod 2 \in \{0, 1\}; \\ a_2 &= (b_2 - a_1)/4; \\ a_4 &= (b_4 - a_1a_3)/2; \\ a_6 &= (b_6 - a_3)/4. \end{aligned}$$

To see this, we may assume that we are seeking coefficients of a reduced model; then $b_2 \in \{-4, -3, 0, 1, 4, 5\}$, and we have $-c_6 \equiv b_2^3 \equiv b_2 \pmod{12}$. The rest is easy; provided that c_4 and c_6 satisfy Kraus's conditions, all the divisions will be exact.

In the following section we answer the second question by giving an algorithm for computing the reduced coefficients of a *minimal* model for any curve E, given either integral invariants satisfying Kraus's conditions, or any integral model for E. We simply determine the maximal integer u such that $c_4' = c_4/u^4$ and $c_6' = c_6/u^6$ satisfy Kraus's conditions, and then compute the reduced coefficients a_i' from these. As with many questions concerning elliptic curves, most of the work goes into determining the powers of 2 and 3 which divide u.

We will assume without further discussion that on any given curve E, points may be added and multiples taken, using standard formulae. The Mordell–Weil group of all rational points on E will be denoted $E(\mathbb{Q})$ as usual. If n is a positive integer, we denote by $E(\mathbb{Q})[n]$ the subgroup of rational points of order dividing n, which is the kernel of the multiplication map from E to itself.

3.2 The Kraus–Laska–Connell algorithm and Tate's algorithm

In this section we give two algorithms. The first was originally given by Laska in [34], and finds a minimal model for a curve E, starting from an integral equation. Essentially the algorithm was to test all positive integers u such that $u^{-4}c_4$ and $u^{-6}c_6$ are integral, to see if they are the invariants of a curve defined over $\mathbb{Z}$. Using Kraus's conditions (see Proposition 3.1.1 above), this procedure can be simplified, since it is possible to compute in advance the exponent d_p of each prime p in the minimal discriminant, and hence compute u at the start. The usual formulae then give the coefficients a_i of the reduced model. Our formulation of the resulting algorithm over $\mathbb{Z}$ is similar to that given in [10], where more general rings are considered: in particular an explicit algorithm is given there for finding local minimal models over arbitrary number fields, and hence global minimal models where they exist. Over $\mathbb{Z}$, the algorithm is extremely simple.

In the pseudocode below,

`ord(p,n)` gives the power of the prime `p` which divides the non-zero integer `n`;

`floor(x)` gives the integral part of the real number `x`;

`a mod p` gives the residue of `a` modulo `p` lying in the range $-\frac{1}{2}p < a \leq \frac{1}{2}p$; in particular, when $p = 2$ or 3 this gives a residue in $\{0, 1\}$ or $\{-1, 0, 1\}$ respectively. Also `inv(a,p)` gives the inverse of `a` modulo `p`, assuming that `gcd(a,p)=1`.

The Laska–Kraus–Connell Algorithm

```
INPUT:       c4, c6 (integer invariants of an elliptic curve E).
OUTPUT:      a1, a2, a3, a4, a6 (coefficients of a reduced minimal model for E).
 1.  BEGIN
 2.  Δ=(c4^3-c6^2)/1728;
(Compute scaling factor u)
 3.  u = 1; g = gcd(c6^2,Δ);
 4.  p_list = prime_divisors(g);
 5.  FOR p IN p_list DO
 6.  BEGIN
 7.       d = floor(ord(p,g)/12);
 8.       IF p=2 THEN
 9.            a = c4/2^(4*d) mod 16; b = c6/2^(6*d) mod 32;
10.            IF (b mod 4 ≠ -1) AND NOT (a=0 AND (b=0 OR b=8))
```

```
11.           THEN d = d-1
12.           FI
13.      ELIF p=3 THEN IF ord(3,c6)=6*d+2 THEN d = d-1 FI
14.      FI;
15.      u = u*p^d
16.  END;
(Compute minimal equation)
17.  c4 = c4/u^4; c6 = c6/u^6;
18.  b2 = -c6 mod 12; b4 = (b2^2-c4)/24; b6 = (-b2^3+36*b2*b4-c6)/216;
19.  a1 = b2 mod 2;
20.  a3 = b6 mod 2;
21.  a2 = (b2-a1)/4;
22.  a4 = (b4-a1*a3)/2;
23.  a6 = (b6-a3)/4
24.  END
```

Next we turn to Tate's algorithm itself. The standard reference for this is Tate's 'letter to Cassels' [65], which appeared in the Antwerp IV volume [2]. There is also a full account in the second volume of Silverman's book [61, Section IV.9]. It may be applied to an integral model of a curve E and a prime p, to give the following data:

• The exponent f_p of p in the conductor N of E (see below);

• the Kodaira symbol of E at p, which classifies the type of reduction of E at p (see [47] or [61, Section IV.9]); these are: I_0 for good reduction; I_n ($n > 0$) for bad multiplicative reduction; and types I_n^*, II, III, IV, II*, III* and IV* for bad additive reduction.

• the local index $c_p = [E(\mathbb{Q}_p) : E^0(\mathbb{Q}_p)]$, where $E^0(\mathbb{Q}_p)$ is the subgroup of the group $E(\mathbb{Q}_p)$ of p-adic points of E, consisting of those points whose reduction modulo p is non-singular. (That this index is finite is implied by the correctness of the algorithm, as observed by Tate in [65].)

In addition, the algorithm detects whether the given model is non-minimal at p, and if so, returns a model which is minimal at p. Thus by applying it in succession with all the primes dividing the discriminant of the original model, one can compute a minimal model at the same time as computing the conductor and the other local reduction data. In practice this makes the Laska–Kraus–Connell algorithm redundant, though much simpler to implement and use if all one needs is the standard model for a curve E.

The conductor N of an elliptic curve E defined over $\mathbb{Q}$ is defined to be

$$N = \prod_p p^{f_p}$$

where $f_p = \mathrm{ord}_p(\Delta) + 1 - n_p$ and n_p is the number of irreducible components on the special fibre of the minimal Néron model of E at p. This Néron model is a more sophisticated object than we wish to discuss here (see [47] or [61] for details): one has to consider E as a scheme over $\mathrm{Spec}(\mathbb{Z}_p)$, and then resolve the singularity at p, to obtain a scheme whose generic fibre is $E/\mathbb{Q}_p$ and whose special fibre is a union of curves over $\mathbb{Z}/p\mathbb{Z}$. In terms of a minimal model for E over $\mathbb{Z}$, all may be computed very simply except when $p = 2$ or $p = 3$ as follows:

$f_p = 0$ if $p \nmid \Delta$;

$f_p = 1$ if $p \mid \Delta$ and $p \nmid c_4$ (then $n_p = \mathrm{ord}_p(\Delta)$);

$f_p \geq 2$ if $p \mid \Delta$ and $p \mid c_4$; moreover, $f_p = 2$ in this case when $p \neq 2, 3$.

To obtain the value of f_p in the remaining cases, and to obtain the Kodaira symbol and the local index c_p, we use Tate's algorithm itself.

In [65], the algorithm is given for curves defined over an arbitrary discrete valuation ring. To apply it to a curve defined over the ring of integers R of a number field K at a prime ideal $\mathfrak{p}$, one would in general have to work in the localization of R at $\mathfrak{p}$; here we can work entirely over $\mathbb{Z}$, since $\mathbb{Z}$ is a principal ideal domain. We have added to the presentation in [65] the explicit coordinate transformations $T(r,s,t,u)$ which are required during the course of the algorithm to achieve divisibility of the coefficients a_i by various power of p. In practice one would ignore the transformations which had taken place while processing each p, unless a scaling by p had taken place on discovering that the model was non-minimal. The most complicated part of the algorithm is the branch for reduction type I_m^*, where one successively refines the model p-adically until certain auxiliary quadratics have distinct roots modulo p. This requires careful book-keeping. The presentation given here closely follows our own implementation of the algorithm, which in turn owes much to an earlier Fortran program written by Pinch. The following sub-procedures are used:

`compute_invariants` computes the b_i, c_i and Δ from the coefficients a_i. Note that c_4, c_6 and Δ do not change unless a scaling is required, since all other transformations have $u = 1$.

`transcoord(r,s,t,u)` applies the coordinate transformation formulae of the previous section to obtain new values for the a_i and other quantities. All calls to this procedure have $u = 1$ except when rescaling a non-minimal equation. In each case we first compute suitable values of r, s and t; usually this requires a separate branch if $p = 2$ or $p = 3$.

`quadroots(a,b,c,p)` returns `TRUE` if the quadratic congruence $ax^2 + bx + c \equiv 0 \pmod{p}$ has a solution, and `FALSE` otherwise. This is used in determining the value of the index c_p.

`nrootscubic(b,c,d,p)` returns the number of roots of the cubic congruence $x^3 + bx^2 + cx + d \equiv 0 \pmod{p}$.

Tate's Algorithm

```
INPUT:        a1, a2, a3, a4, a6 (integer coefficients of E); p (prime).
OUTPUT:       Kp   (Kodaira symbol)
              fp   (Exponent of p in conductor)
              cp   (Local index)
 1.  BEGIN
 2.  compute_invariants(b2,b4,b6,b8,c4,c6,Δ);
 3.  n = ord(p,Δ);
(Test for type I0)
 4.  IF n=0 THEN Kp = "I0"; fp = 0; cp = 1; EXIT FI;
(Change coordinates so that p | a3, a4, a6)
 5.  IF p=2 THEN
 6.       IF p|b2
 7.       THEN r = a4 mod p; t = r*(1+a2+a4)+a6 mod p
 8.       ELSE r = a3 mod p; t = r+a4 mod p
 9.       FI
10.  ELIF p=3 THEN
11.       IF p|b2 THEN r = -b6 mod p ELSE r = -b2*b4 mod p FI;
12.       t = a1*r+a3 mod p
13.  ELSE
14.       IF p|c4 THEN r = -inv(12,p)*b2 ELSE r = -inv(12*c4,p)*(c6+b2*c4) FI;
15.       t = -inv(2,p)*(a1*r+a3);
16.       r = r mod p; t = t mod p
17.  FI;
```

```
18.  transcoord(r,0,t,1);
```
(*Test for types I_n, II, III, IV)*
```
19.  IF p∤c4 THEN
20.       IF quadroots(1,a1,-a2,p) THEN cp = n ELIF 2|n THEN cp = 2 ELSE cp = 1 FI;
21.       Kp = "In"; fp = 1; EXIT
22.  FI;
23.  IF p^2∤a6 THEN Kp = "II"; fp = n; cp = 1; EXIT;
24.  IF p^3∤b8 THEN Kp = "III"; fp = n-1; cp = 2; EXIT;
25.  IF p^3∤b6 THEN
26.       IF quadroots(1,a3/p,-a6/p^2,p) THEN cp = 3 ELSE cp = 1 FI;
27.       Kp = "IV"; fp = n-2; EXIT
28.  FI;
```
(*Change coordinates so that $p \mid a_1, a_2$; $p^2 \mid a_3, a_4$; $p^3 \mid a_6$)*
```
29.  IF p=2
30.  THEN s = a2 mod 2; t = 2*(a6/4 mod 2)
31.  ELSE s = -a1*inv(2,p); t = -a3*inv(2,p)
32.  FI;
33.  transcoord(0,s,t,1);
```
(*Set up auxiliary cubic $T^3 + bT^2 + cT + d$)*
```
34.  b = a2/p; c = a4/p^2; d = a6/p^3;
35.  w = 27*d^2-b^2*c^2+4*b^3*d-18*b*c*d+4*c^3;
36.  x = 3*c-b^2;
```
(*Test for distinct roots: type I_0^*)*
```
37.  IF p∤w THEN Kp = "I*0"; fp = n-4; cp = 1+nrootscubic(b,c,d,p); EXIT
```
(*Test for double root: type I_m^*)*
```
38.  ELIF p∤x THEN
```
(*Change coordinates so that the double root is $T \equiv 0$)*
```
39.       IF p=2 THEN r = c ELIF p=3 THEN r = b*c ELSE r = (b*c-9*d)*inv(2*x,p) FI;
40.       r = p*(r mod p);
41.       transcoord(r,0,0,1);
```
(*Make a_3, a_4, a_6 repeatedly more divisible by p)*
```
42.       m = 1; mx = p^2; my = p^2; cp = 0;
43.       WHILE cp=0 DO
44.       BEGIN
45.            xa2 = a2/p; xa3 = a3/my; xa4 = a4/(p*mx); xa6 = a6/(mx*my);
46.            IF p∤(xa3^2+4*xa6) THEN
47.                 IF quadroots(1,xa3,-xa6,p) THEN cp = 4 ELSE cp = 2 FI
48.            ELSE
49.                 IF p=2 THEN t = my*xa6 ELSE t = my*((-xa3*inv(2,p)) mod p) FI;
50.                 transcoord(0,0,t,1);
51.                 my = my*p; m = m+1;
52.                 xa2 = a2/p; xa3 = a3/my; xa4 = a4/(p*mx); xa6 = a6/(mx*my);
53.                 IF p∤(xa4^2-4*xa2*xa6) THEN
54.                      IF quadroots(xa2,xa4,xa6,p) THEN cp = 4 ELSE cp = 2 FI
55.                 ELSE
56.                      IF p=2 THEN r = mx*(xa6*xa2 mod 2)
```

```
57.                     ELSE r = mx*(-xa4*inv(2*xa2,p) mod p)
58.                     FI;
59.                     transcoord(r,0,0,1);
60.                     mx = mx*p; m = m+1
61.                 FI
62.             FI
63.         END;
64.         fp = n-m-4; Kp = "I*m"; EXIT
65.   ELSE
```

(*Triple root case: types II*, III*, IV* or non-minimal)*
(*Change coordinates so that the triple root is* $T \equiv 0$*)*

```
66.         IF p=3 THEN rp = -d ELSE rp = -b*inv(3,p) FI;
67.         r = p*(rp mod p);
68.         transcoord(r,0,0,1);
69.         x3 = a3/p^2; x6 = a6/p^4;
```

(*Test for type IV*)*

```
70.         IF p|(x3^2+4*x6) THEN
71.             IF quadroots(1,x3,-x6,p) THEN cp = 3 ELSE cp = 1 FI;
72.             Kp = "IV*"; fp = n-6; EXIT
73.         ELSE
```

(*Change coordinates so that* $p^3 \mid a_3$*,* $p^5 \mid a_6$*)*

```
74.             IF p=2 THEN t = x6 ELSE t = x3*inv(2,p) FI;
75.             t = -p^2*(t mod p);
76.             transcoord(0,0,t,1);
```

(*Test for types III*, II*)*

```
77.             IF p^4|a4 THEN Kp = "III*"; fp = n-7; cp = 2; EXIT
78.             ELIF p^6|a6 THEN Kp = "II*"; fp = n-8; cp = 1; EXIT
79.             ELSE
```

(*Equation non-minimal: divide each* a_i *by* p^i *and start again)*

```
80.                 transcoord(0,0,0,p); restart
81.             FI
82.   FI
83.   END
```

In Table 1 we will give the local reduction data for each curve at each 'bad' prime (dividing the discriminant of the minimal model). We also give the factorization of the minimal discriminant and of the denominator of j, as in the earlier tables. To save space we omit the c_4 and c_6 invariants, which are easily computable from the coefficients a_i.

3.3 Computing the Mordell–Weil group I: finding torsion points

In this and the next three sections we will discuss the question of determining the Mordell–Weil group $E(\mathbb{Q})$ of rational points on an elliptic curve E defined over $\mathbb{Q}$. This group is finitely generated, by Mordell's Theorem, and hence has the structure

$$E(\mathbb{Q}) = T \times F$$

where T is the finite torsion subgroup $E(\mathbb{Q})_{\text{tors}}$ of $E(\mathbb{Q})$ consisting of the points of finite order, and F is free abelian of some rank $r \geq 0$:

$$F \cong \mathbb{Z}^r.$$

The problem of computing $E(\mathbb{Q})$ thus subdivides into several parts:

- computing the torsion T;
- computing the rank r;
- finding r independent points of infinite order;
- computing a $\mathbb{Z}$-basis for the free part F.

A related task is to compute the regulator $R(E(\mathbb{Q}))$ (defined below); for this and for the latter two steps we will also need to compute the canonical height $\hat{h}(P)$ of points $P \in E(\mathbb{Q})$, and hence the height pairing $\hat{h}(P, Q)$.

In this section we will treat the easiest of these problems, that of finding the torsion points. In fact, these can be found as a byproduct of the more general search for points on the curve, since their naive height can be bounded (see the remark before Lemma 3.5.2). However, it is also useful to have a self-contained method for determining the torsion.

Using the fact that $E(\mathbb{R})$ is isomorphic either to the circle group S^1 (when $\Delta < 0$) or to $S^1 \times C_2$ (when $\Delta > 0$), where C_k denotes a cyclic group of order k, together with the fact that all finite subgroups of S^1 are cyclic, we see that T is isomorphic either to C_k or to $C_{2k} \times C_2$ for some $k \geq 1$, the latter only being possible when Δ is positive. The number of possible values of k is finite: by a theorem of Mazur [39],[40], a complete list of possible structures of T is

$$\begin{array}{rll} C_k & \text{for} & 1 \leq k \leq 10 \quad \text{or} \quad k = 12; \\ C_{2k} \times C_2 & \text{for} & 1 \leq k \leq 4. \end{array}$$

To determine the torsion subgroup of an elliptic curve defined over $\mathbb{Q}$, we may use a form of the Lutz–Nagell Theorem. (The situation is more complicated over number fields other than $\mathbb{Q}$, on account of the ramified primes.) The first step is to find a model for the curve in which all torsion points are integral. For this it suffices to complete the square (if necessary) to eliminate the xy and y terms, at the expense of a scaling by $u = 2$. Then for $P = (x, y)$ a torsion point, we can use the fact that both P and $2P$ are integral to bound y. For the first step, the following result may be found in [33, Section III.1] and [28, Theorem 5.1]. The original form of this result, due independently to Lutz [36] and Nagell [46], was for curves of the form $y^2 = x^3 + ax + b$, with no x^2 term. While such an equation may be obtained by completing the cube, this would involve a further scaling of coordinates, and so would lead to larger numbers. If $a_1 = a_3 = 0$ we can apply the following result directly; otherwise, put $a = b_2$, $b = 8b_4$ and $c = 16b_6$.

PROPOSITION 3.3.1. *Let E be an elliptic curve defined over $\mathbb{Q}$, given by an equation*

$$y^2 = f(x) = x^3 + ax^2 + bx + c \tag{3.3.1}$$

where $a, b, c \in \mathbb{Z}$. If $P = (x, y) \in E(\mathbb{Q})$ has finite order, then $x, y \in \mathbb{Z}$.

Next we bound the y coordinate of a torsion point $P = (x, y)$ (see [33, Theorem 1.4]).

PROPOSITION 3.3.2. *Let E be as in (3.3.1). If $P = (x_1, y_1)$ has finite order in $E(\mathbb{Q})$ then either $y_1 = 0$ or $y_1^2 \mid \Delta_0$, where*

$$\Delta_0 = 27c^2 + 4a^3c + 4b^3 - a^2b^2 - 18abc.$$

PROOF. If $2P = 0$ then $y_1 = 0$, since $-P = (x_1, -y_1)$. Otherwise $2P = (x_2, y_2)$ with $x_2, y_2 \in \mathbb{Z}$ by Proposition 3.3.1. Using the addition formula on E we find that $2x_1 + x_2 = m^2 - a$ where $m = f'(x_1)/2y_1$ is the slope of the tangent to E at P. Hence $m \in \mathbb{Z}$, so that $y_1 \mid f'(x_1)$. Using $y_1^2 = f(x_1)$, this implies that $y_1^2 | \Delta_0$, since

$$\Delta_0 = (-27f(x) + 54c + 4a^3 - 18ab)f(x) + (f'(x) + 3b - a^2)f'(x)^2. \quad \square$$

This gives us a finite number of values of y to check; for each, we attempt to solve the cubic for $x \in \mathbb{Z}$, to obtain all torsion points on E. Note that we are actually determining all points P such that both P and $2P$ are integral (in the possibly scaled model for E), which includes all torsion points, but may also include points of infinite order. To determine whether a given integral point has finite or infinite order, we simply compute multiples mP successively until either $mP = 0$, in which case P has order m, or mP is not integral, in which case P has infinite order. This does not take long, as the maximum possible order for a torsion point is 12 by Mazur's theorem. If we find points of infinite order at this stage we keep a note of them for later use (see Section 3.5).

The quantity Δ_0 is related to the discriminant Δ of the curve (3.3.1) by $\Delta = -16\Delta_0$. If this is large, there may be many values of y_0 to check when we apply the preceding Proposition to determine the torsion on a given curve. It is possible to save time by using a further result, which states that for an odd prime p of good reduction (that is, $p \nmid 2\Delta$), the reduction map from $E(\mathbb{Q})_{\text{tors}}$ to $E(\mathbb{Z}/p\mathbb{Z})$ is injective. For more details, and worked examples, see either [58, Section VIII.7] or [28, Section V.1].

If we want to know the structure of T and not just its order, note that from Mazur's theorem the only ambiguous cases are when T has order $4k = 4$, 8 or 12 and $\Delta > 0$; we can always tell apart the groups C_{4k} and $C_2 \times C_{2k}$ as the former has only one element of order 2 while the latter has three, and this number is the number of rational (integer) roots of $f(x)$.

To solve the cubic equations $f(x) = y^2$ for x, given y, we use the classical formula of Cardano (see any algebra textbook) to find the complex roots (which we also need in computing the periods in section 3.7 below), and if any of these are real and close to integers we check them using exact integer arithmetic. Testing all divisors of the constant term can be too time-consuming, as it involves factorization of the numbers $y^2 - c$ which may be very large.

Here is the algorithm in pseudocode; for simplicity we only give it for curves with no xy or y terms; in the general case, one works internally with points on a scaled model (including the calculation of the order), converting back to the original model on output. Since we know in advance that no point will have order greater than 12, when computing the order of a point we simply use repeated addition until we reach a non-integral point or the identity `O`. The subroutine `order(P)` returns 0 for a point of infinite order. Also: `square_part(`Δ`)` returns the largest integer whose square divides Δ; `integer_roots` returns a list of the integer roots of a cubic with integral coefficients; and `integral(x)` tests whether its (rational) argument is integral.

Algorithm for finding all torsion points

```
INPUT:        a,b,c (integer coefficients of a nonsingular cubic).
OUTPUT:       A list of all torsion points on y^2=x^3+ax^2+bx+c, with orders.
1.  BEGIN
2.  Δ=27*c^2+4*a^3*c+4*b^3-a^2*b^2-18*a*b*c;
3.  y_list=positive_divisors(square_part(Δ)) ∪ {0};
4.  FOR y IN y_list DO
5.  BEGIN
```

```
 6.       x_list=integer_roots(x^3+a*x^2+b*x+c-y^2);
 7.       FOR x IN x_list DO
 8.       BEGIN
 9.            P=point(x,y);
10.            n=order(P);
11.            IF n>0 THEN OUTPUT P,n FI
12.       END
13.  END
14.  END
```

(*Subroutine to compute order of a point)*

```
SUBROUTINE order(P)
 1.  BEGIN
 2.  n=1; Q=P;
 3.  WHILE integral(x(Q)) AND Q≠O DO
 4.  BEGIN
 5.       n = n+1; Q = Q+P
 6.  END;
 7.  IF Q≠O THEN n=0 FI;
 8.  RETURN n
 9.  END
```

3.4 Heights and the height pairing

In this section we will show how to compute the canonical height $\hat{h}(P)$ of a point $P \in E(\mathbb{Q})$, and hence the height pairing

$$\hat{h}(P,Q) = \frac{1}{2}(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)).$$

We will use this in the following section to find dependence relations among finite sets of points of infinite order, when we are computing a $\mathbb{Z}$-basis $\{P_1, \ldots, P_r\}$ for the free abelian group $E(\mathbb{Q})/T$. Also, the regulator $R(E)$ is given by the determinant of the height pairing matrix:

$$R(E) = \left|\det(\hat{h}(P_i, P_j))\right|.$$

The canonical height $\hat{h}$ is a real-valued quadratic form on $E(\mathbb{Q})$. It differs by a bounded amount (with a bound dependent on E but not on the point P) from the naive or Weil height $h(P)$. For a point $P = (x, y) = (a/c^2, b/c^3) \in E(\mathbb{Q})$ with $a, b, c \in \mathbb{Z}$ and $\gcd(a, c) = 1 = \gcd(b, c)$, the latter is defined to be

$$h(P) = \log \max\{|a|, c^2\}.$$

Now the canonical height may be defined as $\hat{h}(P) = \lim_{n\to\infty} 4^{-n} h(2^n P)$, but this is not practical for computational purposes. For the theory of heights on elliptic curves, see [58, Chapter VIII]. Later (in the next section) we will need an explicit bound on the difference between $\hat{h}(P)$ and $h(P)$.

The height algorithms in this section are taken from Silverman's paper [59]. The *global height* $\hat{h}(P)$ is defined as a sum of *local heights*:

$$\hat{h}(P) = \sum_{p \le \infty} \hat{h}_p(P). \tag{3.4.1}$$

Here the sum is over all finite primes p and the 'infinite prime' ∞ coming from the real embedding of $\mathbb{Q}$. (Over a general number field, there would in general be several of these infinite primes, including complex ones, and the local heights need to be multiplied by certain multiplicities: see [59]).

A remark about normalization[1]: the canonical height must be suitably normalized. In the literature there are two normalizations used, one of which is double the other and is the one appropriate for the Birch–Swinnerton-Dyer conjecture (resulting in a regulator 2^r times as large). In Silverman's paper he uses the other (smaller) normalization. Thus all the formulae here are double those in the paper [59].

The following proposition, which is Theorem 5.2 of [59] (for curves over general number fields) specialized to the case of a curve defined over $\mathbb{Q}$, also applies to a curve defined over $\mathbb{Q}_p$ and to a point $P = (x, y) \in E(\mathbb{Q}_p)$. In the proposition, we refer to the functions ψ_2 and ψ_3 defined on E by

$$\psi_2(P) = 2y + a_1x + a_3, \qquad \text{and} \qquad \psi_3(P) = 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8;$$

thus, ψ_2 vanishes at the 2-torsion points of E and ψ_3 at the 3-torsion.

PROPOSITION 3.4.1. *Let E be an elliptic curve defined over $\mathbb{Q}$ given by a standard Weierstrass equation* (3.1.1) *which is minimal at p, and let $P = (x, y) \in E(\mathbb{Q})$.*

(a) *If*

$$\mathrm{ord}_p(3x^2 + 2a_2x + a_4 - a_1y) \le 0 \qquad \textit{or} \qquad \mathrm{ord}_p(2y + a_1x + a_3) \le 0$$

then

$$\hat{h}_p(P) = \max\{0, -\mathrm{ord}_p(x)\} \log p.$$

(b) *Otherwise, if* $\mathrm{ord}_p(c_4) = 0$ *then set* $N = \mathrm{ord}_p(\Delta)$ *and* $M = \min\{\mathrm{ord}_p(\psi_2(P), \frac{1}{2}N\}$*; then*

$$\hat{h}_p(P) = \frac{M(M-N)}{N} \log p.$$

(c) *Otherwise, if* $\mathrm{ord}_p(\psi_3(P)) \ge 3\mathrm{ord}_p(\psi_2(P))$ *then*

$$\hat{h}_p(P) = -\frac{2}{3}\mathrm{ord}_p(\psi_2(P)) \log p.$$

(d) *Otherwise*

$$\hat{h}_p(P) = -\frac{1}{4}\mathrm{ord}_p(\psi_3(P)) \log p.$$

The first case in Proposition 3.4.1 covers primes p where the point P has good reduction (including all primes where E has good reduction, as well as those where the reduced curve is singular but P does not reduce to the singular point). In the other three cases, P has singular reduction, and the reduction of E at p is multiplicative, additive of types IV or IV*, and additive of types III, III* and I$_m^*$ respectively.

Hence for each point P, the local height $\hat{h}_p(P) = 0$ if p divides neither the discriminant Δ nor c, where c^2 is the denominator of the x-coordinate of the point P. In all cases, $\hat{h}_p(P)$ is a rational multiple of $\log(p)$. The total contribution from the primes dividing c in the global height $\hat{h}(P)$ is therefore (from case (a) of the Proposition) simply $2\log(c)$, and we have

[1] I am grateful to Gross for explaining this to me, after I found that apparently the two sides of the Birch–Swinnerton-Dyer conjecture disagreed by a factor of 2^r!

the following formula, better for practical computation than (3.4.1) since we do not have to factorize c:

$$\hat{h}(P) = \hat{h}_\infty(P) + 2\log(c) + \sum_{p|\Delta, p\nmid c} \hat{h}_p(P). \tag{3.4.2}$$

This formula appears in [62], where it is shown how to compute $\hat{h}(P)$ using little (or no) factorization of Δ, which can be useful in certain situations. We refer the reader to [62] for details.

An algorithm for computing the local height at a finite prime p is given by the following:

Silverman's algorithm for computing local heights: finite primes

```
INPUT:   a1, a2, a3, a4, a6 (integer coefficients of a minimal model for E).
         x,y (rational coordinates of a point P on E).
         p (a prime).
OUTPUT:  the local height of P at p.
 1.  BEGIN
 2.  compute_invariants(b2,b4,b6,b8,c4,Δ);
 3.  N = ord(p,Δ);
 4.  A = ord(p,3*x^2+2*a2*x+a4-a1*y);
 5.  B = ord(p,2*y+a1*x+a3);
 6.  C = ord(p,3*x^4+b2*x^3+3*b4*x^2+3*b6*x+b8);
 7.  M = min(B,N/2);
 8.  IF A ≤ 0 OR B ≤ 0 THEN L = max(0,-ord(p,x))
 9.  ELSE IF ord(p,c4)=0 THEN L = M*(M-N)/N
10.  ELSE IF C ≥ 3*B THEN L = -2*B/3
11.  ELSE L = -C/4
12.  FI;
13.  RETURN L*log(p)
14.  END
```

We must also compute the local component of the height at the infinite prime, $\hat{h}_\infty(P)$. The method here originated with Tate, but was amended by Silverman in [59] to improve convergence, and to apply also to complex valuations. Tate in [66] expressed $\hat{h}_\infty(P)$ as a series

$$\hat{h}_\infty(P) = \log|x| + \frac{1}{4}\sum_{n=0}^{\infty} 4^{-n} c_n$$

where the coefficients c_n are bounded provided that no point on $E(\mathbb{R})$ has x-coordinate zero. Of course, over $\mathbb{R}$ one can shift coordinates to ensure that this condition holds, but the resulting series can have poor convergence properties, and this trick will not work over $\mathbb{C}$. Silverman's solution is to use alternately the parameters x and $x' = x + 1$, switching between them (and between the two associated series c_n and c'_n) whenever $|x|$ or $|x'|$ becomes small (less than $1/2$). The series of coefficients c_n is obtained by repeated doubling of the point P, working with $t = 1/x$ or $t' = 1/x'$ as local parameter. The result is a new series of the above type in which the error in truncating before the Nth term is $O(4^{-N})$, with an explicit constant. In fact (see [59, Theorem 4.2]) the error is less than $\frac{1}{2}10^{-d}$, giving a result correct to d decimal places, if

$$N \geq \frac{5}{3}d + \frac{1}{2} + \frac{3}{4}\log(7 + \frac{4}{3}\log H + \frac{1}{3}\log\max\{1, |\Delta|^{-1}\})$$

where

$$H = \max\{4, |b_2|, 2|b_4|, 2|b_6|, |b_8|\}.$$

The last term vanishes for curves defined over $\mathbb{Z}$, since then we have $|\Delta| > 1$.

In the algorithm which we now give, the quantities `b2', b4', b6'` and `b8'` are those associated with the shifted model of E with $x' = x+1$; the switching flag `beta` indicates which model we are currently working on; `mu` holds the current partial sum; `f` holds the negative power of 4.

Silverman's algorithm for computing local heights: real component

---

```
INPUT:   a1, a2, a3, a4, a6 (integer coefficients of a minimal model for E).
         x (x-coordinate of a point P on E).
         d (number of decimal places required).
OUTPUT:  the real local height of P.

 1.  BEGIN
 2.  compute_invariants(b2,b4,b6,b8);
 3.  H = max(4,|b2|,2*|b4|,2*|b6|,|b8|);
 4.  b2' = b2-12; b4' = b4-b2+6; b6' = b6-2*b4+b2-4; b8' = b8-3*b6+3*b4-b2+3;
 5.  N = ceiling((5/3)*d + (1/2) + (3/4)*log(7+(4/3)*log(H)));
 6.  IF |x|<0.5 THEN t = 1/(x+1); beta = 0 ELSE t = 1/x; beta = 1 FI;
 7.  mu = -log|t|; f = 1;
 8.  FOR n = 0 TO N DO
 9.  BEGIN
10.       f = f/4;
11.       IF beta=1 THEN
12.            w = b6*t^4+2*b4*t^3+b2*t^2+4*t;
13.            z = 1-b4*t^2-2*b6*t^3-b8*t^4;
14.            zw = z+w
15.       ELSE
16.            w = b6'*t^4+2*b4'*t^3+b2'*t^2+4*t;
17.            z = 1-b4'*t^2-2*b6'*t^3-b8'*t^4;
18.            zw = z-w
19.       FI;
20.       IF |w| ≤ 2*|z|
21.       THEN mu = mu+f*log|z|; t = w/z
22.       ELSE mu = mu+f*log|zw|; t = w/zw; beta = 1-beta
23.       FI
24.  END;
25.  RETURN mu
26.  END
```

Finally, to compute the global height $\hat{h}(P)$, we simply add to the infinite local height $\hat{h}_\infty(P)$ the finite local heights $\hat{h}_p(P)$ for all primes p dividing either Δ or the denominator of $x(P)$. Using (3.4.2) this leads to the following algorithm.

Algorithm for computing global canonical heights

```
INPUT:   a1, a2, a3, a4, a6 (integer coefficients of a minimal model for E).
         P=(x,y) (a rational point P on E).
OUTPUT:  the global canonical height ĥ(P) of P.
 1.  BEGIN
 2.  Δ= discr(a1,a2,a3,a4,a6);
 3.  d = denom(x);
 4.  h = real_height(P) + log(d);
 5.  p_list = prime_divisors(Δ);
 6.  FOR p IN p_list DO
 7.  BEGIN
 8.       IF p∤d THEN h = h + local_height(p,P) FI
 9.  END;
10.  RETURN h
11.  END
```

3.5 The Mordell–Weil group II: generators

In this section we will show how we look for rational points of infinite order on an elliptic curve E. In compiling the tables, we usually knew the rank r in advance so that we knew how many independent points to expect to find (and only looked for such points when we knew that $r > 0$); however, this procedure is also useful as an open-ended search when we do not know the rank, as obviously it can provide us with a lower bound for r.

The procedure divides into two parts. First, we have a searching routine which looks for points up to some bound on the naive height (equivalently, some bound on the numerator and denominator of the x-coordinate). As this routine finds points, it gives them to the second routine, which has at each stage a $\mathbb{Z}$-basis for a subgroup A of $E(\mathbb{Q})/T$: initially $A = 0$. This second routine uses the height pairing to determine one of three possibilities: the new point P may be independent of those already found and can then be added to our cumulative list of independent points; the rank of A is thus increased by 1. Secondly, P may be an integral combination of the current basis (modulo torsion) and can then be ignored. Finally, if a multiple kP of P is an integral combination of the current basis for some $k > 1$, we can find a basis for a new subgroup A which contains the old A with index k. Even when we know the rank r in advance, we do not stop as soon as we have a subgroup A of rank r, since A might still have finite index in $E(\mathbb{Q})/T$. To close this final gap we use explicit bounds for the difference between the naive and canonical heights, such as Silverman's result (Proposition 3.5.1) below.

The algorithm we use for the second procedure is a very general one, which can be used in many other similar situations; for example, as part of an algorithm for finding the unit group of a number field, where the first routine somehow finds units. Our algorithm is essentially the same as the 'Algorithm for enlarging sublattices' in the book by Pohst and Zassenhaus [50, Chapter 3.3].

A rational point P on E (given by a standard Weierstrass equation) may be written uniquely as $P = (x, y) = (a/c^2, b/c^3)$ with integers a, b, and c satisfying $\gcd(a, c) = \gcd(b, c) = 1$ and $c \geq 1$. The naive or Weil height of P is $h(P) = \log \max\{|a|, c^2\}$. Initially, we find the point of order 2 in $E(\mathbb{R})$ with minimal x-coordinate x_0; this gives a lower bound for the x-coordinates of all real points on E. We then search for points P with naive height up to some bound B by looping through positive integers $c \leq \exp(B/2)$ and through a coprime to c in the range[2]

[2]If $E(\mathbb{R})$ has three points of order two, with x-coordinates $x_0 < x_1 < x_2$, then we also omit those a for which $c^2 x_1 < a < c^2 x_2$.

$\max\{c^2x_0, -\exp(B)\} \le a \le \exp(B)$. Given a and c, we attempt to solve the appropriate quadratic equation for $b \in \mathbb{Z}$. To speed up this procedure, we use a quadratic sieve: for each denominator c we precompute for about 10 auxiliary sieving primes p the residue classes modulo p to which a must belong if the equation for b is to be soluble modulo p. Each candidate value of a can then first be checked to see if it is admissible modulo each sieving prime before the more time-consuming step of attempting to solve for b. This improvement to the search results in a major time saving in most cases, though for most of the curves in our tables on which we expected to find points of infinite order, such a point was found very quickly anyway. (In some cases we had already found such a point during the search for torsion points.) In practice it may be better to use composite moduli for the sieving.

Each point P found by this search is passed to the second procedure, which tests whether it has infinite order, discarding it if not. At the general stage we will have k independent points P_i for $1 \le i \le k$ (initially $k = 0$) which generate a subgroup A of rank k, and will have stored the $k \times k$ height pairing matrix $M = (h(P_i, P_j))$ and its determinant R. Now we set $P_{k+1} = P$ and compute $\hat{h}(P_i, P_{k+1})$ for $i \le k+1$ to obtain a new height pairing matrix of order $k+1$. If the determinant of this new matrix is non-zero, the new point is independent of the previous ones and we add it to the current list of generators, increment k, replace R by the new determinant, and go on with the point search. If the new determinant is zero, however, we use the values $h(P_i, P)$ to express P_{k+1} as a linear combination of the P_i for $i \le k$, with approximate real coefficients: in fact we have

$$a_1P_1 + a_2P_2 + \ldots + a_kP_k + a_{k+1}P_{k+1} = 0 \quad \text{(modulo torsion)}$$

where for $1 \le i \le k+1$ the coefficient a_i is the $(i, k+1)$ cofactor of the enlarged matrix, which we will have stored during the computation of the new determinant. In particular, a_{k+1} is (up to sign) the previous value of R, and hence is non-zero. Next we find rational approximations to these floating-point coefficients a_i (using continued fractions, or MLLL if available), and clear denominators to obtain a new equation of the same form with coprime integer coefficients a_i, which we can check holds exactly. In this relation we still have $a_{k+1} \ne 0$ (the first k points are independent). The simplest case now is when $a_{k+1} = \pm 1$, for then P_{k+1} is redundant and can be discarded. Similarly, if $a_i = \pm 1$ for some $i \le k$, then we may discard P_i, replacing it by P_{k+1}, and gaining index $|a_{k+1}|$. In general let a_i be the minimal non-zero coefficient (in absolute value); if $|a_i| > 1$, we find a coefficient a_j not divisible by a_i (which must exist since the coefficients are coprime) and write $a_j = a_iq + b$ where $0 < b < |a_i|$. Now since

$$a_iP_i + a_jP_j = a_iP_i + (a_iq + b)P_j = a_i(P_i + qP_j) + bP_j$$

we may replace the generator P_i by $P_i + qP_j$, replace the coefficient a_j by b (which is smaller than $|a_i|$), and replace i by j. After a finite number of steps we obtain a minimal coefficient $a_i = 1$ and can discard the current generator P_i, leaving a new set of k independent generators which generate a group larger than before by a finite index equal to the original value of $|a_{k+1}|$.

In this way, we will be able to find a $\mathbb{Z}$-basis for the subgroup A of the Mordell–Weil group (modulo torsion) which is generated by the points of naive height less than the bound B. Often we know the rank r of our curve in advance, so that we can increase B until A has rank r. Then A has finite index in $E(\mathbb{Q})$, and we must enlarge it to give the whole of $E(\mathbb{Q})$. There are various methods one can use here, all of which rely on having explicit bounds for the difference between the naive and canonical heights on the curve E. The simplest general bound here is a result of Silverman (see [60]). One can certainly often obtain better bounds for individual curves, and there are also more complicated results which apply in general and which usually give much better bounds, such as the main result of [57].

For simplicity we will only give Silverman's version of the bound. In the following proposition,[3] the height of a rational number a/b with $\gcd(a,b)=1$ is $h(a/b)=\log\max\{|a|,|b|\}$, and $\log^+(x)=\log\max\{1,|x|\}$ for $x\in\mathbb{R}$.

PROPOSITION 3.5.1. *Let E be an elliptic curve defined by a standard Weierstrass equation over $\mathbb{Z}$, with discriminant Δ and j-invariant j. Set $2^*=2$ if $b_2\neq 0$, or $2^*=1$ if $b_2=0$. Define*

$$\mu(E)=\frac{1}{6}\left(\log|\Delta|+\log^+(j)\right)+\log^+(b_2/12)+\log(2^*).$$

Then for all $P\in E(\mathbb{Q})$,

$$-\frac{1}{12}h(j)-\mu(E)-1.922\leq\hat{h}(P)-h(P)\leq\mu(E)+2.14.$$

This result is easiest to apply in the rank 1 case, as follows. Suppose we have a rational point P of infinite order on E, of height $\hat{h}(P)$. If P is not a generator it is a multiple $P=kQ$ (modulo torsion) of some generator Q, where $k\geq 2$, so that $\hat{h}(Q)\leq\frac{1}{4}\hat{h}(P)$. By the preceding proposition we can bound the naive height of Q and adjust the bound B in our search accordingly. If a further search up to this bound finds no more points, then P was a generator after all; otherwise we are sure to find a generator.

Similar techniques are possible in higher rank situations, using estimates from the geometry of numbers. See the papers [70] and [57] for more details.

We may also remark that since P has finite order if and only if $\hat{h}(P)=0$, the proposition implies that all torsion points have naive height $h(P)\leq\frac{1}{12}h(j)+\mu(E)+1.922$, giving us another way of finding all the rational torsion points.

For the general case, the following simple result[4] may be used.

LEMMA 3.5.2. *Let $B>0$ be such that*

$$S=\{P\in E(\mathbb{Q})\mid\hat{h}(P)\leq B\}$$

contains a complete set of coset representatives for $2E(\mathbb{Q})$ in $E(\mathbb{Q})$. Then S generates $E(\mathbb{Q})$.

PROOF. Let A be the subgroup of $E(\mathbb{Q})$ modulo torsion generated by the points in S. Suppose that A is a proper subgroup; then we may choose $Q\in E(\mathbb{Q})-A$ with $\hat{h}(Q)$ minimal, since $\hat{h}$ takes a discrete set of values. By hypothesis, there exist $P\in A$ and R such that $Q=P+2R$; certainly $R\notin A$, so that $\hat{h}(R)\geq\hat{h}(Q)$ by minimality. Now using the fact that $\hat{h}$ is quadratic and non-negative we obtain a contradiction:

$$\begin{aligned}\hat{h}(P)&=\frac{1}{2}(\hat{h}(Q+P)+\hat{h}(Q-P))-\hat{h}(Q)\\&\geq\frac{1}{2}\hat{h}(2R)-\hat{h}(Q)\\&=2\hat{h}(R)-\hat{h}(Q)\geq\hat{h}(Q)>B.\quad\square\end{aligned}$$

We have two ways of using this in practice. First of all, it is possible to obtain from the two-descent procedure which we use to determine the rank (see the next section), a set of coset representatives for $E(\mathbb{Q})$ modulo $2E(\mathbb{Q})$. Computing the heights of these points we can find

[3] When referring to [60], recall that our $\hat{h}$ is double Silverman's; also, the constant 1.922 appearing here is a (normalized) correction, due to Bremner, of the constant in Silverman's paper.

[4] Attributed in [60] to Zagier, it is also exercise 5 on page 84 of Cassels' book [8].

a B for which the Lemma holds, to which we add the maximum difference between naive and canonical heights from the preceding proposition to get a bound on the naive heights of a set of generators.

Alternatively, assuming that we know the rank r, we first run our search until we find r independent points P_i. Now it is easy to check whether a point P is twice another: if any subset of the P_i sums to $2Q$ for some Q we replace one of the P_i in the sum by Q and gain index 2. After a finite number of steps (since we are in a finitely-generated group) we obtain independent points which are independent modulo 2, and proceed as before.

Again, we have only presented here the most straightforward strategies for enlarging a set of r independent points in $E(\mathbb{Q})$ to a full $\mathbb{Z}$-basis; this is a topic of active research, with new ideas being developed rapidly: see the paper [57] for some recent advances.

Putting the pieces together, we can determine a set of generators for $E(\mathbb{Q})$ modulo torsion, and then compute the regulator, provided that we know its rank. If we do not know the rank, we at least can obtain lower bounds for the rank. Together with the torsion points found in section 3.3, we will have determined the Mordell–Weil group $E(\mathbb{Q})$ explicitly. Computing the rank is the subject of the next section.

3.6 The Mordell–Weil group III: the rank

For an elliptic curve E defined over the rationals, the rank of the Mordell–Weil group $E(\mathbb{Q})$ is by far the hardest of the elementary quantities associated with E to compute, both theoretically and in terms of implementation. Strictly speaking, the two-descent algorithms we will describe are not algorithms at all, as they are not guaranteed to terminate in all cases. One part of the procedure involves establishing whether or not certain curves of genus one have rational points, when they are known to have points everywhere locally (that is, over $\mathbb{R}$ and over the p-adic field $\mathbb{Q}_p$ for all primes p): there is no known algorithm to decide this in general. Moreover, even without this difficulty, for curves with large coefficients and no rational points of order two, the general two-descent algorithm takes too long to run in practice. For simplicity, we will refer to the procedures as rank algorithms, although their output in certain cases will be bounds on the rank rather than its actual value.

We originally decided to implement a general two-descent procedure in order to check that the modular curves we had computed did have their rank equal to the analytic rank, which we knew, as described in the previous chapter. This was a somewhat thankless task, as it involved a large programming effort, and a large amount of computer time to run the resulting program, in order to verify that approximately 2500 numbers did in fact have the values 0, 1 or 2 which we were already sure were correct. Since the project started, the major theoretical advances by Kolyvagin, Rubin and others meant that all the cases of rank 0 or 1 were known anyway, which left just 18 cases of conjectured rank 2 to verify. In the end we were able to verify these cases, and to check all but a few dozen of the rank 0 or 1 curves; we also obtained extra information by the two-descent procedure, such as the 2-rank of the Tate–Shafarevich group Ш, and a set of coset representatives for $E(\mathbb{Q})/2E(\mathbb{Q})$.

Since the original implementation, the algorithm has been much improved in many ways (notably the syzygy sieve in the search for quartics, the systematic use of group structure in the 2-isogeny case, and the use of quadratic sieving in searching for rational points on homogeneous spaces: see below for details). Our program `mwrank`,[5] based on the algorithm, now works well on a much larger set of curves, including some of fairly high rank such as a curve of Fermigier [23] with rank 13 and 2-torsion (see the example below), and several curves with no 2-torsion and ranks 6, 7 and 8. However, curves with extremely large coefficients, such as Nagao's curve of rank (at least) 21 (see [45]), are beyond the reach of this algorithm owing to the enormous

[5] Available from the author's ftp site: see the Introduction for details.

search regions required. One can also use the program `mwrank` to find points on curves which are too large to find by the search methods of the previous section.

We will not describe here the theory of two-descent, which is the basis of the algorithm, in great detail. Roughly speaking, one has an injective homomorphism from $E(\mathbb{Q})/2E(\mathbb{Q})$ into a finite elementary abelian 2-group, the 2-Selmer group, and attempts to determine the image; if this has order 2^t then the rank of $E(\mathbb{Q})$ is t, $t-1$ or $t-2$ according to whether the number of points of order 2 in $E(\mathbb{Q})$ is 0, 1 or 3 (respectively). This procedure applies to arbitrary curves, and is called *general two-descent*. When E has a rational point P of order two, there is a rational 2-isogeny $\phi : E \to E' = E/\langle P\rangle$ and a dual isogeny $\phi' : E' \to E$. We may then proceed differently, using a procedure we call *two-descent via 2-isogeny*: we embed each of $E'/\phi(E)$ and $E/\phi'(E')$ into finite subgroups of $\mathbb{Q}^*/(\mathbb{Q}^*)^2$, which are easy to write down. This is in contrast to the general two-descent, where one has to work hard to find the Selmer group itself. A full description of two-descent can be found in the standard references such as the books by Silverman [58], Husemöller [27], Knapp [28], or Cassels [8], but the descriptions given there are only easy to apply when E has all its 2-torsion rational. For the general case where there are no rational points of order 2, the main reference is one of the original papers [3] by Birch and Swinnerton–Dyer on their Conjecture, and we followed that paper closely in writing the first version of our program. More detail on the invariant theory, which has resulted in substantial improvements to the general two-descent algorithm, can be found in the paper [20]; a very full description of the algorithm, together with its extension to real quadratic number fields (see also [19]), can be found in Serf's thesis [52].

Both algorithms involve the classification of certain curves, associated with the given elliptic curve E, called *principal homogeneous spaces*. These are twists of E: curves of genus 1 isomorphic to E over an extension field, but not (necessarily) over $\mathbb{Q}$ itself; they need not have rational points, so need not themselves be elliptic curves. When they do have rational points, these map to rational points on E; the maps $H \to E$ are called 2-*coverings* and have degree 4 (in the general two-descent) or 2 (in the 2-isogeny descent). The homogeneous spaces which arise in both algorithms have equations of the form

$$H : \qquad y^2 = g(x) = ax^4 + bx^3 + cx^2 + dx + e \tag{3.6.1}$$

where $g(x)$ is a quartic polynomial with rational coefficients. For brevity we will usually refer to these principal homogeneous spaces simply as *quartics*. The invariants I and J of $g(x)$ (see below for their definition) are related to the invariants c_4 and c_6 of either E or the 2-isogenous curve E'. In the case of descent via 2-isogeny, $g(x)$ will in fact be a quadratic in x^2. We will be interested in whether the quartic H has points over $\mathbb{Q}$ or one of its completions, the p-adics $\mathbb{Q}_p$ or the reals $\mathbb{R}$. Such a point will either be an affine point (x, y) satisfying the equation (3.6.1), or one of the two points at infinity on the projective completion of H, which are rational if and only if a is a square.

In all cases, a quartic with a (global) rational point (x, y) will lead to a rational point on the original curve E, and the set of all the rational points thus obtained will cover the cosets of $2E(\mathbb{Q})$ in $E(\mathbb{Q})$; thus we will be able to determine the rank of $E(\mathbb{Q})$, and at the same time obtain a set of points which generates a subgroup of the Mordell-Weil group $E(\mathbb{Q})$ of odd, finite index. Quartics with no global rational point which are everywhere locally soluble arise from non-trivial elements in the Tate–Shafarevich group of E (or of E'); if these exist, we will only obtain upper and lower bounds for the rank. This is because we currently have no general procedure for proving that a quartic with no rational points does have none. In practice, moreover, it is often impossible to distinguish between such a quartic and one with rational points which are all very large, and hence outside the search region; this happens when a curve has some very large generators, and in such cases also we may only be able to give bounds for

the rank. Further work on these questions is clearly needed, and is currently the focus of much active research.

Since the covering maps $H \to E$ have degree 2 or 4, the rational points on H tend to be smaller (in the sense of naive height) than the rational points they map to on E; this makes them easier to find by search. Here is an example of this: the curve $y^2 = x^3 - 673$ has rank 2, with generators $P_1 = (29, 154)$ and $P_2 = (33989323537/61761^2, -1384230292401340/61761^3)$. The second generator, which would take a very long time to find by searching on the curve itself, is obtained from the rational point $(x, y) = (191/97, 123522/97^2)$ on the quartic with coefficients $(a, b, c, d, e) = (-2, 4, -24, 164, -58)$. This is much easier to find: our program takes less than a second to find the rank and both generators of this curve (but in this time it does not prove that they are generators, only that they generate a subgroup of finite odd index in the Mordell-Weil group).

Before we describe the two main two-descent algorithms, we will present algorithms for determining local solubility and for attempting to determine global solubility of a quartic equation such as (3.6.1), as these are used in both the algorithms.

Checking local solubility.

Here we present an algorithm for determining the local solubility of a curve of the form (3.6.1), where $g(x)$ is a square-free quartic polynomial with integer coefficients. It is easy to generalize this algorithm in two ways: firstly, one might be interested in polynomials of higher degree (when studying curves of higher genus, for example); secondly, working over a general number field K, one would replace the p-adic field $\mathbb{Q}_p$ here with the appropriate completion of K. These extensions are quite straightforward.

Solubility at the infinite prime (that is, over the reals) is easily determined. If $g(x)$ has a real root then it certainly takes positive values, so that H has real points; if $g(x)$ has no real roots, then the values of $g(x)$ have constant sign, and we merely have to check that $a > 0$.

Regarding the finite primes, we first observe that there are only a finite number which need checking in each case, for if p is an odd prime not dividing the discriminant of g, then H certainly has points modulo p which are nonsingular and hence lift to p-adic points. For the other primes, we present an algorithm first given in [3].

It suffices to determine solubility in $\mathbb{Z}_p$ for either $g(x)$ or $g^*(x) = ex^4 + dx^3 + cx^2 + bx + a$, and in the latter case we may assume $x \in p\mathbb{Z}_p$. Given x_k modulo p^k, one tries to lift to a p-adic point (x, y) with $x \equiv x_k \pmod{p^k}$. In [3], conditions are given for this to be possible; more precisely, one of three possibilities may occur (given k and x_k): either a lifting is definitely possible, and we may terminate the algorithm with a positive result; or it is definitely not possible, and we reject this value of x_k; or it is impossible to decide without considering x_k modulo a higher power of p. The test for this lifting is given below in the two subroutines called `lemma6` and `lemma7`, named after the corresponding results in [3]. This leads to a recursive algorithm which is guaranteed to terminate since in any given case there is an exponent k such that it is possible[6] to determine p-adic solubility by considering solubility modulo p^k. All this is an exercise in Hensel's Lemma; the prime $p = 2$ needs to be considered separately. For the details, we refer to the pseudocode below, or to [3]. Further information on local solubility may be found in [56] and [57].

Here is the pseudocode for these algorithms. Note that for any given elliptic curve, all the homogeneous spaces considered will have the same discriminant as the curve (up to a power of 2), so that in practice we would not need to factorize the discriminant of each quartic.

[6]In fact, if $k > \mathrm{ord}_p(\mathrm{disc}(g))$, and also $k \geq 2$ when $p = 2$, then the third possibility cannot occur in algorithms `lemma6` and `lemma7`.

Algorithm for determining local solubility of a quartic

```
INPUT:        a, b, c, d, e (integer coefficients of a quartic g(x))
OUTPUT:       TRUE/FALSE    (solubility of y^2=g(x) in ℝ and in ℚ_p for all p)
 1.  BEGIN
 2.  IF NOT R_soluble(a,b,c,d,e) THEN RETURN FALSE FI;
 3.  IF NOT Qp_soluble(a,b,c,d,e,2) THEN RETURN FALSE FI;
 4.  Δ = discriminant(a,b,c,d,e);
 5.  p_list = odd_prime_factors(Δ);
 6.  FOR p IN p_list DO
 7.  BEGIN
 8.       IF NOT Qp_soluble(a,b,c,d,e,p) THEN RETURN FALSE FI
 9.  END;
10.  RETURN TRUE
11.  END
```

(*Subroutine for determining real solubility)*

```
SUBROUTINE R_soluble(a,b,c,d,e)
INPUT:        a, b, c, d, e (integer coefficients of a quartic g(x))
OUTPUT:       TRUE/FALSE    (solubility of y^2=g(x) in ℝ)
 1.  BEGIN
 2.  IF a>0 THEN RETURN TRUE FI;
 3.  x_list = real_roots(a*x^4+b*x^3+c*x^2+d*x+e=0);
 4.  IF length(x_list)>0 THEN RETURN TRUE FI;
 5.  RETURN FALSE
 6.  END
```

(*Subroutine for determining p-adic solubility)*

```
SUBROUTINE Qp_soluble(a,b,c,d,e,p)
INPUT:        a, b, c, d, e (integer coefficients of a quartic g(x))
              p             (a prime)
OUTPUT:       TRUE/FALSE    (solubility of y^2=g(x) in ℚ_p)
 1.  BEGIN
 2.  IF Zp_soluble(a,b,c,d,e,0,p,0) THEN RETURN TRUE FI;
 3.  IF Zp_soluble(e,d,c,b,a,0,p,1) THEN RETURN TRUE FI;
 4.  RETURN FALSE
 5.  END
```

(*Recursive $\mathbb{Z}_p$-solubility subroutine)*

```
SUBROUTINE Zp_soluble(a,b,c,d,e,x_k,p,k)
INPUT:        a, b, c, d, e (integer coefficients of a quartic g(x))
              p             (a prime)
              x_k           (an integer)
              k             (a non-negative integer)
OUTPUT:       TRUE/FALSE    (solubility of y^2=g(x) in ℤ_p, with x≡x_k (mod p^k))
 1.  BEGIN
 2.  IF p=2
 3.  THEN code = lemma7(a,b,c,d,e,x_k,k)
 4.  ELSE code = lemma6(a,b,c,d,e,x_k,p,k)
```

```
 5.  FI;
 6.  IF code=+1 THEN RETURN TRUE FI;
 7.  IF code=-1 THEN RETURN FALSE FI;
 8.  FOR t = 0 TO p-1 DO
 9.  BEGIN
10.       IF Zp_soluble(a,b,c,d,e,x_k+t*p^k,p,k+1) THEN RETURN TRUE FI
11.  END;
12.  RETURN FALSE
13.  END
```

($\mathbb{Z}_p$ lifting subroutine: odd p)

```
SUBROUTINE lemma6(a,b,c,d,e,x,p,n)
 1.  BEGIN
 2.  gx = a*x^4+b*x^3+c*x^2+d*x+e;
 3.  IF p_adic_square(gx,p) THEN RETURN +1 FI;
 4.  gdx = 4*a*x^3+3*b*x^2+2*c*x+d;
 5.  l = ord(p,gx); m = ord(p,gdx);
 6.  IF (l≥m+n) AND (n>m) THEN RETURN +1 FI;
 7.  IF (l≥2*n) AND (m≥n) THEN RETURN 0 FI;
 8.  RETURN -1
 9.  END
```

($\mathbb{Z}_2$ lifting subroutine)

```
SUBROUTINE lemma7(a,b,c,d,e,x,n)
 1.  BEGIN
 2.  gx = a*x^4+b*x^3+c*x^2+d*x+e;
 3.  IF p_adic_square(gx,2) THEN RETURN +1 FI;
 4.  gdx = 4*a*x^3+3*b*x^2+2*c*x+d;
 5.  l = ord(p,gx); m = ord(p,gdx);
 6.  gxodd = gx; WHILE even(gxodd) DO gxodd = gxodd/2;
 7.  gxodd = gxodd (mod 4);
 8.  IF (l≥m+n) AND (n>m) THEN RETURN +1 FI;
 9.  IF (n>m) AND (l=m+n-1) AND even(l) THEN RETURN +1 FI;
10.  IF (n>m) AND (l=m+n-2) AND (gxodd=1) AND even(l) THEN RETURN +1 FI;
11.  IF (m≥n) AND (l≥2*n) THEN RETURN 0 FI;
12.  IF (m≥n) AND (l=2*n-2) AND (gxodd=1) THEN RETURN 0 FI;
13.  RETURN -1
14.  END
```

A few further remarks on these algorithms: firstly, only trivial changes need to be made to the algorithms `Qp_soluble` and `Zp_soluble` to make them apply to more general equations of the form $y^2 = g(x)$ where $g(x)$ is a non-constant squarefree integer polynomial. This is relevant for work on curves of higher genus, and was observed by S. Siksek. Secondly, extensions to more general p-adic fields are also useful in studying curves over number fields, and again the extensions of Lemma 6 and Lemma 7 in [3] are not difficult. See the theses [56] and [52] for details of such extensions.

Lastly, D. Simon observed that in our application of the algorithms `lemma6` and `lemma7`, we only care whether there is a solution, not necessarily that there is a solution congruent to the given $x \pmod{p^k}$; hence line 6 of subroutine `lemma6` and line 8 of subroutine `lemma7` can both be replaced by:

```
IF l>2*m THEN RETURN +1 FI.
```

Checking global solubility.

To determine whether an equation (3.6.1) has a rational point is much harder than the corresponding local question. All we can do at present is search (efficiently) for a point up to a certain height, after checking that there is no local obstruction. The only satisfactory way known at present to decide on the existence of rational points on these homogeneous spaces is to carry out so-called higher descents; as mentioned above, this is the subject of current work (see [63], for example), and we will not consider it further here.

Our strategy is to look first for a small rational point, using a very simple procedure with low overheads; if this fails, we check for local solubility; if this passes, we start a much more thorough search for a global point, using a quadratic sieving procedure rather similar to the one described in the previous section for finding points on the elliptic curve itself. (In fact, such a sieve-assisted search may be used to find rational points on any curve given by an equation of the form $y^2 = g(x)$ where $g(x)$ is a polynomial in x.) The philosophy here is that there is no point in looking hard for rational points unless one is sure of local solubility, but also that there is no point in checking local solubility when there is an obvious global point.

To carry out the sieve-assisted search, for each possible denominator of x one precomputes, for each of several sieving moduli m, the residues to which the numerator of x must belong if the right-hand side of the equation is to be a square modulo m. In addition, it is easy to see that for every odd prime p dividing the denominator of the x-coordinate of a rational point, we must have $(\frac{a}{p}) = +1$; so provided that the leading coefficient a is not a square (in which case the points at infinity are rational anyway), we precompute a list of primes p for which $(\frac{a}{p}) = -1$, and discard possible denominators divisible by any of these primes. For $p = 2$ a similar condition holds.[7] One also obviously restricts the search to ranges of x for which $g(x)$ is positive; depending on the number of real roots of g and the sign of a, this splits the search into up to three intervals. Finally, in the case of two-descent via 2-isogeny, where the quartics are polynomials in x^2 and thus even, we may restrict to positive x.

For reasons of space, we will only give here the code for a simple point search with no sieving.

Algorithm for searching for a rational point on a quartic: simple version

```
INPUT:        a, b, c, d, e (integer coefficients of a quartic g(x))
              k1, k2        (lower and upper bounds)
OUTPUT:       TRUE/FALSE    (solubility of y^2=g(x) in ℚ with x=u/w
                            and k1 ≤ |u|+w ≤ k2)

 1.  BEGIN
 2.  FOR n = k1 TO k2 DO
 3.  BEGIN
 4.       IF n=1 THEN
 5.            IF square(a) RETURN TRUE FI;
 6.            IF square(e) RETURN TRUE FI
 7.       ELSE
 8.            FOR u = 1 TO n-1 DO
 9.            BEGIN
10.                 IF gcd(u,n)=1
11.                 THEN
12.                      w = n-u;
13.                      IF square(a*u^4+b*u^3w+c*u^2w^2+d*uw^3+e*w^4) RETURN TRUE FI;
```

[7] I am grateful to J. Gebel for this idea, which saves considerable time in practice.

```
14.                     IF square(a*u^4-b*u^3w+c*u^2w^2-d*uw^3+e*w^4) RETURN TRUE FI
15.                 FI
16.             END
17.         FI
18.   END;
19.   RETURN FALSE
20.   END
```

We will now describe the two main two-descent algorithms: two-descent via 2-isogeny for use when E has a rational point of order 2, and general two-descent in the general case. We only use general two-descent when there is no point of order 2, so that the first method does not apply. The situation is not appreciably simpler when E has all three of its points of order two rational than when there is just one rational point of order two, and so we will not bother to consider this case separately.

Method 1: descent using 2-isogeny.

Suppose that E has a rational point P of order 2. By a change of coordinates we may assume that E has equation

$$E: \quad y^2 = x(x^2 + cx + d)$$

where $P = (0, 0)$, and $c, d \in \mathbb{Z}$. Explicitly, in terms of a Weierstrass equation, let x_0 be a root of the cubic $x^3 + b_2x^2 + 8b_4x + 16b_6$, and set $c = 3x_0 + b_2$, $d = (c + b_2)x_0 + 8b_4$. If $a_1 = a_3 = 0$, then we can avoid a scaling factor of 2 by letting x_0 be a root of $x^3 + a_2x^2 + a_4x + a_6$, and setting $c = 3x_0 + a_2$, $d = (c + a_2)x_0 + a_4$. The 2-isogenous curve $E' = E/\langle P\rangle$ has equation

$$E': \quad y^2 = x(x^2 + c'x + d')$$

where

$$c' = -2c \qquad \text{and} \qquad d' = c^2 - 4d.$$

The nonsingularity condition on E is equivalent to $dd' \neq 0$. The 2-isogeny $\phi\colon E \to E'$ has kernel $\{0, P\}$ and in general maps (x, y) to $\left(\frac{y^2}{x^2}, \frac{y(x^2-d)}{x^2}\right)$. The dual isogeny $\phi'\colon E' \to E$ maps (x, y) to $\left(\frac{y^2}{4x^2}, \frac{y(x^2-d')}{8x^2}\right)$.

For each factorization $d = d_1d_2$, with d_1 square-free, we consider the homogeneous space

$$H(d_1, c, d_2): \quad v^2 = d_1u^4 + cu^2 + d_2.$$

Let $n_1 = n_1(c, d)$ be the number of factorizations of d for which the quartic $H(d_1, c, d_2)$ has a rational point, and $n_2 = n_2(c, d)$ the number for which the quartic has a point everywhere locally. Define $n_1' = n_1(c', d')$ and $n_2' = n_2(c', d')$ similarly. Then it is not hard to show by rather explicit calculation (see below and the references given) that $E(\mathbb{Q})/\phi'(E'(\mathbb{Q}))$ is isomorphic to the subgroup of $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ generated by the factors d_1 for which $H(d_1, c, d_2)$ has a rational point. Thus

$$|E(\mathbb{Q})/\phi'(E'(\mathbb{Q}))| = n_1,$$

which must therefore be a power of 2, say $n_1 = 2^{e_1}$; similarly,

$$|E'(\mathbb{Q})/\phi(E(\mathbb{Q}))| = n_1' = 2^{e_1'}.$$

It then follows (see below) that

$$\operatorname{rank}(E(\mathbb{Q})) = \operatorname{rank}(E'(\mathbb{Q})) = e_1 + e_1' - 2. \tag{3.6.2}$$

With luck one will find rational points on all the quartics which have them everywhere locally; then $n_1 = n_2$, and there is no ambiguity in the result. However there will be cases in which the number $\tilde{n}_1$ of quartics on which we can find a rational point is strictly less than n_2. In such cases, we will only have upper and lower bounds for n_1, and similarly for n_1', leading to upper and lower bounds for the rank. This can happen for two reasons: either there is a rational point on some quartic, but our search bound was too small to find it; or the quartic has points everywhere locally but no global rational point.

The quartics H which have points everywhere locally but not globally come from elements of order 2 in the Tate–Shafarevich groups $Ш(E/\mathbb{Q})$ and $Ш(E'/\mathbb{Q})$. There is an exact sequence

$$0 \to E(\mathbb{Q})/\phi'(E'(\mathbb{Q})) \to S^{(\phi')}(E'/\mathbb{Q}) \to Ш(E'/\mathbb{Q})[\phi'] \to 0$$

coming from Galois cohomology; here $S^{(\phi')}(E'/\mathbb{Q})$ is the Selmer group of order n_2 whose elements are represented by the homogeneous spaces $H(d_1, c, d_2)$ which are everywhere locally soluble, and $Ш(E'/\mathbb{Q})$ is the Tate–Shafarevich group of E'. The injective map $E(\mathbb{Q})/\phi'(E'(\mathbb{Q})) \to S^{(\phi')}(E'/\mathbb{Q})$ is induced by taking a point (x, y) in $E(\mathbb{Q})$ with $x \neq 0$ to the space $H(d_1, c, d_2)$ where $d_1 = x$ modulo squares: if $x = d_1u^2$ and $v = uy/x$ then (u, v) is a rational point on $H(d_1, c, d_2)$. The point $P = (0, 0)$ maps to d modulo squares. Conversely, if (u, v) is a rational point on $H(d_1, c, d_2)$ then $(x, y) = (d_1u^2, d_1uv)$ is a rational point on E. (In proving these statements, one has to check that two rational points on E have the same x-coordinate modulo squares if and only if their difference is in $\phi'(E'(\mathbb{Q}))$; for example, the image of P is d, which is a square if and only if $P \in \phi'(E'(\mathbb{Q}))$.) It follows that n_1 is the order of $E(\mathbb{Q})/\phi'(E'(\mathbb{Q}))$, as stated above, and hence that

$$|Ш(E'/\mathbb{Q})[\phi']| = n_2/n_1.$$

Similarly, from the exact sequence

$$0 \to E'(\mathbb{Q})/\phi(E(\mathbb{Q})) \to S^{(\phi)}(E/\mathbb{Q}) \to Ш(E/\mathbb{Q})[\phi] \to 0$$

with similarly defined maps, we obtain

$$|Ш(E/\mathbb{Q})[\phi]| = n_2'/n_1'.$$

Thus the result is only genuinely ambiguous when either $Ш(E/\mathbb{Q})[\phi]$ or $Ш(E'/\mathbb{Q})[\phi']$ is non-trivial, so that not all elements of the Selmer groups are obtained from rational points on the elliptic curves. This is rare for the curves in the tables, but obviously must be taken into account in general. A typical situation is to have $n_2n_2' = 16$ and $n_1n_1' \geq 4$, when one suspects that $r = 0$ with $|Ш(E/\mathbb{Q})[2]| = 4$ or $|Ш(E'/\mathbb{Q})[2]| = 4$, but where it is possible instead that $r = 2$ and $|Ш(E/\mathbb{Q})[2]| = |Ш(E'/\mathbb{Q})[2]| = 1$. Curve 960D1 in the tables is an example of this, although in this case since the curve is modular and we know that $L(E, 1) \neq 0$, it must have rank 0 by the result of Kolyvagin mentioned earlier. We can also deduce this by working with the 2-isogenous curves 960D3 and 960D2, where there is no ambiguity: here $n_1 = n_2 = n_1' = n_2' = 2$, showing that the rank is certainly 0. (Note that isogenous curves have the same rank, but not necessarily the same order of Ш, which can work to our advantage in cases like this.) Returning to the pair 960D1-960D2 where we compute $n_1 = n_2 = 1$, $n_2' = 16$ and $n_1' \geq 4$, now we know that the rank is in fact zero we can conclude that $n_1' = 4$, and that $|Ш(E/\mathbb{Q})[\phi]| = 4$. The nontriviality of $Ш(E/\mathbb{Q})$ in this case is confirmed by the Birch–Swinnerton-Dyer conjecture, which for this curve predicts that Ш has order 4 (see Table 4).

Local solubility of $H(d_1, c, d_2)$ is automatic for all primes p which do not divide $2dd'$; for those p which do divide $2dd'$ we may apply the general criteria of Birch and Swinnerton-Dyer.

Local solubility in $\mathbb{R}$ is easy to determine here: if $d' < 0$ then we require $d_1 > 0$, while if $d' > 0$ then either $d_1 > 0$ or $c + \sqrt{d'} > 0$ is necessary. Thus if either $d' < 0$, or $d' > 0$ and $c + \sqrt{d'} < 0$, then we only consider positive divisors d_1 of d, and need not apply the general test for solubility in $\mathbb{R}$.

Each rational point (u, v) on $H(d_1, c, d_2)$ maps, as observed above, to the point (d_1u^2, d_1uv) on E; modulo $\phi'(E'(\mathbb{Q}))$, this is independent of the rational point (u, v), and only depends on d_1 modulo squares. Similarly, a rational point (u, v) on $H(d_1', c', d_2')$ maps to a point on E', and hence via the dual isogeny ϕ' to the point

$$\left(\frac{v^2}{4u^2}, \frac{v(d_1'u^4 - d_2')}{8u^3}\right)$$

in $E(\mathbb{Q})$. The set of n_1n_1' points in $E(\mathbb{Q})$ thus determined (by adding the points constructed in this way) cover the cosets of $E(\mathbb{Q})/2E(\mathbb{Q})$, either once each, when $|E(\mathbb{Q})[2]| = 4$, which is when d' is a square, or twice, when d' is not a square and $|E(\mathbb{Q})[2]| = 2$. Thus, when $|E(\mathbb{Q})[2]| = 2$ we have

$$\frac{n_1n_1'}{2} = |E(\mathbb{Q})/2E(\mathbb{Q})| = 2^{r+1},$$

while if $|E(\mathbb{Q})[2]| = 4$ we have

$$n_1n_1' = |E(\mathbb{Q})/2E(\mathbb{Q})| = 2^{r+2};$$

hence $2^r = n_1n_1'/4$ in both cases, proving (3.6.2).

When counting n_1 and n_2 (and similarly, n_1' and n_2'), it is very useful to use the fact that each is a power of 2, being the order of an elementary abelian 2-group. This is particularly important when d (or d') has many distinct prime factors. Let A_0 be the group of all divisors of d modulo squares, of order n_0 (say). Then A_0 is generated by -1 and the primes dividing d, so that $n_0 = 2^{e_0}$ where e_0 is the number of distinct prime factors of d, plus 1. Within A_0 we must determine the subgroups A_1 and A_2 of orders n_1 and n_2, consisting of those divisors d_1 of d for which the corresponding homogeneous space is everywhere locally or globally soluble, respectively.

We can effectively reduce the size of the set A_0 of divisors to be searched by a factor up to 8 as follows: as observed above, if either $d' < 0$, or $d' > 0$ and $c + \sqrt{d'} < 0$, then we need only consider positive divisors d_1 of d, cutting in two the number of elements of A_0 which may lie in A_1. Secondly, we may take advantage of the fact that we know the rational point $(0, 0)$ on E; thus we know that d is in A_2 (though possibly just the identity if d is a square); similarly, if d' is a square then $x^2 + cx + d$ factorizes, say as $(x - x_2)(x - x_3)$, and we know that x_2 and x_3 also lie in A_2.

More generally, whenever we find in the course of our systematic search through the elements of A_0 that the element d_1 lies in A_2, we can effectively factor out d_1 and reduce the number of remaining values to check by a factor of 2. Of course, this requires careful book-keeping in the implementation; for simplicity, we omit these refinements from the pseudocode below, where we simply loop over all square-free divisors of d and d'.

As an example of the saving that can be made, consider the curve of rank 13 constructed by Fermigier in [23]; this is of the form $y^2 = x(x^2 + cx + d)$ with

$$c = 36861504658225 \qquad \text{and}$$

$$d = 1807580157674409809510400 = 2^{15} \cdot 3^4 \cdot 5^2 \cdot 7^2 \cdot 17 \cdot 23 \cdot 29 \cdot 41 \cdot 103 \cdot 113 \cdot 127 \cdot 809,$$

so that d has 12 distinct prime factors and $2^{13} = 8192$ square-free divisors. Since d is non-square we can cut the set in half, say by excluding all d_1 divisible by the largest prime factor 809, leaving 4096 values to test. In our implementation, the results of the test are as follows:

- 7 non-trivial values of d_1 give rational points after searching, as well as $d_1 = 1$ which gives the trivial point;
- 120 further values are in the subgroup A_2 generated by these 7 values and need not be tested;
- 122 further values were tested and found to be not everywhere locally soluble, hence not in A_1;
- 3846 further values were discarded as being a product of an element of A_2 and an element not in A_1, and hence not in A_1.

Thus in this case we find that $n_1 = n_2 = 256$, after only having to search for points on seven homogeneous spaces. Working with the isogenous curve, we obtain $n_1' = n_2' = 128$ after only searching six homogeneous spaces for points. Thus $e_1 = 8$, $e_1' = 7$ and the rank is 13. Note that in the course of computing this value, we have searched precisely 13 homogeneous spaces, and the points we thereby construct give 13 generators of $E(\mathbb{Q})/2E(\mathbb{Q})$ modulo torsion. Adding $P = (0, 0)$ to this list gives 14 points which generate $E(\mathbb{Q})/2E(\mathbb{Q})$ (which has order 2^{14}), and which therefore generate a subgroup of finite odd index in the full Mordell-Weil group $E(\mathbb{Q})$.

The situation is not always this simple, however, even for curves where Ш[2] is trivial, since there may be homogeneous spaces with rational points which are hard to find. For example, consider Fermigier's curve of rank 14 from [23], with $c = 2429469980725060$ and $d = 27513070338817213683364775638 8$ (which has 14 prime factors). When we run our program using a (logarithmic) bound of 10 in the search for rational points on the quartics, we find $n_1 \geq 64$, $n_1' \geq 128$, while $n_2 = n_2' = 256$. Here the correct values are $n_1 = n_1' = 256$, giving $r = 14$, but we only find $11 \leq r \leq 14$; and in the process, we have had to search many more homogeneous spaces for rational points.

Here is the pseudo-code which implements the algorithm just described. The main routine aborts if either the input curve is singular (this is useful if one wants to apply the algorithm systematically to a range of inputs) or if there is no point of order two. The latter is detected in lines 6–7, where an integer root to a monic cubic with integer coefficients is found (if it exists). Most of the work is done in the subroutine `count(c,d,p_list)` which determines $n_2(c, d)$ and, as far as possible, $n_1(c, d)$. Here `p_list` is the set of 'bad' primes dividing $2dd'$ where local solubility needs to be checked, which we only compute once. There are two calls to the subroutine `rational_point(a,b,c,d,e,k1,k2)`, which seeks a rational u/w with $k_1 \leq |u| + w \leq k_2$ such that $g(u/w)$ is a rational square, where $g(x) = ax^4 + bx^3 + cx^2 + dx + e$. (Here $w > 0$ and $\gcd(u, w) = 1$.) In the first call we carry out a quick check for 'small' points; then we look further, having first checked for everywhere local solubility. The particular parameters `lim1`, `lim2` for the search will probably be decided at run time. The subroutines `Qp_soluble` and `rational_point` are implementations of the algorithms given earlier (though in practice we would use a more efficient algorithm for the second call to `rational_point`, as explained above).

Algorithm for computing rank: rational 2-torsion case

```
INPUT:        a1, a2, a3, a4, a6 (coefficients of E)
OUTPUT:       r_min, r_max       (bounds for rank of E)
              S,S'               (upper bounds for #Ш(E)[φ] and #Ш(E')[φ'])
 1.  BEGIN
 2.  IF a1=a3=0
 3.  THEN s2 = a2; s4 = a4; s6 = a6
 4.  ELSE s2 = a1*a1+4*a2; s4 = 8*(a1*a3+2*a4); s6 = 16*(a3*a3+4*a6)
 5.  FI;
 6.  x_list = integer_roots(x^3+s2*x^2+s4*x+s6=0);
```

```
 7.  IF length(x_list)=0 THEN abort ELSE x0 = x_list[1] FI;
 8.  c = 3*x0+s2; d = (c+s2)*x0 + s4;
 9.  c' = -2*c; d' = c^2-4*d;
10.  IF d*d'=0 THEN abort FI;
11.  p_list = prime_divisors(2*d*d');
12.  (n1,n2) = count(c,d,p_list);
13.  (n1',n2') = count(c',d',p_list);
14.  e1 = log_2(n1); e2 = log_2(n2);
15.  e1' = log_2(n1'); e2' = log_2(n2');
16.  r_min = e1+e1'-2; r_max = e2+e2'-2;
17.  S = n2'/n1'; S' = n2/n1;
18.  RETURN r_min, r_max, S, S'
19.  END
```

(*Main counting subroutine)*

```
SUBROUTINE count(c,d,p_list)
 1.  BEGIN
 2.  n1 = n2 = 1; d' = c^2-4*d;
 3.  d1_list = squarefree_divisors(d);
 4.  FOR d1 IN d1_list DO
 5.  BEGIN
 6.       IF rational_point(d1,0,c,0,d/d1,1,lim1)
 7.       THEN n1 = n1+1; n2 = n2+1
 8.       ELSE
 9.            IF everywhere_locally_soluble(c,d,d',d1,p_list)
10.            THEN
11.                 n2 = n2+1;
12.                 IF rational_point(d1,0,c,0,d/d1,lim1+1,lim2)
13.                 THEN n1 = n1+1
14.                 FI
15.            FI
16.       FI
17.  END;
18.  RETURN (n1, n2)
19.  END
```

(*Subroutine to check for everywhere local solubility)*

```
 1.  SUBROUTINE everywhere_locally_soluble(c,d,d',d1,p_list)
 2.  BEGIN
 3.  IF d'<0 AND d1<0 THEN RETURN FALSE FI;
 4.  IF d'>0 AND d1<0 AND (c+sqrt(d'))<0 THEN RETURN FALSE FI;
 5.  FOR p IN p_list DO
 6.  BEGIN
 7.       IF NOT Qp_soluble(d1,0,c,0,d/d1,p) THEN RETURN FALSE FI
 8.  END;
 9.  RETURN TRUE
10.  END
```

Method 2: general two-descent.

We now turn to the general two-descent, which applies whether or not E has a rational point of order 2. Again, the basic idea is to associate to E a collection of 2-covering quartic

curves (or homogeneous spaces) H. These have equations of the form

$$H: \qquad y^2 = g(x) = ax^4 + bx^3 + cx^2 + dx + e \tag{3.6.1}$$

with $a, b, c, d, e \in \mathbb{Q}$, such that the *invariants*

$$I = 12ae - 3bd + c^2 \qquad \text{and} \qquad J = 72ace + 9bcd - 27ad^2 - 27eb^2 - 2c^3$$

are related to the c_4 and c_6 invariants of E via

$$I = \lambda^4 c_4 \qquad \text{and} \qquad J = 2\lambda^6 c_6$$

for some $\lambda \in \mathbb{Q}^*$. Two such quartics $g_1(x)$, $g_2(x)$ are *equivalent* if

$$g_2(x) = \mu^2(\gamma x + \delta)^4 g_1\left(\frac{\alpha x + \beta}{\gamma x + \delta}\right)$$

for some α, β, γ, δ and $\mu \in \mathbb{Q}$, with μ and $\alpha\delta - \beta\gamma$ non-zero. The invariants of $g_1(x)$ and $g_2(x)$ are then related by the scaling factor $\lambda = \mu(\alpha\delta - \beta\gamma)$:

$$I(g_2) = \mu^4(\alpha\delta - \beta\gamma)^4 I(g_1),$$
$$J(g_2) = \mu^6(\alpha\delta - \beta\gamma)^6 J(g_1).$$

We set $\Delta = 4I^3 - J^2 = 27\mathrm{disc}(g)$, and call Δ the discriminant.

In particular, by scaling up the coefficients, we may assume that the invariants I and J are integral. The number of equivalence classes of quartics with given invariants (up to a scaling factor λ) which are everywhere locally soluble is finite. One of our tasks will be to determine, for a given integral quartic, an equivalent integral one with minimal invariants. This process is closely analogous to the one considered earlier in this chapter, using Kraus's conditions or Tate's algorithm to determine minimal models for elliptic curves. Indeed, we will see below that if c_4 and c_6 are invariants of a minimal model for the elliptic curve E, then $I = c_4$ and $J = 2c_6$ are also minimal, except possibly at the prime 2. (We may lose minimality at 2 because the equations (3.6.1) we use for homogeneous spaces are not completely general, not having terms in y, xy or x^2y; to remove these by completing the square involves a scaling by a factor of 2.)

We now explain the relationship between equivalence classes of soluble quartics with invariants I and J and rational points on the elliptic curve. More details of this relationship, including proofs, may be found in [20]. For convenience, we again start by making a coordinate transformation: if c_4 and c_6 are the integral invariants of our curve E, we set $I = c_4$ and $J = 2c_6$, and replace E by the isomorphic curve

$$E_{I,J}: \quad Y^2 = F(X) = X^3 - 27IX - 27J. \tag{3.6.3}$$

This is the model on which the rational points we construct will naturally lie; it is then a simple matter to transfer them back to the original model for E. For simplicity, we will still continue to refer to the curve simply as E when this will not cause confusion.

Associated to each quartic g there are two so-called *covariants*, which we denote g_4 and g_6:

$$\begin{aligned}
g_4(X,Y) &= (3b^2 - 8ac)X^4 + 4(bc - 6ad)X^3Y + 2(2c^2 - 24ae - 3bd)X^2Y^2 \\
&\quad + 4(cd - 6be)XY^3 + (3d^2 - 8ce)Y^4, \\
g_6(X,Y) &= (b^3 + 8a^2d - 4abc)X^6 + 2(16a^2e + 2abd - 4ac^2 + b^2c)X^5Y \\
&\quad + 5(8abe + b^2d - 4acd)X^4Y^2 + 20(b^2e - ad^2)X^3Y^3 \\
&\quad - 5(8ade + bd^2 - 4bce)X^2Y^4 - 2(16ae^2 + 2bde - 4c^2e + cd^2)XY^5 \\
&\quad - (d^3 + 8be^2 - 4cde)Y^6.
\end{aligned} \tag{3.6.4}$$

Wherever convenient, we will also denote by g the homogenized polynomial $g(X,Y) = aX^4 + bX^3Y + cX^2Y^2 + dXY^3 + eY^4$. These three homogeneous polynomials satisfy an algebraic identity, or syzygy:

$$27g_6^2 = g_4^3 - 48Ig^2g_4 - 64Jg^3. \tag{3.6.5}$$

Later we will also need a simpler form of this syzygy; set

$$p = g_4(1,0) = 3b^2 - 8ac \qquad \text{and} \qquad r = g_6(1,0) = b^3 + 8a^2d - 4abc; \tag{3.6.6}$$

these quantities are called *seminvariants* of g. Substituting $(X,Y) = (1,0)$ in the covariant syzygy (3.6.5) gives an identity (the seminvariant syzygy) between these seminvariants:

$$27r^2 = p^3 - 48Ia^2p - 64Ja^3. \tag{3.6.7}$$

We will make use of this equation in our search for quartics with given invariants, where it will allow us to set up a quadratic sieve.

It follows from the covariant syzygy (3.6.5), by simple substitution, that the map

$$\xi : \quad (x,y) \mapsto \left(\frac{3g_4(x,1)}{(2y)^2}, \frac{27g_6(x,1)}{(2y)^3}\right) \tag{3.6.8}$$

maps rational points (x,y) on H (satisfying $y^2 = g(x,1)$) to rational points on $E_{I,J}$, thus defining a rational map ξ, of degree 4, from $H(\mathbb{Q})$ to $E_{I,J}(\mathbb{Q})$. We are using affine coordinates here; the points at infinity on H map to $\left(\frac{3p}{4a}, \frac{\pm 27r}{(4a)^{3/2}}\right)$, which are rational if and only if a is a square.

We now have the following facts (see [20] for details):

• If $R \in H(\mathbb{Q})$ with $P = \xi(R) \in E_{I,J}(\mathbb{Q})$, then the coset of P modulo $2E_{I,J}(\mathbb{Q})$ is independent of R, and of the particular quartic g up to equivalence; in fact, equivalences between quartics induce rational maps between the associated homogeneous spaces, and the covariant property of g_4 and g_6 ensures that corresponding rational points on the homogeneous spaces have the same image in $E_{I,J}(\mathbb{Q})$.

• Each rational point $P = (x,y) \in E_{I,J}(\mathbb{Q})$ arises as the image of a rational point on some quartic g with invariants I and J: explicitly, one can take the rational point at infinity on the quartic with coefficients $(a,b,c,d,e) = (1, 0, -x/6, y/27, I/12 - x^2/432)$; the equivalence class of g depends only on the coset of P modulo $2E_{I,J}(\mathbb{Q})$.

• The equivalence classes of everywhere locally soluble quartics with invariants I and J form a finite elementary abelian 2-group, isomorphic to the 2-Selmer group $S^{(2)}(E/\mathbb{Q})$.

• The equivalence classes of soluble quartics with invariants I and J form a finite elementary abelian 2-group isomorphic to $E(\mathbb{Q})/2E(\mathbb{Q})$; the identity is the *trivial* class, consisting of quartics with a rational root.

• More generally, when E has no 2-torsion, for any extension field K of $\mathbb{Q}$ there is a bijection between the roots of $g(x)$ in K and the solutions $Q \in E_{I,J}(K)$ to the equation $2Q = P$ (where $P = \xi(R)$ for $R \in H(\mathbb{Q})$ as above). In particular, non-trivial quartics are irreducible in this case. We will use this fact with $K = \mathbb{R}$ later.

We therefore classify the set of equivalence classes of quartics with invariants I and J as follows:

(0) the trivial class consists of those quartics $g(x)$ which have a rational root. These are elliptic curves isomorphic to E over $\mathbb{Q}$.
(1) those which have a rational point: these are also elliptic curves, isomorphic to E over $\mathbb{Q}$.
(2) those which have points everywhere locally.
(3) those which fail to have points everywhere locally.

Let the number of inequivalent quartics in the first three sets be $n_0 = 1$, n_1 and n_2. (Those in the last set will not be used.) Because of the group structure, each of these numbers is a power of 2. We write $n_i = 2^{e_i}$ for $i = 1, 2$.

As in the case of descent via 2-isogeny, Galois cohomology gives an exact sequence

$$0 \to E(\mathbb{Q})/2E(\mathbb{Q}) \to S^{(2)}(E/\mathbb{Q}) \to Ш(E/\mathbb{Q})[2] \to 0.$$

Thus the quotient of $S^{(2)}(E/\mathbb{Q})$ by the image of $E(\mathbb{Q})$ is isomorphic to $Ш(E/\mathbb{Q})[2]$, the 2-torsion subgroup of the Tate–Shafarevich group $Ш(E/\mathbb{Q})$. So it is the points of order 2 in $Ш(E/\mathbb{Q})$, if any, which account for the possible existence of homogeneous spaces which have points everywhere locally but not globally, and we have

$$|Ш(E/\mathbb{Q})[2]| = n_2/n_1.$$

As before, the potential practical difficulty lies in determining whether each homogeneous space H has a rational point, as there is no known algorithm to do this in general. Again, for the vast majority of the curves in the tables, we found a rational point easily on each space which was everywhere locally soluble, which not only determined the rank of E, but also implied that the Tate–Shafarevich group had no 2-torsion. The only example with $n_1 < n_2$ in the tables (for a curve with no 2-torsion) is curve 571A1, where $n_1 = 1$ and $n_2 = 4$; here the rank is 0, and $|Ш(E/\mathbb{Q})[2]| = 4$; the Birch–Swinnerton-Dyer conjecture predicts $|Ш(E/\mathbb{Q})| = 4$.

The steps of the algorithm are as follows: first we determine the pair or pairs of integral invariants (I, J) such that every quartic associated with our curve E is equivalent to one with integer coefficients and these invariants. There will be either one or two such pairs. For each pair (I, J), we find a finite set of quartics with invariants (I, J) such that every non-trivial, everywhere locally soluble quartic with these invariants is equivalent to one in the list. This is the most time-consuming step, as the search region can be very large when I and J are large. Now we must test the quartics in our list pairwise for equivalence, discarding those equivalent to earlier ones; look for rational points; and test everywhere local solubility. Again, there may be quartics where we do not find rational points despite their having points everywhere locally, so that although we can always (given enough time) determine n_2, we may in some cases only find bounds on n_1. Since $n_1 = |E(\mathbb{Q})/2E(\mathbb{Q})|$, we can then compute the rank r, or bounds on the rank. Usually, E will have no rational 2-torsion, or we would probably be using descent via 2-isogeny, and then simply $2^r = n_1$.

We now consider each of these steps in more detail.

Step 1: Determining the invariants (I, J).

Given an integral quartic g with invariants I and J, we must consider the question of whether there exists an equivalent integral quartic with smaller invariants. The smaller invariants will have the form $\lambda^{-4}I$, $\lambda^{-6}J$ with $\lambda \in \mathbb{Q}^*$. In [3, Lemmas 3–5], conditions are stated under which g is equivalent to an integral quartic with invariants $p^{-4}I$, $p^{-6}J$ for a prime p; we call such a quartic *p-reducible*, otherwise *p-minimal*. Clearly a necessary condition for reducibility is that $p^4 \mid I$ and $p^6 \mid J$. We say that the pair (I, J) is *p-reducible* if every integral quartic with these invariants *which is p-adically soluble* is equivalent to an integral quartic with invariants $p^{-4}I$ and $p^{-6}J$.

The question of p-reducibility is almost completely settled by the following proposition. The result is simplest for primes greater than 3, but even for these it is important to note that the assumption of p-adic solubility is necessary for reduction to be possible when the divisibility conditions are satisfied.

PROPOSITION 3.6.1. *Let I and J be integers such that $\Delta = 4I^3 - J^2 \neq 0$.*

(1) *If p is a prime and $p \geq 5$, then (I, J) is p-reducible if and only if $p^4 \mid I$ and $p^6 \mid J$.*

(2) (I, J) *is* 3*-reducible if and only if either* $3^5 \mid I$ *and* $3^9 \mid J$*, or* $3^4 \,||\, I$*,* $3^6 \,||\, J$ *and* $3^{15} \mid \Delta$*.*
(3) (I, J) *is* 2*-reducible if* $2^6 \mid I$*,* $2^9 \mid J$ *and* $2^{10} \mid 8I + J$*.*

This proposition is stated in [3] as Lemmas 3–5, but only the proof of Lemma 3 (covering the case $p \geq 5$) is given there. Complete proofs in all cases (which are elementary though somewhat lengthy) can be found in [52].

Note that for $p = 2$ we only have sufficient conditions for reducibility. Because of this, we will sometimes have to consider two pairs of invariants, a smaller pair (I_0, J_0) and a larger pair $(16I_0, 64J_0)$. However, when searching for integral quartics with the larger invariants, we may assume that the quartic cannot be 2-reduced, and this provides us with useful congruence conditions on the coefficients of such a quartic. We state these here.

PROPOSITION 3.6.2. *Let* g *be an integral* 2*-adically soluble quartic whose invariants satisfy* $2^4 \mid I$ *and* $2^6 \mid J$*, such that*

(1) g *is not equivalent to an integral quartic with invariants* $2^{-4}I$ *and* $2^{-6}J$*;*
(2) g *is not equivalent to an integral quartic with the same invariants* I *and* J *and smaller leading coefficient* a*.*

Then the coefficients of g *satisfy*

(a) $2 \nmid a$*,* $2^2 \mid b$*,* $2 \mid c$*,* $2^4 \nmid e$ *and* $2^4 \nmid a + b + c + d + e$*; or*
(b) $2 \,||\, a$*,* $2^2 \mid b$*,* $2^2 \mid c$*,* $2^3 \nmid e$ *and* $2^3 \nmid a + b + c + d + e$*.*

Moreover, if $2^6 \mid I$ *and* $2^7 \mid J$*, then we must have*

(a') $2 \nmid a$*,* $2^2 \mid b$*,* $2^2 \,||\, c$*,* $2^3 \mid d$*, and* $2^2 \,||\, e$*; or*
(b') $2 \nmid a$*,* $2^2 \mid b$*,* $2^2 \,||\, c - 2a + 3b$*,* $2^3 \mid d - b$ *and* $2^2 \,||\, a + c + e$*.*

The first set of conditions stated here were given in [3]; the second set are from [52], which contains complete proofs in both cases.

Using this proposition, we may ensure that few of the quartics we find when searching the larger pair of invariants are equivalent to one with smaller invariants. More significantly in terms of running time, we have extra congruence conditions to apply when searching for the larger invariants, which speeds up this search.

It would appear that rational points in $E(\mathbb{Q})$ whose quartics have the larger pair of invariants lie in certain components of the 2-adic locus $E(\mathbb{Q}_2)$. Further study of this would be very useful, since if the search for quartics with the larger pair of invariants could be eliminated or curtailed, it could result in a major saving of time in the algorithm.

In practice, suppose that our original curve E is given by a minimal equation, with invariants c_4 and c_6. We set $I = c_4$ and $J = 2c_6$. Clearly the pair (I, J) is p-minimal for $p \geq 5$: for if $p^4 \mid I$ and $p^6 \mid J$ then $p^{-4}c_4$ and $p^{-6}c_6$ would be integral invariants of an elliptic curve, contradicting minimality of E, and similarly the pair $(p^4 I, p^6 J)$ is certainly p-reducible by Proposition 3.6.1(1). Less obvious is that (I, J) is also 3-minimal; using Kraus's conditions, it is easy to check first that $(3^4 I, 3^6 J)$ is certainly 3-reducible (one needs here that $\mathrm{ord}_3(c_6) \neq 2$), and then that (I, J) itself is not 3-reducible, using Proposition 3.6.1(2).

For $p = 2$, the best we can do is the following. First set $I = c_4$ and $J = 2c_6$. Replace (I, J) by $(2^{-4}I, 2^{-6}J)$ if $2^4 \mid I$ and $2^6 \mid J$; the resulting pair (I, J) (which will not be further divisible by 2) will be the basic pair of invariants. Then we also use the pair $(16I, 64J)$ unless $4 \mid I$, $8 \mid J$ and $16 \mid (2I + J)$.

The result of this step is then to produce either one or two pairs of invariants (I, J). In the latter case, the following steps must be carried out with both pairs separately.

Step 2: Finding the quartics with given I and J.

We now have a fixed pair of invariants (I, J) with $\Delta = 4I^3 - J^2 \neq 0$, and we wish to find all integral quartics with these invariants, up to equivalence. We classify the quartics $g(x)$ into

types, according as $g(x)$ has no real roots (type 1), four real roots (type 2) or two real roots (type 3). When $\Delta < 0$ only type 3 is possible, while if $\Delta > 0$, only types 1 and 2 are possible. For each relevant type, we now determine a finite list of quartics of that type with the given invariants such that every soluble quartic with these invariants is equivalent to at least one on the list. We can ignore quartics which are negative definite (type 1 with $a < 0$), since they will not be soluble over $\mathbb{R}$. For each type, we will determine a finite region of (a, b, c)-space such that every quartic with invariants I and J is equivalent to at least one in this region.

As observed above, the number of real roots of $g(x)$ is equal to the number of points $Q \in E(\mathbb{R})$ satisfying $2Q = P$, where $P \in E(\mathbb{R})$ is the image under the map ξ of any real point on the homogeneous space H with equation $y^2 = g(x)$. When $\Delta < 0$, the real locus is in one component, and $E(\mathbb{R})$ is isomorphic to the circle group, which is 2-divisible with two 2-torsion points, so in this case the equation $2Q = P$ has exactly two solutions for all $P \in E(\mathbb{R})$. This agrees with the observation just made, that quartics with negative discriminant Δ will all have exactly two real roots.

Consider further the case $\Delta > 0$. Now $E(\mathbb{R})$ has two components, the connected component of the identity $E^0(\mathbb{R})$ and a second component which we call the 'egg'. There are four 2-torsion points, and $2E(\mathbb{R}) = E^0(\mathbb{R})$. There are therefore two possibilities for a point $P \in E(\mathbb{R})$ and its associated real quartic: if $P \in E^0(\mathbb{R})$, then there are four solutions Q to $2Q = P$, and P will be associated to a quartic of type 2 with four real roots. On the other hand, if $P \notin E^0(\mathbb{R})$, then there are no solutions and the quartic associated to P will be of type 1, with no real roots.

The image of $E(\mathbb{Q})$ in $E(\mathbb{R})/2E(\mathbb{R})$ has order 2 or 1, depending on whether or not there are any rational points on the egg. Thus there are two sub-cases to the case $\Delta > 0$: if $E(\mathbb{Q}) \subset E^0(\mathbb{R})$, then there are no rational points on the egg, the index is 1, and there will be *no* soluble quartics of type 1; on the other hand, if $E(\mathbb{Q}) \not\subset E^0(\mathbb{R})$, then there are rational points on the egg, the index is 2, and there are equal numbers of (equivalence classes of) soluble quartics of types 1 and 2. Those of type 2 will lead to rational points on $E(\mathbb{Q}) \cap E^0(\mathbb{R})$, while those of type 1 will lead to rational points on the egg.

To take advantage of this in practice, when $\Delta > 0$ we will first look for quartics of type 2; let the number of these be n_1^+, where n_1/n_1^+ is either 1 or 2. At this stage we will already know the rank to within one, since if we set $r^+ = \log_2(n_1^+)$ then (assuming no rational 2-torsion) we have either $r = r^+$ or $r = r^+ + 1$. Then we start to look for quartics of type 1; as soon as we find one which is soluble, then we may abort the search for type 1 quartics at that point, and assert that $r = r^+ + 1$. On the other hand, if we complete the search for quartics of type 1 without finding any soluble ones, then we will know that $r = r^+$, and we will have proved that there are no rational points on the egg. An example of the second possibility happens with the curve $E = [0, 0, 1, -529, -3042]$ (which is the -23-twist of the curve $[0, 0, 1, -1, 0]$ with conductor 37 and rank 1), which has rank 1 with generator $(46, 264)$ on the identity component, and no rational points on the egg.[8]

If we happened to know in advance that there were rational points on the egg (perhaps by a short preliminary search for such points with small height), then we would already know that $r = r^+ + 1$, and we would not need to search for type 1 quartics at all.

In order to find all integral quartics of a given type (up to equivalence) we proceed as follows. First, following [3], we determine bounds on the coefficients a, b and c. We also set up a sieve based on the seminvariant syzygy (3.6.7) to speed up our search through this region of (a, b, c)-space. For triples (a, b, c) in the region which pass the sieve, we solve for d and e and ensure that they are integral. Finally, we check that the quartic we have constructed satisfies any further congruence conditions we require (for example, when we are using the larger pair of invariants).

[8]Thanks to Nelson Stephens for this example.

The method for bounding the coefficients which is developed in [3] involves using the auxiliary (resolvent) cubic equation

$$\phi^3 - 3I\phi + J = 0 \tag{3.6.9}$$

which will have one real root (type 3) or three real roots (types 1 and 2), since its discriminant is 27Δ. Indeed, ϕ is a root of (3.6.9) if and only if $(-3\phi, 0)$ is a point of order 2 on the curve $E_{I,J}$.

In each case, the bound for b arises simply from the fact that the quartics $g(x)$ and $g(x+k)$ are equivalent, and the coefficients of the latter are $(a, b+4ak, \dots)$, so that we may assume that b is reduced modulo $4a$. Also, note that the bounds on c are effectively bounds on the seminvariant $8ac - 3b^2 = -p$, which is how they arise in [3].

Bounds for (a,b,c): Type 1. Here we may assume $a > 0$ for real solubility. Order the three real roots of (3.6.9) as $\phi_1 > \phi_2 > \phi_3$, and set $K = (4I - \phi_1^2)/3$. Then the bounds on a, b, c are

$$\begin{aligned} 0 < a &\le \frac{K + K^{\frac{1}{2}}\phi_1}{3K^{\frac{1}{2}} + \phi_1 + 2\phi_2}; \\ -2a < b &\le 2a; \\ \frac{4a\phi_2 + 3b^2}{8a} \le c &\le \frac{4a\phi_1 + 3b^2}{8a}. \end{aligned}$$

Bounds for (a,b,c): Type 2. This subdivides into subtypes according as $a > 0$ or $a < 0$. For $a > 0$ we take $\phi_1 > \phi_2 > \phi_3$ and search the region

$$\begin{aligned} 0 < a &\le \frac{I - \phi_2^2}{3(\phi_2 - \phi_3)}; \\ -2a < b &\le 2a; \\ \frac{4a\phi_2 - \frac{4}{3}(I - \phi_2^2) + 3b^2}{8a} \le c &\le \frac{4a\phi_3 + 3b^2}{8a}. \end{aligned}$$

Then for $a < 0$ we take $\phi_1 < \phi_2 < \phi_3$ and search over

$$\begin{aligned} 0 < -a &\le \frac{I - \phi_2^2}{3(\phi_3 - \phi_2)}; \\ -2|a| < b &\le 2|a|; \\ \frac{4a\phi_2 - \frac{4}{3}(I - \phi_2^2) + 3b^2}{8a} \ge c &\ge \frac{4a\phi_3 + 3b^2}{8a}. \end{aligned}$$

Bounds for (a,b,c): Type 3. Here we let ϕ be the unique real root of (3.6.9), and search

$$\begin{aligned} \frac{1}{3}\phi - \sqrt{\frac{4}{27}(\phi^2 - I)} \le a &\le \frac{1}{3}\phi + \sqrt{\frac{4}{27}(\phi^2 - I)}; \\ -2|a| < b &\le 2|a|; \\ \frac{9a^2 - 2a\phi + \frac{1}{3}(4I - \phi^2) + 3b^2}{8|a|} \le c.\mathrm{sign}(a) &\le \frac{4a\phi + 3b^2}{8|a|}. \end{aligned}$$

The syzygy sieve. Recall the seminvariant syzygy

$$27r^2 = p^3 - 48Ia^2p - 64Ja^3 = s(a,p), \tag{3.6.7}$$

say, where $p = 3b^2 - 8ac$ and $r = b^3 + 8a^2d - 4abc$. For fixed I, J the expression $s(a,p)$ is a polynomial in a, b and c, which we require to be 27 times an integer square. We can set up a quadratic sieve as follows: for each of several sieving moduli m we create and initialize an $m \times m$ array indicating whether $s(a,p)$ is 27 times a square modulo m, for each pair (a,p) modulo m. We take one of the moduli to be 9 and use it to force the right-hand side of (3.6.7) to be divisible by 27; it will certainly be positive, as this is ensured by the bounds on c.

For each (a,b,c) in the region searched, we check that it passes the sieving test; it is then quite likely that $s(a,p)$ will be 27 times a square, since it is so modulo a large modulus and is positive. We then test whether this is the case, discarding (a,b,c) if not, and if so we then find r. We can take $r > 0$, since the quartics with coefficients (a,b,c,d,e) and $(a,-b,c,-d,e)$ are equivalent, with opposite signs of their respective r-seminvariants. In fact, we treat the triples $(a,\pm b,c)$ together in practice.

Implementation note: It is worth pointing out that a large proportion of the running time of our algorithm is spent testing whether large integers are squares (given that they are positive and congruent to squares modulo several carefully chosen moduli), and find their integer square root if so. This is needed here, and in our searches for rational points, both on the elliptic curve directly, and on the homogeneous spaces. Hence it is crucial that we have access to efficient procedures for this in the multiprecision integer package we use.

Solving for d and e. Given integers a, b, c, r satisfying (3.6.7) with $p = 3b^2 - 8ac$, we can solve for d and e, setting

$$d = (r - b^3 + 4abc)/(8a^2) \qquad \text{and} \qquad e = (I + 3bd - c^2)/(12a).$$

This will certainly give rational values for d and e; we must check that they are integral, discarding the triple (a,b,c) if not. If they are, we have integral coefficients (a,b,c,d,e) of a quartic $g(x)$ with invariants I and J in the search region, which we add to our list for further processing.

Solving for the roots of $g(x)$. For later use, when we check for triviality, and again when we search for rational points on the homogeneous spaces, we will need to know the real roots of the quartic $g(x)$ we have constructed. Although the formulae for finding the roots of quartic are well-known, we give them here: since we already know the roots of the resolvent cubic, there is very little work remaining.

For $i = 1,2,3$ we set $z_i = (4a\phi_i + p)/3$ where the ϕ_i are the three roots of (3.6.9). The product of these quantities is r^2 (from (3.6.7) again), and we form their square roots with product r by setting $w_1 = \sqrt{z_1}$, $w_2 = \sqrt{z_2}$, and $w_3 = r/(w_1w_2)$. Then the roots of $g(x)$ are

$$\begin{aligned} x_1 &= (\ \ w_1 + w_2 - w_3 - b)/(4a), \\ x_2 &= (\ \ w_1 - w_2 + w_3 - b)/(4a), \\ x_3 &= (-w_1 + w_2 + w_3 - b)/(4a), \\ x_4 &= (-w_1 - w_2 - w_3 - b)/(4a). \end{aligned}$$

We will not give here a pseudo-code algorithm for the search for quartics, as it is straightforward in principle, although in practice it needs careful book-keeping. As this is the most

time-consuming part of the whole procedure, particularly when the second, larger, pair of invariants must be used, it is important to make the implementation code as efficient as possible.

At the end of this step we will have a list of quartics with the desired invariants. We now discard any which are equivalent to earlier ones, or are not locally soluble at some prime p, and try to find rational roots on the remainder. In practice we may choose to apply these tests in a different order, such as not bothering to check equivalences between quartics which are not locally soluble.

Step 3: Testing triviality.

For each quartic $g(x)$ in the list, we already know its roots x to reasonable precision. If x is rational, then ax is integral, which we can test. If we suspect that ax is equal to an integer n to within some working tolerance, we can check whether n/a is a root of $g(x)$ using exact arithmetic.

Step 3: Testing equivalence of quartics.

With each quartic we find with the right invariants, we store its coefficients, type, roots and seminvariants p and r. We also compute and store the number of roots of the quartic (including roots at infinity) modulo each of several primes not dividing its discriminant, as these numbers are clearly invariant under equivalence.[9]

When testing equivalence of two quartics, we first check that their invariants and type are the same, as well as their numbers of roots modulo these primes. If this is the case, we use a general test for equivalence (valid over any field) from [20], which we state here.[10]

PROPOSITION 3.6.3. *Let g_1 and g_2 be quartics over the field K, both having the same invariants I and J, and with leading coefficients a_i and seminvariants p_i and r_i for $i = 1, 2$. Then g_1 is equivalent to g_2 over K if and only if the quartic $u^4 - 2pu^2 - 8ru + s$ has a root in K, where*

$$\begin{aligned} p &= (32a_1a_2I + p_1p_2)/3, \\ r &= r_1r_2, \qquad \text{and} \\ s &= (64I(a_1^2p_2^2 + a_2^2p_1^2 + a_1a_2p_1p_2) - 256a_1a_2J(a_1p_2 + a_2p_1) - p_1^2p_2^2)/27. \end{aligned}$$

The quantities p, r and s in this proposition will be integers when g_1 and g_2 are integral. Converting the proposition into an algorithm is straightforward.

Step 5: Testing local and global solubility.

This is carried out using the procedures and strategy described earlier.

Step 6: Final computation of the rank.

The number of quartics found (up to equivalence) which are everywhere locally soluble is n_2, the order of the 2-Selmer group. This must be a power of 2, say $n_2 = 2^{e_2}$, which serves as a check on our procedures. The number n_1 with a rational point is also a power of 2, say $n_1 = 2^{e_1}$, equal to the order of $E(\mathbb{Q})/2E(\mathbb{Q})$. If we have found rational points on all n_2 locally soluble quartics, then certainly $n_1 = n_2$, so that $Ш(E/\mathbb{Q})[2]$ is trivial, and the rank of $E(\mathbb{Q})$ is $e_1 - e_0$ where $|E(\mathbb{Q})[2]| = 2^{e_0}$ with $e_0 = 0$, 1 or 2. The rank is equal to the Selmer rank $e_2 - e_0$ in this case. (Usually $e_0 = 0$ when we are using this method.)

As before, we may not have found global points on all the locally soluble quartics; if the number on which we have points is $\tilde{n}_1$ with $\tilde{n}_1 < n_2$ then we only know that $\tilde{n}_1 \le n_1 \le n_2$. If

[9]This was suggested to us by S. Siksek.

[10]The algorithm presented here only applies to quartics. In the First Edition we presented a different algorithm, described in [3], which is messier to implement, but which generalizes more readily to more general situations, such as testing the equivalence of binary forms of higher degree.

$\tilde{n}_1$ is not a power of 2, we will know that $n_1 > \tilde{n}_1$, so that at least some of our locally soluble quartics must have rational points which we have not found. In this case, we replace $\tilde{n}_1$ by the next highest power of 2, say $\tilde{n}_1 = 2^{\tilde{e}_1}$. Then we have bounds on the rank, namely

$$\tilde{e}_1 - e_0 \leq e_1 - e_0 = \mathrm{rank}(E(\mathbb{Q})) \leq e_2 - e_0,$$

and on the order of $Ш(E/\mathbb{Q})[2]$:

$$|Ш(E/\mathbb{Q})[2]| \leq n_2/\tilde{n}_1.$$

One final point: from the Selmer conjecture, we expect the Selmer rank $e_2 - e_0$ to differ from the actual rank $e_1 - e_0$ by an even number, so that $e_2 \equiv e_1 \pmod 2$. This would also follow from the conjecture that $Ш(E/\mathbb{Q})$ is finite, since then its order is known to be a perfect square, so that n_2/n_1 must be a square. So if we find that $e_2 \not\equiv \tilde{e}_1 \pmod 2$, then we suspect that the rank is at least one more than our lower bound, and can output a comment to this effect, though of course we will not have proved that the rank is greater than our lower bound. In some cases, such as for a modular curve where we know the sign of the functional equation, we may have other conjectural evidence for the parity of the rank.

Step 7: Recovering points on E.

Each quartic $g(x)$ for which the homogeneous space $y^2 = g(x)$ has a rational point R leads to a rational point $P = \xi(R)$ on the model $E_{I,J}$ of our curve E, via the formula (3.6.8) given above. If we apply this formula to all the inequivalent quartics with rational points which we found in computing the rank of E, we will have a complete set of coset representatives for $2E(\mathbb{Q})$ in $E(\mathbb{Q})$, provided that $\tilde{n}_1 = n_1$. In cases where we have rounded up $\tilde{n}_1$ to the nearest power of 2, we will still have generators for $E(\mathbb{Q})/2E(\mathbb{Q})$, and can fill in the missing coset representatives if we wish.

This completes our description of algorithms for determining the Mordell-Weil group $E(\mathbb{Q})$.

3.7 The period lattice

In this section we show how to compute the complex periods for an elliptic curve defined over the complex numbers. We used this in our investigation of modular curves to check that the exact integral equations we found (after rounding the approximate computed values of c_4 and c_6) did have the correct periods; and also in our method for computing isogenous curves, which we describe in the following section.

Let E be an elliptic curve defined over the complex numbers $\mathbb{C}$, given by a Weierstrass equation. We wish to compute periods λ_1 and λ_2 which are a $\mathbb{Z}$-basis for the period lattice Λ of E. We do this using Gauss's arithmetic–geometric mean (AGM) algorithm. Write the equation for E in the form

$$\left(y + \frac{a_1x + a_3}{2}\right)^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4} = (x - e_1)(x - e_2)(x - e_3),$$

where the roots e_i are found as complex floating-point approximations (using Cardano's formula, say). Then the periods are given by

$$\begin{aligned} \lambda_1 &= \frac{\pi}{\mathrm{AGM}(\sqrt{e_3 - e_1}, \sqrt{e_3 - e_2})}, \\ \lambda_2 &= \frac{\pi i}{\mathrm{AGM}(\sqrt{e_3 - e_1}, \sqrt{e_2 - e_1})}. \end{aligned} \tag{3.7.1}$$

Notice that in general this involves the AGM of pairs of complex numbers. This is a multi-valued function: at each stage of the AGM algorithm we replace the pair (z, w) by $(\sqrt{zw}, \frac{1}{2}(z+w))$, and must make a choice of complex square root. It follows from work of Cox (see [11]) that while a different set of choices does lead to a different value for the AGM, the periods we obtain this way will nevertheless always be a $\mathbb{Z}$-basis for the full period lattice Λ. We have found this to be the case in practice, where we always choose a square root with positive real part, or with positive imaginary part when the real part is zero. The computation of λ_1 and λ_2 by this method is very fast, as the AGM algorithm converges extremely quickly, even in its complex form. As a check on the values obtained, in each case we recomputed the invariants c_4 and c_6 of each curve from these computed periods λ_1 and λ_2, using the standard formulae given in Chapter II; in every case we obtained the correct values (known exactly from the coefficients of the minimal Weierstrass equation) to within computational accuracy.

If the curve is defined over $\mathbb{R}$, we can avoid the use of the complex AGM, and also arrange that λ_1 is a positive real period, as follows. First suppose that all three roots e_i are real; order the roots so that $e_3 > e_2 > e_1$, and take the positive square root in the above formulae. Then we may use the usual AGM of positive reals in (3.7.1), and thus obtain a positive real value for λ_1 and a pure imaginary value for λ_2. This is the case where the discriminant $\Delta > 0$ and the period lattice is rectangular. When $\Delta < 0$ there is one real root, say e_3, and $e_2 = \overline{e_1}$. If $\sqrt{e_3 - e_1} = z = s + it$ with $s > 0$ then $\sqrt{e_3 - e_2} = \overline{z} = s - it$, so that $\lambda_1 = \pi/\mathrm{AGM}(z, \overline{z}) = \pi/\mathrm{AGM}(|z|, s)$ which is also real and positive.

3.8 Finding isogenous curves

Given an elliptic curve E defined over $\mathbb{Q}$, we now wish to find all curves E' isogenous to E over $\mathbb{Q}$. The set of all such curves is finite (up to isomorphism), and any two curves in the isogeny class are linked by a chain of isogenies of prime degree l. Thus it suffices to be able to compute l-isogenies for prime l, if we can determine those l for which rational l-isogenies exist. The latter question can be rather delicate in general, and we have to have a completely automatic algorithmic procedure if we are to apply it to several thousand curves, such as we had to when preparing the tables.

When the conductor N of E is square-free, so that E has good or multiplicative reduction at all primes, E is called semi-stable. In this case, a result of Serre (see [53]) says that either E or the isogenous curve E' has a rational point of order l, and so by Mazur's result already mentioned, l can only be 2, 3, 5 or 7. Moreover, if a curve E possesses a rational point of order l, then the congruence $1 + p - a_p \equiv 0 \pmod{l}$ holds for all primes p not dividing Nl, so the presence of such a point is easy to determine, even if it is not E itself but the isogenous curve E' which possesses the rational l-torsion, since the trace of Frobenius a_p is isogeny-invariant.

If E is not semi-stable we argue as follows. The existence of a rational l-isogeny is purely a function of the j-invariant j of E: in fact, pairs (E, E') of l-isogenous curves parametrize the modular curve $X_0(l)$ whose non-cuspidal points are given by the pairs $(j(E), j(E'))$. For $l = 2$, 3, 5, 7 or 13 the genus of $X_0(l)$ is zero, and infinitely many rational j occur. The only other values of l for which rational l-isogenies occur are $l = 11, 17, 19, 37, 43, 67$, and 163, and these occur for only a small finite number of j-invariants (see below). The fact that no other l occur is a theorem of Mazur (see [39] and [40]), related to the theorem limiting the rational torsion which we quoted earlier in Section 3.3 of this chapter. These extra values occur only for curves with CM (complex multiplication, see the next section), apart from $l = 17$ (where $X_0(l)$ has genus 1) and the exotic case $l = 37$ studied by Mazur and Swinnerton–Dyer in [41] (where $X_0(l)$ has genus 2).

For isogenies of non-prime degree m, the degrees which occur are: $m \leq 10$, and $m = 12$, 16 18, and 25 (where $X_0(l)$ has genus 0, infinitely many cases); and finally $m = 14$, 15, 21, and

27. The latter occur first for conductors $N = 49$ (with CM), $N = 50$, $N = 162$ and $N = 27$ (with CM) respectively. See [2, pages 78–80] for more details.

Thus our procedure is:

- If N is square-free, try $l = 2, 3, 5, 7$ only;
- else try $l = 2, 3, 5, 7$ and 13 in all cases; and
- if $j(E) = -2^{15}$, -11^2, or $-11 \cdot 131^3$, try also $l = 11$;
- if $j(E) = -17^2 \cdot 101^3/2$ or $-17 \cdot 373^3/2^{17}$, try also $l = 17$;
- if $j(E) = -96^3$, try also $l = 19$;
- if $j(E) = -7 \cdot 11^3$ or $-7 \cdot 137^3 \cdot 2083^3$, try also $l = 37$;
- if $j(E) = -960^3$, try also $l = 43$;
- if $j(E) = -5280^3$, try also $l = 67$;
- if $j(E) = -640320^3$, try also $l = 163$.

Now we turn to the question of finding all curves (if any) which are l-isogenous to our given curve E for a specific prime l. The kernel of the isogeny is a subgroup A of $E(\overline{\mathbb{Q}})$ which is defined over $\mathbb{Q}$, but the points of A may not be individually rational points. If we have the coordinates of the points of a subgroup of E of order l defined over K, we may use Vélu's formulae in [68] to find the corresponding l-isogenous curve. Finding such coordinates by algebraic means is troublesome, except when the subgroup is point-wise defined over K, and instead we resort to a floating-point method.

The case $l = 2$ is simpler to describe separately. Obviously in this case the subgroup of order 2 defined over $\mathbb{Q}$ must consist of a single rational point P of order 2 together with the identity. We have already found such points, if any, in computing the torsion. There will be 0, 1 or 3 of them according to the number of rational roots of the cubic $4x^3 + b_2x^2 + 2b_4x + b_6$. If x_1 is such a root, then $P = (x_1, y_1)$ has order 2, where $y_1 = -(a_1x_1 + a_3)/2$. As a special case of Vélu's formulae we find that the isogenous curve E' has coefficients $[a'_1, a'_2, a'_3, a'_4, a'_6] = [a_1, a_2, a_3, a_4 - 5t, a_6 - b_2t - 7w]$ where

$$t = (6x_1^2 + b_2x_1 + b_4)/2 \qquad \text{and} \qquad w = x_1t.$$

Note that the point (x_1, y_1) need not be integral even when E has integral coefficients a_i, but that $4x_1$ and $8y_1$ are certainly integral, by the formulae given; thus the model just given for the isogenous curve may need scaling by a factor of 2 to make it integral.

The simpler formula for a curve in the form $y^2 = x^3 + cx^2 + dx$ and the point $P = (0, 0)$ was given in the previous section: the formulae just given take the curve $[0, c, 0, d, 0]$ to $[0, c, 0, -4d, -4cd]$, which transforms to $[0, -2c, 0, c^2 - 4d, 0]$ after replacing x by $x - c$. The relation between the two formulae is given by $c = 12x_1 + b_2$ and $d = 16t$.

For reference we give here similar algebraic formulae for l-isogenies for $l = 3$ and $l = 5$, from Laska's book [35]. In each case we assume that the curve E is given by an equation of the form $y^2 = x^3 + ax + b$, and the isogenous curve E' by $y^2 = x^3 + Ax + B$. Each subgroup of E of order l is determined by a rational factor of degree $(l - 1)/2$ of the l-division polynomial of degree $(l^2 - 1)/2$, whose roots are the x-coordinates of the points in the subgroup. The simplest case is $l = 3$, where there is just one x-coordinate, which must be rational.

$l = 3$. Let ξ be a root of the 3-division polynomial $3x^4 + 6ax^2 + 12bx - a^2$. Then the 3-isogenous curve E' is given by

$$A = -3(3a + 10\xi^2)$$
$$B = -(70\xi^3 + 42a\xi + 27b).$$

$l = 5$. Let $x^2 + h_1x + h_2$ be a rational factor of the 5-division polynomial $5x^{12} + 62ax^{10} + 380bx^9 - 105a^2x^8 + 240abx^7 - (300a^3 + 240b^2)x^6 - 696a^2bx^5 - (125a^4 + 1920ab^2)x^4 - (1600b^3 + 80a^3b)x^3 - (50a^5 + 240a^2b^2)x^2 - (100a^4b + 640ab^3)x + (a^6 - 32a^3b^2 - 256b^4)$. Then the 5-isogenous curve E' is given by

$$A = -19a - 30(h_1^2 - 2h_2)$$
$$B = -55b - 14(15h_1h_2 - 5h_1^3 - 3ah_1).$$

A similar formula is given in [35] for $l = 7$, where A and B are given in terms of a, b and the coefficients of a factor $x^3 + h_1x^2 + h_2x + h_3$ of the 7-division polynomial. Rather than take up space by giving the latter here, we refer the reader to [35, page 72].

Now we turn to Vélu's formulae in the case of an odd prime l. Let $P = (x_1, y_1)$ be a point of order l in $E(\overline{\mathbb{Q}})$, and set $kP = (x_k, y_k)$ for $1 \le k \le (l-1)/2$. Define

$$t_k = 6x_k^2 + b_2x_k + b_4 \qquad \text{and} \qquad u_k = 4x_k^3 + b_2x_k^2 + 2b_4x_k + b_6,$$

and then set

$$t = \sum_{k=1}^{(l-1)/2} t_k \qquad \text{and} \qquad w = \sum_{k=1}^{(l-1)/2} (u_k + x_kt_k).$$

Then the isogenous curve E' has coefficients $[a_1, a_2, a_3, a_4 - 5t, a_6 - b_2t - 7w]$ as before. Again, these may not be integral, even when the original coefficients were; but since the x_k are the roots of a polynomial of degree $(l-1)/2$ with integral coefficients and leading coefficient l^2 (the so-called l-division equation), we must have l^2x_k integral. Thus a scaling factor of l will certainly produce an integral equation.

We make these remarks on integrality as our method is to find the coordinates x_k and y_k as real floating-point approximations, and thus to determine the coefficients of any curves l-isogenous to E over $\mathbb{R}$; there will always be exactly two such curves over $\mathbb{R}$, but of course they will not necessarily be defined over $\mathbb{Q}$. As we will only know the coefficients a_i' of the isogenous curves approximately, we wish to ensure that if they are rational then they will in fact be integral, so that we will be able to recognize them as such.

First we find the period lattice Λ of E, as described in the previous section. The $\mathbb{Z}$-basis $[\lambda_1, \lambda_2]$ of Λ is normalized as follows: there are two cases to consider, according as $\Delta > 0$ (first or 'harmonic' case) or $\Delta < 0$ (second or 'anharmonic' case). In both cases λ_1 is real (the least positive real period); in the first case, λ_2 is pure imaginary, while in the second case, $2\lambda_1 - \lambda_2$ is pure imaginary. We can also ensure that $\tau = \lambda_2/\lambda_1$ is in the upper half-plane; however we can not simultaneously arrange that τ is in the usual fundamental region for $\mathrm{SL}(2, \mathbb{Z})$, and this needs to be remembered when evaluating the Weierstrass functions below.

Of the $l + 1$ subgroups of $\mathbb{C}/\Lambda$ of order l, the two defined over $\mathbb{R}$ are the one generated by $z = \lambda_1/l$ (in both cases), and in the first case, the one generated by $z = \lambda_2/l$, or in the second case, the one generated by $z = (\lambda_1 - 2\lambda_2)/l$. Thus z/λ_1 is either $1/l$, τ/l, or $(1 - 2\tau)/l$. Let $\wp(z; \tau)$ denote the Weierstrass $\wp$-function relative to the lattice $[1, \tau]$. Then we have

$$x_k = \wp(kz\lambda_1^{-1}; \tau)\lambda_1^{-2} - \frac{1}{12}b_4 \qquad \text{and} \qquad y_k = \frac{1}{2}\left(\wp'(kz\lambda_1^{-1}; \tau)\lambda_1^{-3} - a_1x_k - a_3\right).$$

Here we have had to take account of the lattice scaling $[\lambda_1, \lambda_2] = \lambda_1[1, \tau]$, and also of the fact that $(\wp(z), \wp'(z))$ is a point on the model of E of the form $y^2 = 4x^3 - g_2x - g_3 = 4x^3 - (c_4/12)x - (c_6/216)$ rather than a standard model where the coefficient of x^3 is 1.

We evaluate these points of order l numerically for $k = 1, 2, \ldots, (l-1)/2$, for each of the two values of z (depending on whether we are in case 1 or case 2). Substituting into Vélu's

formulae, we obtain in each case the real coefficients a_i' of a curve which is l-isogenous to E over $\mathbb{R}$. If these coefficients are close to integers we round them and check that the resulting curve over $\mathbb{Q}$ has the same conductor N as the original curve E. If not, we also test the curve with coefficients $l^i a_i'$.

The resulting program finds l-isogenous curves very quickly for any given prime l. We run it for all primes l in the set determined previously, applying it recursively to each new curve found until we have a set of curves closed under l-isogeny for these values of l. Since the set of primes l for which a rational l-isogeny exists is itself an isogeny invariant, once we have finished processing the first curve in the class, we will already know which primes l to use for all the remaining curves.

Some care needs to be taken with a method of computation such as this, where we use floating-point arithmetic to find integers. The series we use to compute the periods and the Weierstrass function and its derivative all converge very quickly, so that we can compute the a_i' to whatever precision is available, though of course in practice some rounding error is bound to arise. When we test whether a floating-point number is 'approximately an integer' in the program, we must make a judgement on how close is close enough. With too relaxed a test, we will find too many curves are 'approximately integral'; usually these will fail the next hurdle, where we test the conductor, but this takes time to check (using Tate's algorithm). On the other hand, too strict a test might mean that we miss some rational isogenies altogether, which is far more serious. In compiling the tables, there was only one case which caused trouble after the program had been finely tuned. The resulting error resulted in a curve (916B1) being erroneously listed as 3-isogenous to itself in the first (preprint) edition of the tables; this is possible only when a curve has complex multiplication, which is not the case here, though it does not often occur even in the complex multiplication case (see the remarks in the next section). Unfortunately the error was not noticed in the automatic generation of the typeset tables, and I am grateful to Elkies for spotting it.[11] The curve $E = [0, 0, 0, -1013692, 392832257]$ has three real points of order 2, two of which are equal to seven significant figures; the period ratio is approximately $7i$. One of the curves 2-isogenous to E over $\mathbb{R}$ has coefficients $[0, 0, 0, -1013691.999999999992, 392832257.000000006]$, which are extremely close to those of E itself. Thus this new curve, which is not defined over $\mathbb{Q}$, passed both our original tests (the coefficients are extremely close to integers, and the rounded coefficients are those of a curve of the right conductor, namely E itself). After this example was discovered, we inserted an extra line in the program, to print a warning whenever a supposedly isogenous curve was the original curve itself, and reran the program on all 2463 isogeny classes (which only takes a few minutes of machine time). The result was that expected, namely that 916B1 is the only curve for which this phenomenon occurs within the range of the tables[12]. There is no example of a curve actually l-isogenous to itself with conductor less than 1000.

Our original implementation of this algorithm in Algol68 used a precision of approximately 30 significant figures for its real and complex arithmetic, which was sufficient to find all the isogenous curves up to conductor 1000. However, our implementation in C++ misses several isogenous curves when using standard double precision, with approximately 15 digits (though this runs very quickly); we need to use a multiprecision floating-point package (such as the one included in `LiDIA`) to obtain a satisfactory working program, though the resulting code runs very much slower. In our extended computations to conductor 5077, we have computed the isogenies independently using both a C++/`LiDIA` program and a PARI program, and the results agree.

When we were initially persuaded to extend the tables to include isogenous curves as well as the modular curves themselves, we were afraid that the total number of resulting curves

[11] This error also somehow survived into the first edition of this book, despite these comments in the text.

[12] Another example of the same type occurs for curve 1342C3, where the period ratio is approximately $9.5i$.

would be rather larger than it turned out to be. On average, we found that the number of curves per isogeny class was 5113/2463, or just under 2.08. We do not know of any asymptotic analysis, or even a heuristic argument, which would predict an average number of two curves per class. However, it is dangerous to generalize from the limited amount of data which we have available. In the extended computations to conductor 5077, the ratio slowly diminishes; for all curves up to this conductor, the ratio is 31570/17583 = 1.795.

3.9 Twists and complex multiplication

Traces of Frobenius.

If E is given by a standard minimal Weierstrass equation over $\mathbb{Z}$, then for all primes p of good reduction the trace of Frobenius a_p is given by

$$a_p = 1 + p - |E(\mathbb{F}_p)|\,.$$

If E has bad reduction at p, this same formula gives the correct value for the pth Fourier coefficient of the L-series of E.

Since in our applications we never needed to compute a_p for large primes p, we used a very simple method to count the number of points on E modulo p. First, for all primes p in the desired range (say $3 \le p \le 1000$; $p = 2$ would be dealt with separately), we precompute the number $n(t,p)$ of solutions to the congruence $s^2 \equiv t \pmod p$. Then we simply compute

$$a_p = p - \sum_{x=0}^{p-1} n(4x^3 + b_2x^2 + 2b_4x + b_6, p).$$

This was sufficient for us to compute a_p for all $p < 1000$ for all the curves in the table, which we did to compare with the corresponding Hecke eigenvalues. For large p, there are far more efficient methods, such as the baby-step giant-step method or Schoof's algorithm (see [51]). Details of these may be found in [9]. More recently, even better algorithms have been developed, by Atkin, Elkies, Morain, Müller and others. For example, Morain and Lercier in 1995 successfully computed the number of points on the curve $[0,0,0,4589,91228]$ over $\mathbb{F}_p$ for $p = 10^{499} + 153$, a prime with 500 decimal digits. This took 4200 hours of computer time.

Twists.

A twist of a curve E over $\mathbb{Q}$ is an elliptic curve defined over $\mathbb{Q}$ and isomorphic to E over $\overline{\mathbb{Q}}$ but not necessarily over $\mathbb{Q}$ itself. Thus the set of all twists of E is the set of all curves with the same j-invariant as E. These can be simply described, as follows.

First suppose that $c_4 \ne 0$ and $c_6 \ne 0$; equivalently, $j \ne 1728$ and $j \ne 0$ (respectively). Then the twists of E are all quadratic, in that they become isomorphic to E over a quadratic extension of $\mathbb{Q}$. For each integer d (square-free, not 0 or 1), there is a twisted curve $E*d$ with invariants d^2c_4 and d^3c_6, which is isomorphic to E over $\mathbb{Q}(\sqrt{d})$. If E has a model of the form $y^2 = f(x)$ with $f(x)$ cubic, then $E*d$ has equation $dy^2 = f(x)$. A minimal model for $E*d$ may be found easily by the Laska–Kraus–Connell algorithm. The conductor of $E*d$ is only divisible by primes dividing ND, where D is the discriminant of $\mathbb{Q}(\sqrt{d})$. The simplest case is when $\gcd(D,N) = 1$; then $E*d$ has conductor ND^2. More generally, if $D^2 \nmid N$ then $E*d$ has conductor $\mathrm{lcm}(N, D^2)$, but if $D^2 \mid N$ then the conductor may be smaller; for example, $(E*d)*d$ is isomorphic to E, so has conductor N again.

Twisting commutes with isogenies, in the sense that if two curves E, F are l-isogenous then so are their twists $E*d$, $F*d$. If E has no complex multiplication (see below), then the structure of the isogeny class of E is a function of $j(E)$ alone.

The trace of Frobenius of $E * d$ at a prime p not dividing N is $\chi(p)a_p$, where χ is the quadratic character associated to $\mathbb{Q}(\sqrt{d})$ and a_p is the trace of Frobenius of E. Thus if E is modular, attached to the newform f, then $E * d$ is also modular and attached to the twisted form $f \otimes \chi$, in the notation of Chapter 2.

When $j = 0$ (or equivalently $c_4 = 0$), E has an equation of the form $y^2 = x^3 + k$ with $k \in \mathbb{Z}$ non-zero and free of sixth powers. Such curves have complex multiplication by $\mathbb{Z}[(1+\sqrt{-3})/2]$. Two such curves with parameters k, k' are isomorphic over $\mathbb{Q}(\sqrt[6]{k/k'})$.

Similarly, when $j = 1728$ (or equivalently $c_6 = 0$), E has an equation of the form $y^2 = x^3 + kx$ with $k \in \mathbb{Z}$ non-zero and free of fourth powers. Such curves have complex multiplication by $\mathbb{Z}[\sqrt{-1}]$. Two such curves with parameters k, k' are isomorphic over $\mathbb{Q}(\sqrt[4]{k/k'})$.

Complex multiplication.

Each of the 13 imaginary quadratic orders $\mathfrak{O}$ of class number 1 has a rational value of $j(\mathfrak{O}) = j(\omega_1/\omega_2)$, where $\mathfrak{O} = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$. Elliptic curves E with $j(E) = j(\mathfrak{O})$ have complex multiplication: their ring of endomorphisms defined over $\mathbb{C}$ is isomorphic to $\mathfrak{O}$. In all other cases the endomorphism ring of an elliptic curve defined over $\mathbb{Q}$ is isomorphic to $\mathbb{Z}$, since an elliptic curve with complex multiplication by an order of class number $h > 1$ has a j-invariant which is not rational, but algebraic of degree h over $\mathbb{Q}$.

We give here a table of triples (D, j, N) where $j = j(\mathfrak{O})$ for an order of discriminant D, and N is the smallest conductor of an elliptic curve defined over $\mathbb{Q}$ with this j-invariant. All but the last three values ($D = -43, -67, -163$) have $N < 1000$ and so occur in the tables.

D	-4	-16	-8	-3	-12	-27	-7	-28	-11	-19	-43	-67	-163
j	12^3	66^3	20^3	0	$2 \cdot 30^3$	$-3 \cdot 160^3$	-15^3	255^3	-32^3	-96^3	-960^3	-5280^3	-640320^3
N	32	32	256	27	36	27	49	49	121	361	43^2	67^2	163^2

If E has complex multiplication by the order $\mathfrak{O}$ of discriminant D, then the twist $E * D$ is isogenous to E, though not usually isomorphic to E (over $\mathbb{Q}$). Indeed, the only cases where E is isomorphic to $E * D$ are $D = -4$ and $D = -16$ with $j(E) = 1728$: the curves $y^2 = x^3 + 16kx$ and $y^2 + 256kx$ are twists of, and isomorphic to, $y^2 = x^3 + kx$. Since curves are isogenous if and only if they have the same L-series by Faltings's Theorem (see [22]), this implies that E has complex multiplication if and only if $a_p = \chi(p)a_p$ for all primes p, where χ is the quadratic character as above. Thus $a_p = 0$ for half the primes p, namely those for which $\chi(p) = -1$. This gives an alternative way of recognizing a curve with complex multiplication, from its traces of Frobenius. This is particularly convenient in the case of modular curves, where we compute the a_p first, and will always know when a newform f, and hence the associated curve E_f, has complex multiplication. For, in such a case, we must have $D^2 \mid N$ and $f = f \otimes \chi$, which we may easily check from the tables.