Elliptic Curve Handbook

Ian Connell
February, 1999

http://www.math.mcgill.ca/connell/

Foreword

The first version of this handbook was a set of notes of about 100 pages handed out to the class of an introductory course on elliptic curves given in the 1990 fall semester at McGill University in Montreal. Since then I have added to the notes, holding to the principle: *If I look up a certain topic a year from now I want all the details right at hand, not in an "exercise", so if I've forgotten something I won't waste time.* Thus there is much that an ordinary text would either condense, or relegate to an exercise. But at the same time I have maintained a solid mathematical style with the thought of sharing the handbook.

Montreal,
August, 1996.

Contents

Chapter 1

Introduction to Elliptic Curves.

1.1 The a, b, c 's and $\Delta, j, \ldots$

We begin with a series of definitions of **elliptic curve** in order of increasing generality and sophistication. These definitions involve technical terms which will be defined at some point in what follows.

The most concrete definition is that of a curve E given by a nonsingular **Weierstrass equation**:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{1}$$

The coefficients a_i are in a field K and $E(K)$ denotes the set of all solutions $(x, y) \in K \times K$, together with the point O "at infinity" — to be explained in §1.3. We will see later why the a's are numbered in this way; to remember the Weierstrass equation think of the terms as being in a graded ring with

$$\begin{array}{cccccc} \text{weight} & \text{of} & x & = & 2 \\ \text{"} & \text{"} & y & = & 3 \\ \text{"} & \text{"} & a_i & = & i \end{array}$$

so that each term in the equation has weight 6. (This also "explains" the absence of a_5.)

A slightly more general definition is: a plane nonsingular cubic with a rational point (**rational** means the coordinates are in the designated field K and does not refer to the rational field $\mathbf{Q}$, unless of course $K = \mathbf{Q}$). An example of such a curve that is not a Weierstrass equation is the Fermat curve

$$x^3 + y^3 = 1, \quad \text{with points} \quad (x, y) = (1, 0),\ (0, 1),$$

assuming the characteristic of K, denoted $\operatorname{char} K$, is not 3. In Corollary 1.4.2 we will see how to transform such an equation into Weierstrass form.

More general still: a nonsingular curve of genus 1 with a rational point. (As we will explain later, conic sections — circles, ellipses, parabolas, and hyperbolas — have genus 0 which implies that they are not elliptic curves.) An example that is not encompassed by the previous definitions is

$$y^2 = 3x^4 - 2, \quad \text{with points} \quad (x, y) = (\pm 1, \pm 1),$$

assuming $\operatorname{char} K \neq 2, 3$. Proposition 1.2.1 below explains how to transform such quartic equations into Weierstrass form (without using $\sqrt[4]{3}$ or $\sqrt{-2}$!).

Alternative terminology which emphasizes the algebraic group structure: abelian variety of dimension 1.

More abstractly: E is a scheme over a base scheme S (e.g. spec K) which is proper, flat and finitely presented, equipped with a section ...: there is little point to state all the technicalities at this time. Suffice it to say that the work of Tate, Mazur and many others makes it plain that it is essential to know the language of schemes to understand the deeper arithmetic properties of elliptic curves. (More easily said than done!)

Now let us begin to fill in some details. Consider a Weierstrass equation (1), which we denote as E. If $\operatorname{char} K \neq 2$ we can complete the square by defining $\eta = y + (a_1 x + a_3)/2$:

$$\eta^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4} \tag{2}$$

where
$$b_2 = a_1^2 + 4a_2, \quad b_4 = a_1 a_3 + 2a_4, \quad b_6 = a_3^2 + 4a_6. \tag{3}$$

If $\operatorname{char} K \neq 3$ we can complete the cube by setting $\xi = x + b_2/12$:

$$\eta^2 = \xi^3 - \frac{c_4}{48}\xi - \frac{c_6}{864} \tag{4}$$

where
$$c_4 = b_2^2 - 24b_4, \quad c_6 = -b_2^3 + 36b_2 b_4 - 216b_6. \tag{5}$$

One then defines

$$b_8 = a_1^2 a_6 - a_1 a_3 a_4 + 4a_2 a_6 + a_2 a_3^2 - a_4^2, \tag{6}$$

and
$$\Delta = -b_2^2 b_8 - 8b_4^3 - 27b_6^2 + 9b_2 b_4 b_6. \tag{7}$$

The subscripts on the b's and c's are their weights. We refer to (1), (2) and (4) as the ***a*-form**, ***b*-form** and ***c*-form** respectively. The definitions (3) and (5)–(7) are made for all E, regardless of the characteristic of K, and the condition that the curve be nonsingular, and so define an elliptic curve, is that $\Delta \neq 0$, as we will explain in §1.5. Then one defines $j = c_4^3/\Delta$. For example

$$\textit{when} \ \ \operatorname{char} K = 2, \quad \Delta \neq 0 \Longrightarrow a_1 \ \textit{and} \ \ a_3 \ \ \textit{are not both zero.}$$

Thus
$$\kappa := 2y + a_1 x + a_3$$
is nonzero† for every elliptic curve E in any characteristic. When $\operatorname{char} K \neq 2$ we have $\kappa = 2\eta$ and κ is determined up to sign by x. Note that
$$\kappa^2 = 4x^3 + b_2 x^2 + 2b_4 x + b_6$$
is valid in all characteristics.

The **covariants** c_4, c_6 and the **discriminant** Δ have weights 4,6,12 respectively. The quantity j defined above when $\Delta \neq 0$ is called the ***j*-invariant**, or simply the **invariant** of E; its weight is 0.

It is often convenient to include Δ as a third covariant. Thus we say that
$$y^2 + y = x^3 - x^2 \qquad (\mathbf{A11})$$
has covariants 16, -152, -11, meaning that $c_4 = 16$, $c_6 = -152$ and $\Delta = -11$. The label **A11** is the standard catalog name of this elliptic curve as in [AntIV]; we put the letter first, rather than 11A, so that A11 can be used as the name of this curve in computer programs such as **aPecs**; see the appendix to this chapter. In [Cre92], which extends the catalog of [AntIV], the labelling has been modified (with the former notation given in parentheses) — this curve is denoted $\boldsymbol{A_3 11}$; by force of habit, we will use the notation of [AntIV] for curves contained in that catalog, and then use Cremona's notation for curves that are only in the larger catalog.

For convenience of reference, we collect these various definitions in a box:

$$\boxed{\begin{array}{lcl}
b_2 & = & a_1^2 + 4a_2\,, \\
b_4 & = & a_1 a_3 + 2a_4\,, \\
b_6 & = & a_3^2 + 4a_6, \\
b_8 & = & a_1^2 a_6 - a_1 a_3 a_4 + 4a_2 a_6 + a_2 a_3^2 - a_4^2\,, \\
c_4 & = & b_2^2 - 24b_4\,, \\
c_6 & = & -b_2^3 + 36 b_2 b_4 - 216 b_6\,, \\
\Delta & = & -b_2^2 b_8 - 8b_4^3 - 27b_6^2 + 9 b_2 b_4 b_6\,, \\
\kappa & = & 2y + a_1 x + a_3\,, \\
4b_8 & = & b_2 b_6 - b_4^2\,, \\
1728\Delta & = & c_4^3 - c_6^2\,, \\
j & = & c_4^3/\Delta \;=\; 1728 + c_6^2/\Delta\,.
\end{array}}$$

The last three lines in the box are identities that one can verify on the computer.

†as an element of the field $L = K(x, y)$ obtained as a quadratic extension $K(x)(y)$ of the transcendental extension $K(x)$, where y is defined by equation (1). As will be discussed in §1.6, L is called the *function field* of E, and $P = (x, y) \in E(L)$ is called a *generic point*.

Examples:

1. Suppose $\operatorname{char} K \neq 2$. Then Δ is 16 times the polynomial discriminant[†] of the cubic on the right side of the b-form (2):

$$\operatorname{Dis}\left(x^3 + (b_2/4)x^2 + (b_4/2)x + b_6/4\right) = \Delta/16.$$

Hence $\Delta = 0$ iff the cubic has a multiple root.

2. If $\operatorname{char} K \neq 2$ or 3, an alternative to the c-form is

$$\eta'^2 = \xi'^3 - 27\bar{c}_4\xi' - 54\bar{c}_6, \qquad \eta' = 6^3\eta, \qquad \xi' = 6^2\xi.$$

Caution: We have put bars on the c's because with the displayed values for the Weierstrass coefficients $a_1 = 0, \ldots, a_6 = -54\bar{c}_6$ the formulas give $c_4 = 6^4\bar{c}_4$, $c_6 = 6^6\bar{c}_6$. In the case of (4), bars are not necessary: the calculated c's are the same as the c's in the equation.

3.

$$y^2 = x^3 + bx + c$$

has covariants $\qquad -48b, \quad -864c, \quad -16(4b^3 + 27c^2).$

Thus provided $\Delta = -2^4 3^3 c^2 \neq 0$, $y^2 = x^3 + c$

has $\qquad c_4 = 0 \quad \text{and} \quad j = 0;$

and provided $\Delta = -64b^3 \neq 0$, $y^2 = x^3 + bx$

has $\qquad c_6 = 0 \quad \text{and} \quad j = 1728 = 12^3.$

4. "Generic j": provided $j \neq 0, 1728$,

$$y^2 + xy = x^3 - \frac{36}{j - 1728}x - \frac{1}{j - 1728}$$

has j-invariant $= j$; the covariants are

$$c_4 = -c_6 = \frac{j}{j - 1728}, \quad \text{and} \quad \Delta = \frac{j^2}{(j - 1728)^3}.$$

5. When K is the real field $\mathbf{R}$ we can take the equation in c-form $\eta^2 = \xi^3 + \cdots$. The cubic has either 1 or 3 real roots according as the discriminant Δ is negative or positive; thus as a real manifold there are 1 or 2 components. We will see in §1.3 that the addition of the point O at ∞ will compactify the curve.

On the following interleaving sheet there are plots of three examples (the same ones used in [Sil86,p.47]).

[†] in the usual sense $\operatorname{Dis}(f) = (-1)^{n(n-1)/2}\operatorname{Resultant}(f, f')$ where $n = \deg(f)$:

$$\begin{aligned}
\operatorname{Dis}(X^2 + aX + b) &= a^2 - 4b, \\
\operatorname{Dis}(X^3 + aX^2 + bX + c) &= -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2, \\
\text{in particular,}\quad \operatorname{Dis}(X^3 + bX + c) &= -4b^3 - 27c^2, \quad \text{and} \\
\operatorname{Dis}(X^4 + bX^2 + cX + d) &= 16b^4d - 4b^3c^2 - 128b^2d^2 \\
&\qquad + 144bc^2d - 27c^4 + 256d^3.
\end{aligned}$$

1.2 Quartic to Weierstrass

If K is a field, K^* denotes the multiplicative group and $\overline{K}$ denotes an algebraic closure.

Let F be a nonzero homogeneous polynomial in the variables U, W over the field K. Recall that a *root* of F is a ratio $U : W = \alpha : \beta$ corresponding to a linear factor $\beta U - \alpha W$ of F, where one but not both of α, β may be 0. For example, the homogeneous quartic

$$U^2W^2 - UW^3$$

has the double root $U : W = 1 : 0$ and the two simple roots $0 : 1$, $1 : 1$. If the degree of F is n then over $\overline{K}$, F has precisely n roots, some of which may be coincident.

Let K be a field of characteristic $\neq 2$, and consider the curve defined by an equation over K of the form $v^2 =$ a quartic in u with a rational point $(u, v) = (p, q)$. Replacing u by $u + p$ we can assume that $p = 0$:

$$v^2 = au^4 + bu^3 + cu^2 + du + q^2.$$

When $q \neq 0$, such a curve is birationally equivalent to one given by a Weierstrass equation:

Proposition 1.2.1 *Let K be a field with* $\operatorname{char} K \neq 2$ *and u, v transcendentals over K satisfying*

$$v^2 = au^4 + bu^3 + cu^2 + du + q^2, \qquad (\P)$$

where $a, b, c, d \in K$, and $q \in K^*$. Then

$$\begin{aligned} x &= (2q(v+q) + du)/u^2, \\ y &= (4q^2(v+q) + 2q(du + cu^2) - d^2u^2/2q)/u^3 \end{aligned}$$

satisfy the Weierstrass equation with

$$\begin{aligned} a_1 &= d/q, & a_2 &= c - d^2/4q^2, \\ a_3 &= 2qb, & a_4 &= -4q^2a, \end{aligned}$$

$$a_6 = a_2a_4 = a(d^2 - 4q^2c).$$

The discriminant Δ of this Weierstrass equation is 0 *iff the homogeneous quartic*

$$aU^4 + bU^3W + cU^2W^2 + dUW^3 + q^2W^4$$

has a repeated root in $\overline{K}$, i.e., *iff either $a = b = 0$ or the polynomial on the right in* (¶) *has a repeated root in $\overline{K}$.*

The inverse transformation is given by

$$u = (2q(x+c) - d^2/2q)/y, \qquad v = -q + u(ux - d)/2q.$$

In this birational correspondence, the point $(u, v) = (0, -q)$ *on* (¶) *corresponds to the point* $(x, y) = (-a_2, a_1a_2 - a_3)$ *on the Weierstrass curve.*†

Remarks. The proposition essentially covers all cases where $\Delta \neq 0$, as we can indicate now by anticipating some definitions and results that will be given later. Consider

$$v^2 = au^4 + bu^3 + cu^2 + du + e, \tag{\#}$$

where at least one of a, b is nonzero, and the polynomial on the right has no repeated roots in $\overline{K}$. Then (#) is birationally equivalent over K to a Weierstrass equation iff this curve has a rational place, which means that either

- there is a rational point $(u, v) = (p, q)$, and then either
  - (i) $q \neq 0$ — replace u with $u + p$ so the equation becomes that treated directly by the proposition; or
  - (ii) $q = 0$ — replace u, v with $1/u + p, v/u^2$ to obtain an equation of the type dealt with in (iv) below;
- or there is a rational place at ∞. This means that either
  - (iii) $a = q^2 \in K^{*2}$ — there are two rational places at ∞ (*cf.* Proposition 2.2.8(b)): replacing u by $1/u$ and v by v/u^2 puts (#) in the form treated by the proposition; or
  - (iv) $a = 0$ — (#) is essentially already in Weierstrass form: take $u = x/b$, $v = y/b$. When $e = q^2 \in K^{*2}$, this gives a Weierstrass equation different from that of the proposition; but the two Weierstrass equations can be transformed birationally one into the other.

The meaning of the inverse transformation is this: if x, y satisfy the Weierstrass equation then u, v defined as rational functions in x, y in this way satisfy (¶).

Proof. For all but the last statement of the proposition the verification is by direct calculation, nowadays best performed on the computer. (The theorem of Riemann-Roch discussed in Chapter 6 gives the theoretical explanation; see Corollary 6.1.17). For example to see when $\Delta = 0$, one calculates

- when $a \neq 0$, then $\Delta = 16D$ where D is the discriminant of the quartic on the right of (¶);
- when $a = 0$, $b \neq 0$, then $\Delta = 16b^2D$ where D is the discriminant of the cubic on the right of (¶);
- when $a = b = 0$ then $\Delta = 0$.

†John Cremona suggested adding this last statement.

To obtain the image of $(0, -q)$ we cannot simply substitute $u = 0$, $v = -q$ into the formulas for x and y since we get the indeterminate form $0/0$. L'Hôpital's rule affords the quickest way to obtain the answer: we differentiate the numerator and denominator of x twice with respect to u, and those of y three times, using $dv/du = (4au^3 + \cdots + d)/2v$ obtained by differentiating (¶), and then cancel common factors such as 3! from the numerator and denominator of the resulting fractions. Again the computer makes this relatively painless (and may tempt the reader to find the point (u, v) corresponding to $(x, y) = (-a_2, 0)$). The validity of the method for all K with $\operatorname{char} K \neq 2$ depends on the fact that the functions have perfectly usable Taylor expansions (there is no problem with factorials in denominators) which are most easily described in the field of formal power series as follows.

Regard u as an indeterminate so that the field of rational functions $K(u)$ is canonically a subfield of the field $K((u))$ of formal power series, *i.e.,* series of the form $\sum_N^\infty k_n u^n$ for some $N \in \mathbf{Z}$, $k_n \in K$. Now (¶) defines a quadratic extension $L = K(u)(v)$ of $K(u)$ and there are two embeddings $\phi : L \longrightarrow K((u))$ corresponding to the two square roots of $au^4 + \cdots$. The one that is relevant here is

$$\begin{aligned}\phi(v) &= -q\left(1 + \frac{d}{q^2}u + \frac{c}{q^2}u^2 + \frac{b}{q^2}u^3 + \frac{a}{q^2}u^4\right)^{1/2} \\ &= -q - \frac{d}{2q}u + \left[\frac{d^2}{8q^3} - \frac{c}{2q}\right]u^2 + \cdots .\end{aligned}$$

Induction (or at worst a reference to the general binomial theorem in [Con82]) shows that 2 is the only prime that occurs in denominators, and substitution yields

$$\begin{aligned}x &= \left[\frac{d^2}{4q^2} - c\right] + \left[\frac{-d^3}{8q^4} + \frac{cd}{2q^2} - b\right]u + \cdots, \\ y &= \left[\frac{-d^3}{4q^3} + \frac{cd}{q} - 2bq\right] + \left[\frac{5d^4}{32q^5} - \cdots\right]u + \cdots.\end{aligned}$$

When $u = 0$ these expressions reduce to $-a_2$ and $a_1a_2 - a_3$ respectively. ■

Example The curve $v^2 = 3u_1^4 - 2$ was mentioned in §1.1 (in a different notation) as an example of a curve of genus 1 with a rational point $(u_1, v) = (1, 1)$. To apply the proposition we substitute $u_1 = u + 1$, obtaining the curve

$$v^2 = 3u^4 + 12u^3 + 18u^2 + 12u + 1.$$

With $a, b, c, d, q = 3, 12, 18, 12, 1$, we find that

$$\begin{aligned}x &= 2(6u + v + 1)/u^2, \\ y &= 4(-9u^2 + 6u + v + 1)/u^3,\end{aligned}$$

satisfy the Weierstrass equation

$$y^2 + 12xy + 24y = x^3 - 18x^2 - 12x + 216.$$

We obtain a simpler Weierstrass equation by completing the square on the left and then the cube on the right: the equation becomes

$$y_1^2 = x_1^3 + 24x_1$$

where $$x_1 = x + 6, \quad y_1 = y + 6x + 12.$$

Using the notation

$$(u_1, v) \mapsto (u, v) \mapsto (x, y) \mapsto (x_1, y_1),$$

the transformation formulas give

$$(-1, 1) \mapsto (-2, 1) \mapsto (-5, 23) \mapsto (1, 5),$$

$$(-1, -1) \mapsto (-2, -1) \mapsto (-6, 24) \mapsto (0, 0),$$

and L'Hôpital yields

$$(1, -1) \mapsto (0, -1) \mapsto (18, -240) \mapsto (24, -120).$$

The inverse transformations yield, *e.g.,*

$$(x_1, y_1) = (1, -5) \mapsto (u_1, v) = \left(-\frac{33}{13}, -\frac{1871}{169}\right).$$

J. Fearnley raised the question: starting with different rational points on the same quartic, how are the Weierstrass equations given by the proposition related? We will see in a later chapter that the Riemann-Roch theorem implies that one can pass from one equation to any other one by a transformation of the form $x = \alpha^2 x_1 + r,\ y = \alpha^3 y_1 + s\alpha^2 x_1 + t$, where $\alpha, r, s, t \in K$, $\alpha \neq 0$. In the language of §4.1, the elliptic curves are isomorphic.

The above proposition can be 'reverse engineered': given a point $Q = (x_0, y_0)$ satisfying a Weierstrass equation E, one can write down an equation (¶): $v^2 = au^4 + \cdots$ as in the proposition, and birational transformations between E and (¶), such that Q corresponds to $(0, -q)$. The first step is to transform the equation of E to a new Weierstrass equation E' whose coefficients satisfy $a_6' = a_2' a_4'$ and such that Q is transformed to $(-a_2', a_1' a_2' - a_3')$ as in the proposition. For reference purposes we put the details in a

Corollary 1.2.2 *Let K be a field of characteristic $\neq 2$, let E be a Weierstrass equation with coefficients $a_1 \ldots, a_6 \in K$, and let $Q = (x_0, y_0) \in E(K)$.*

(a) *Define*

$$x' = x + (x_0 + a_2)/2, \quad y' = y + (y_0 + a_1 x_0 + a_3).$$

Then x', y' satisfy the Weierstrass equation with coefficients

$$\begin{aligned}
a_1' &= a_1, \\
a_2' &= -(3x_0 + a_2)/2 = -x_0', \\
a_3' &= -(2y_0 + a_1(5x_0 + a_2)/2 + a_3) = -y_0' + a_1' a_2', \\
a_4' &= a_1 y_0 + (a_1^2 + a_2/2)x_0 + 3x_0^2/4 + a_1 a_3 - a_2^2/4 + a_4, \\
a_6' &= a_2' a_4'.
\end{aligned}$$

In terms of the new x', y'-coordinates,

$$Q = (x_0', y_0') = (-a_2', a_1' a_2' - a_3').$$

(b) *Define*

$$\begin{aligned}
u &= \frac{2(x - x_0)}{y + y_0 + a_1 x_0 + a_3}, \\
v &= \frac{2x + x_0 + a_2}{4} u^2 - \frac{a_1}{2} u - 1.
\end{aligned}$$

Then

$$v^2 = au^4 + bu^3 + cu^2 + du + 1 \qquad (\P')$$

where

$$a = -a_4'/4, \quad b = a_3'/2, \quad c = {a_1'}^2/4 + a_2', \quad d = a_1'.$$

The inverse transformations are

$$\begin{aligned}
x &= (2(v + 1) + du)/u^2 - (x_0 + a_2)/2, \\
y &= (4(v + 1) + 2(du + cu^2) - d^2 u^2/2)/u^3 - (y_0 + a_1 x_0 + a_3).
\end{aligned}$$

In this birational correspondence, Q corresponds to the point $(u, v) = (0, -1)$ on $(\P')$.

Proof. The verification of (a) amounts to some easy calculations, and (b) to applying the formulas in the proposition where we have chosen $q = 1$. (There is no real loss of generality in the proposition if we take $q = 1$ — this corresponds to replacing v with qv and dividing ($\P$) by q^2.) ■

We mention two points concerning the calculation of E':

1. If E satisfies $a_6 = a_2 a_4$ it is still usually necessary to make the transformation to E' in order to have $Q = (-a_2', a_1' a_2' - a_3')$.

2. Another application of the transformation produces no change: $x'' = x'$ and $y'' = y'$, hence $a''_i = a'_i$.

Because the reciprocal quartic $a + b/u + c/u^2 + d/u^3 + 1/u^4$ will arise on several occasions, it is worthwhile to introduce special notation. It turns out to be convenient to substitute $1/u = m/2 - a_1/4$, which produces a quartic polynomial $\cdots + m^4/16$. Multiplying this by 16 and using the notation $\eta_0 = y_0 + (a_1x_0 + a_3)/2$, the resulting quartic is

$$\boxed{\mathrm{Quar}_Q(m) = (b_2^2/16 - 2b_4 - b_2x_0/2 - 3x_0^2) - 8\eta_0 m - (b_2/2 + 6x_0)m^2 + m^4.}$$

Combining the relation $1/u = m/2 - a_1/4$ with those connecting u, v with x, y, we have

Corollary 1.2.3 *With K and E as in the previous corollary, for each point $Q \in E(K)$ the quartic curve*

$$v^2 = \mathrm{Quar}_Q(m)$$

is birationally equivalent with E. ■

Here is a numerical example over $K = \mathbf{Q}$:

$$E : y^2 = x^3 - x^2 + x, \quad Q = (0,0), \tag{A24}$$

$$E' : y'^2 = x'^3 + \frac{1}{2}x'^2 + \frac{3}{4}x' + \frac{3}{8}, \quad (x'_0, y'_0) = (-1/2, 0),$$

$$\mathrm{Quar}_Q(m) = -3 + 2m^2 + m^4.$$

The significance of the fact that this polynomial has rational roots $m = \pm 1$ will be revealed in §1.7.1.

Here are three examples of E with $\Delta = 0$: $y^2 = x^3$, $y^2 = x^3 + x^2$, and $y^2 = x^3 - x^2$. For these three E, $\mathrm{Quar}_{(0,0)}(m)$ is, respectively,

$$m^4, \quad (m-1)^2(m+1)^2, \quad (m^2+1)^2.$$

We quote from [Ada-Ra80, p.483] a specialized form of the previous corollary that will be used later.

Corollary 1.2.4 *Let $P = (p, q)$ be a point on $E : y^2 = x^3 + bx + c$, all defined over the field K of characteristic $\neq 2$, and define*

$$u := \frac{y+q}{x-p}, \quad v := 2x + p - \left(\frac{y+q}{x-p}\right)^2.$$

Then

$$v^2 = u^4 - 6pu^2 - 8qu - (4b + 3p^2).$$

The inverse transfomation is

$$x = (u^2 + v - p)/2, \quad y = (u^3 + uv - 3pu - 2q)/2. \quad ■$$

The procedure for transforming the general cubic $s_1u^3 + s_2u^2v + \cdots + s_9v + s_{10} = 0$ to a Weierstrass equation involves projective coordinates and projective transformations and so will be given in §1.4 after these necessary preliminaries.

1.3 Projective coordinates.

When we call E a plane curve we are referring to the projective plane $\mathbf{P}^2$. Let us recall the definition of n-dimensional projective space $\mathbf{P}^n(K)$ over a field K. From **affine space** $\mathbf{A}^{n+1}(K)$, which consists of all $n+1$-tuples $(X_0, \ldots, X_n) \in K^{n+1}$, we remove the origin $(0, \ldots, 0)$ and divide by the equivalence relation given by the action of the multiplicative group K^*: $(X_0, \ldots, X_n)$ and $(Y_0, \ldots, Y_n)$ are equivalent if $\exists \lambda \in K^*$ such that $Y_i = \lambda X_i$ $\forall i$. (This relation is reflexive since $1 \in K^*$; symmetric since $\lambda \in K^* \Rightarrow \lambda^{-1} \in K^*$; and transitive since $\lambda, \mu \in K^* \Rightarrow \lambda\mu \in K^*$.) Thus $\mathbf{P}^2(K)$ consists of all triples (X, Y, Z) where not all of X, Y, Z are 0 and where we identify (X, Y, Z) with $(\lambda X, \lambda Y, \lambda Z)$ for $\lambda \in K^*$.

If K' is an overfield of K then there is a natural inclusion $\mathbf{P}^n(K) \subset \mathbf{P}^n(K')$; for if $(X_0, \ldots, X_n)$, $(Y_0, \ldots, Y_n)$ represent points in $\mathbf{P}^n(K)$ and $\lambda \in K'^*$ is such that $Y_i = \lambda X_i$, $\forall i$, then $\lambda \in K^*$ since at least one $X_i \neq 0$. On the other hand, if $P \in \mathbf{P}^n(K')$ is represented by $(X_0, \ldots, X_n) \in K'^{n+1}$, then $P \in \mathbf{P}^n(K)$ iff $\exists \lambda \in K'^*$ such that all $\lambda X_i \in K^*$, equivalently, if X_j is any nonzero coordinate, then $X_i/X_j \in K$ for all i. We then say that P is **defined over K**.

$\overline{K}$ always denotes an algebraic closure of K and we normally abbreviate $\mathbf{P}^n(\overline{K})$ to $\mathbf{P}^n$.

Recall that a homogeneous polynomial F of degree d in the $n+1$-variable polynomial ring $K[U_0, \ldots, U_n]$, *i.e.,* a nonzero linear combination of monomials $U_0^{d_0} \cdots U_n^{d_n}$ with $d_0 + \cdots + d_n = d$, has the property

$$F(\lambda U_0, \ldots, \lambda U_n) = \lambda^d F(U_0, \ldots, U_n) \quad \forall \lambda \in K.$$

In fact this can be taken as the definition when K is infinite; alternatively, if this relation is true for a nonzero polynomial F and a transcendental λ, then F is homogeneous of degree d. It follows that '$F(P) = 0$' is unambiguously true or false for a point $P \in \mathbf{P}^n(K)$. The **zero set of F over K** is

$$\mathsf{Z}_K(F) = \{P \in \mathbf{P}^n(K) : F(P) = 0\}.$$

$\mathsf{Z}(F)$ stands for $\mathsf{Z}_{\overline{K}}(F)$.

A **hyperplane** in $\mathbf{P}^n(K)$ is the zero set of a linear homogeneous equation $c_0X_0 + \cdots + c_nX_n = 0$ where the c_i are not all 0. A **linear subspace** of $\mathbf{P}^n(K)$ is an intersection of hyperplanes, in other words the set of points whose coordinates satisfy a system of linear homogeneous equations. The usual elimination procedure of linear algebra removes redundant equations so that one has a system of r equations where r is the rank of the coefficient matrix. The **dimension**

of the linear subspace is defined to be $n-r$. Thus the dimension of $\mathbf{P}^n(K)$ is n. **Lines** and **planes** are linear subspaces of dimension 1 and 2 respectively; in $\mathbf{P}^2(K)$, hyperplanes are lines.

We know from linear algebra that the rank of a matrix does not change when we view it as being defined over a larger field. Thus the dimension of a linear subspace determined by a set of equations defined over K does not change when K is replaced by an overfield K'.

Linear coordinate changes are given by invertible $(n+1)\times(n+1)$ matrices $A=(a_{ij})$:

$$X_i' = a_{i0}X_0 + \cdots + a_{in}X_n.$$

We indicate this with the matrix notation $AX = X'$, where X is the column vector with entries $X_0,\ldots,X_n$, and similarly for X'. Since λA for $\lambda \in K^*$ gives the same transformation, one is led to the **projective general linear group**:

$$PGL_n(K) = GL_{n+1}(K)/\langle K^* I\rangle,$$

the quotient of the general linear group of invertible $(n+1)\times(n+1)$ matrices by the normal subgroup of nonzero scalar matrices.

Clearly the property of being a linear subspace of dimension $n-r$ is preserved under a linear change of coordinates; in particular, lines remain lines. Also the set of homogeneous polynomials of degree d is mapped to itself.

For later use we make the simple observation that for any given point a coordinate system can be chosen so that the point has coordinates $(0,\ldots,0,1)$, for example. More generally,

Proposition 1.3.1 *Let $P_i = (a_{0i},\ldots,a_{ni})$, $i=1,\ldots,n+1$ be points in $\mathbf{P}^n(K)$ not contained in any hyperplane*, i.e., *the $(n+1)\times(n+1)$ matrix M whose i-th column is $(a_{0i},\ldots,a_{ni})$ is invertible. Then under the linear change of coordinates $M^{-1}X = X'$, the new coordinates of $P_1,\ldots,P_n$ are*

$$(1,0,\ldots,0),\ (0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)$$

respectively. ■

We choose the line $Z=0$ as the **line at infinity** in $\mathbf{P}^2(K)$. This choice is arbitrary; unlike affine space which has the origin as a distinguished point,[†] the projective plane has no distinguished point or line. But having made this choice the points are of two types:

(i) the "affine" points with $Z \neq 0$: $(X,Y,Z) = (x,y,1)$ where $x = X/Z$, $y = Y/Z$

(ii) the points "at infinity" with $Z=0$: $(X,Y,0)$.

[†]Here affine space is regarded as a vector space; however when regarded as an algebraic variety there is no distinguished point.

A "compact" visualization of $\mathbf{P}^2(\mathbf{R})$ is the closed disc with antipodal (diametrically opposite) points identified.

This picture can be obtained by projecting from the center of a hemisphere to the affine plane:

This sets up a bijection between the points on the affine plane and the interior points of the disc (the hemisphere flattened out). The points $(X, Y, 0)$ at ∞ are in bijection with the lines through the origin in the affine plane: the line through $(0, 0)$ and (X, Y) is the same as that through $(0, 0)$ and $(\lambda X, \lambda Y)$. And these lines are in bijection with pairs of antipodal points on the circle bounding the disc.

If we rewrite the Weierstrass equation

$$f = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 = 0$$

in projective coordinates by substituting $x = X/Z$, $y = Y/Z$ and multiplying by Z^3 we get

$$F = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0$$

which is a homogeneous polynomial of degree 3. At infinity this reduces to $-X^3 = 0$, hence $X = 0$ and E has the unique point at infinity

$$(0, Y, 0) = (0, 1, 0).$$

This serves as the designated rational point O of E. How the general curve of genus 1 with a rational point O is converted into Weierstrass form will be explained when we discuss the Riemann-Roch theorem; indeed that theorem will be needed to define the genus of a curve.

A basic topic in the algebraic geometry of $\mathbf{P}^2$ is the analysis of the points of intersection of two curves. The general discussion is quite involved and for now we give only the simplest results.

Proposition 1.3.2 *Let K be any field.*

(a) *Let F be a nonzero homogeneous polynomial of degree d in the two variables U_0 and U_1 defined over K, say*

$$F = \prod_{i=1}^{d} (\alpha_i U_0 + \beta_i U_1)$$

for appropriate α_i, $\beta_i \in \overline{K}$. Then in $\mathbf{P}^1$, $\mathsf{Z}(F)$ consists of d points $(\beta_i, -\alpha_i)$, possibly some coincident.

(b) *If L and C are nonzero homogeneous polynomials of degrees 1 and $d \geq 1$ respectively in three variables defined over K, then in $\mathbf{P}^2$, either $\mathsf{Z}(L) \subset \mathsf{Z}(C)$, or $\mathsf{Z}(L) \cap \mathsf{Z}(C)$ consists of d points, possibly some coincident.*

(c) *If C_1 and C_2 are nonconstant homogeneous polynomials in $K[U_0, U_1, U_2]$, then in $\mathbf{P}^2$, the set $\mathsf{Z}(C_1) \cap \mathsf{Z}(C_2)$ is nonempty. This set is finite iff C_1 and C_2 have no common factor, and in that case all points in $\mathsf{Z}_{K'}(C_1) \cap \mathsf{Z}_{K'}(C_2)$ for any field $K' \supset \overline{K}$ are defined over $\overline{K}$.*

Remarks. Anticipating definitions to be made in Chapter 6, a *plane curve* over K is effectively a homogeneous polynomial F in the variables U_0, U_1, U_2, and the *degree* of the curve is the degree of F. Thus two lines in $\mathbf{P}^2$, if not coincident, intersect in a unique point; and a line intersects a curve of degree d either in precisely d points (properly counted), or else is entirely contained in that curve, in which case the line is a *component* of the curve.

Statement (c) can be formulated as: two curves in $\mathbf{P}^2$ intersect in at least one point; the intersection is finite iff the curves have no component in common, and then all intersection points are algebraic over K. Bézout's theorem says that two curves with degrees d_1 and d_2 and without common components intersect in exactly $d_1 d_2$ points — properly counted. However the precise statement requires a number of preliminaries, including a discussion of singular points.

Proof. (a) is obvious.

(b) Let $L = c_0 U_0 + c_1 U_1 + c_2 U_2$ where, say, $c_2 \neq 0$. Substituting $U_2 = (-c_0/c_2)U_0 + (-c_1/c_2)U_1$ into G yields a homogeneous polynomial in U_0, U_1 which is either 0 or nonzero of degree d. The statement now follows from part (a).

(c) Let C_i be homogeneous in U_1, U_2, U_3 of degree $d_i > 0$. By Proposition 1.3.1, choose a coordinate system so that $(0, 0, 1)$ is on neither C_i. Then

$$C_i = c_{i0} + c_{i1} U_2 + \cdots + c_{id_i} U_2^{d_i}$$

where both c_{id_i} are nonzero constants and c_{ij}, if not 0, is homogeneous in U_0, U_1 of degree $d_i - j$.

As polynomials in the variable U_2 over the ring $K[U_0, U_1]$, their resultant R is a polynomial in $K[U_0, U_1]$, and there exist λ, $\mu \in K[U_0, U_1, U_2]$ such that

$$\lambda C_1 + \mu C_2 = R. \tag{¶}$$

In fact λ and μ are homogeneous of degrees $d_1(d_2-1)$ and $d_2(d_1-1)$ respectively, and therefore R, if not 0, is homogeneous in U_0, U_1 of degree $d_1 d_2$. All of this follows from a formula that we quote from [Con82,p.213]:

$$R = \begin{vmatrix} c_{1d_1} & \cdots & c_{10} & & & & U_2^{d_2-1}C_1 \\ & c_{1d_1} & \cdots & c_{10} & & & U_2^{d_2-2}C_1 \\ & \ddots & & & \ddots & & \vdots \\ & & c_{1d_1} & \cdots & c_{10} & & U_2 C_1 \\ & & & c_{1d_1} & \cdots & c_{11} & C_1 \\ c_{2d_2} & \cdots & c_{20} & & & & U_2^{d_1-1}C_2 \\ & \ddots & & & \ddots & & \vdots \\ & & c_{2d_2} & \cdots & c_{20} & & U_2 C_2 \\ & & & c_{2d_2} & \cdots & c_{21} & C_2 \end{vmatrix}$$

where entries in a row outside the subscript limits of c_{ij} are 0. Expansion of this determinant along the right column simultaneously gives λ, μ and R.

$R = 0$ iff the C_i have a common factor F which is a nonconstant polynomial in U_2 over the field $K(U_0, U_1)$; in fact, since factors of homogeneous polynomials are again homogeneous, F is a homogeneous polynomial of positive degree in the three variables. Then the two curves share the component $\mathsf{Z}(F)$.

If $R \neq 0$, let $\alpha U_0 + \beta U_1$ be a factor of R as in (a) where, say, $\alpha \neq 0$. Let $\widetilde{C_1}$ denote the image of C_1 under the substitution $U_0 \mapsto (-\beta/\alpha)U_1$, and similarly for $\widetilde{C_2}$ and $\widetilde{R}$. Since the leading coefficient of C_i is a constant, the degree of $\widetilde{C_i}$ in U_2 is still d_i, and therefore $\widetilde{R}$ is the resultant of $\widetilde{C_1}, \widetilde{C_2}$ as polynomials in U_2 over $K(U_1)$. The fact that $\widetilde{R} = 0$ means that these polynomials have a factor $\gamma U_1 + \delta U_2$ in common. Hence the point $(\beta\delta, -\alpha\delta, \alpha\gamma)$ lies on the intersection of the two curves.

Suppose $R \neq 0$ and $P = (a_0, a_1, a_2)$ lies on both curves. Since $c_{id_i} \neq 0$, therefore a_0 and a_1 are not both 0. Under the substitutions $U_i \mapsto a_i$, R becomes $\widetilde{R} = 0$ by (¶), hence $a_1 U_0 - a_0 U_1$ is a factor of R. Multiplying the coordinates of P by an appropriate λ, we can assume that $a_0, a_1 \in \overline{K}$, and then from either of the equations $\widetilde{C_i} = 0$ we conclude that $a_2 \in \overline{K}$ also. ■

1.4 Cubic to Weierstrass: Nagell's algorithm

Let K be a field of characteristic $\neq 2$ or 3 and consider the curve defined by an equation over K of the form $0 =$ a cubic in u and v with a rational point (p, q).

This time we can translate both variables. Replacing u by $u+p$ and v by $v+q$ we can assume that the rational point is $(0,0)$:

$$s_1u^3 + s_2u^2v + s_3uv^2 + s_4v^3 + s_5u^2 + s_6uv + s_7v^2 + s_8u + s_9v = 0. \qquad (\P)$$

Let f denote the polynomial on the left of (¶).

We now describe the algorithm, due to Nagell [Nag28], to transform f into Weierstrass form, or to discover that the curve is not elliptic.

Step 1. Interchange u and v if necessary to ensure $s_9 \neq 0$. (If both s_8 and s_9 are 0 then $(0,0)$ is a singular point (see §1.5) and the curve is not elliptic.)

Step 2. Substitute $u = U/W$, $v = V/W$ and clear denominators to obtain the homogenized form

$$F = F_3 + F_2W + F_1W^2 = 0$$

where

$$\begin{aligned} F_3 &= s_1U^3 + s_2U^2V + s_3UV^2 + s_4V^3, \\ F_2 &= s_5U^2 + s_6UV + s_7V^2, \\ F_1 &= s_8U + s_9V. \end{aligned}$$

The rational point P with (u,v)-coordinates $(0,0)$ has projective coordinates $(U,V,W) = (0,0,1)$. The tangent line at P, given by $F_1 = 0$, meets the curve in the point $Q = (-e_2s_9, e_2s_8, e_3)$ where $e_i = F_i(s_9, -s_8)$, $i = 2,3$. The e_i cannot both be 0 because that would make the tangent a component and the curve would be reducible — not elliptic; $e_2 = 0$ means that $P = Q$ is a flex (the tangent has triple contact with the curve at P), while $e_3 = 0$ means that Q is at infinity. If $e_3 \neq 0$ make the coordinate change $U = U' - (s_9e_2/e_3)W'$, $V = V' + (s_8e_2/e_3)W'$, $W = W'$, while if $e_3 = 0$ make the change $U = U' - s_9W'$, $V = V' + s_8W'$, $W = U'$. In either case Q is now at the origin $(U',V',W') = (0,0,1)$ and the tangent at P is $s_8U' + s_9V' = 0$. We can now return to affine coordinates $u' = U'/W', v' = V'/W'$; projective coordinates were really only needed to deal with the case when Q was at infinity.

Step 3. If the equation in terms of u', v' is $f' = f'_3 + f'_2 + f'_1 = 0$ where $f'_i = f'_i(u',v')$ denotes the homogeneous part of f' of degree i, then

$$u'^2 f'_3(1,t) + u'f'_2(1,t) + f'_1(1,t) = 0$$

where $t = v'/u'$. Thus

$$u' = \frac{-\phi_2 \pm \sqrt{\delta}}{2\phi_3}, \qquad v' = tu', \qquad (*)$$

where $\phi_i = f'_i(1,t)$ and $\delta = \phi_2^2 - 4\phi_1\phi_3$. The values of t such that $\delta = 0$ are the slopes of the tangents to the curve that pass through Q, and one of these values is $t_0 = -s_8/s_9$. Write $t = t_0 + 1/\tau$ so that $\rho = \tau^4\delta$ is a cubic polynomial in τ.

Step 4. Finally, if

$$\rho = c\tau^3 + d\tau^2 + e\tau + k$$

then $c \neq 0$ (since $c = 0$ implies that the original curve is not elliptic) and the substitutions $\tau = x/c$, $\rho = y^2/c^2$ give the Weierstrass equation

$$y^2 = x^3 + dx^2 + cex + c^2k.$$

The relations between the original variables u, v and x, y can be traced back starting with $(*)$ where

$$t = t_0 + c/x, \qquad \delta = c^2y^2/x^4.$$

1.4.1 Example 1: Selmer curves

By a **Selmer curve** we understand a homogeneous cubic equation of the form

$$aU^3 + bV^3 + cW^3 = 0 \quad \text{where} \quad abc \neq 0,$$

or an affine version such as

$$au^3 + bv^3 = c.$$

The coefficients appear symmetrically: in the homogeneous case we can permute the variables to obtain a permutation of a, b, c; in the affine case, to interchange a and c for instance, we can substitute $1/u$, $-v/u$ for u, v.

Let us apply Nagell's algorithm:

Proposition 1.4.1 *Let the Selmer curve*

$$au^3 + bv^3 = c, \quad \textit{where} \quad abc \neq 0,$$

be defined over a field K of characteristic $\neq 2$ or 3, and (permuting a, b, c if necessary) assume that $\theta := \sqrt[3]{c/b} \in K$. Then the Selmer curve is birationally equivalent to the Weierstrass curve

$$y^2 = x^3 - 432a^2b^2c^2$$

under the mutually inverse transformations

$$u = -\frac{6b\theta^2x}{y - 36abc}, \quad v = \frac{y + 36abc}{y - 36abc}\theta,$$

$$x = -\frac{12ab\theta^2u}{v - \theta}, \quad y = 36abc\frac{v + \theta}{v - \theta}.$$

Remark. Replacing u, v with $\theta b/u$, $-\theta v/u$ transforms the Selmer curve to $u^3+v^3=ab^2$, which is dealt with in the first corollary below. Thus the proposition is not really more general, but it is convenient to have the details displayed for the symmetrical *abc*-equation; a similar remark applies to the second corollary.
Proof. Replacing v with $v+\theta$ yields a cubic of the form (¶) of the previous section with

$$s_1 = a, \quad s_4 = b, \quad s_7 = 3b\theta, \quad s_9 = 3b\theta^2,$$

and the remaining $s_i = 0$. We find $e_2 = 0$, $e_3 = 27abc^2$. Hence no transformation is needed in step 2 and

$$\phi_3 = a + bt^3, \quad \phi_2 = 3b\theta t^2, \quad \phi_1 = 3b\theta^2 t,$$

$$\delta = -3b\theta^2 t(4a + bt^3).$$

$$t_0 = 0, \quad t = 1/\tau, \quad \rho = -12ab\theta^2\tau^3 - 3b\theta^2.$$

Hence the Weierstrass equation is

$$y^2 = x^3 - 432a^2b^2c^2$$

where x, y are as stated in the proposition. ■

We single out a particular example that will be referred to later:

Corollary 1.4.2 *Let K be a field of characteristic $\neq 2$ or 3, and let $a \in K^*$. Then the twisted Fermat curve*

$$u^3 + v^3 = a$$

is birationally equivalent to the Weierstrass curve

$$y^2 = x^3 - 432a^2$$

under the mutually inverse transformations

$$u = \frac{36a - y}{6x}, \qquad v = \frac{36a + y}{6x},$$

$$x = \frac{12a}{u+v}, \qquad y = 36a\frac{v-u}{v+u}.$$

Proof. We substitute $u = 1/u_1$ and $v = -v_1/u_1$, apply the proposition with $b = c = \theta = 1$, then translate the formulas back using $u_1 = 1/u$, $v_1 = -v/u$. ■

For example, if $a = \alpha^3 + 1$ then

$$(u, v) = (1, \alpha) \longmapsto (x, y) = (12(\alpha^2 - \alpha + 1), 36(\alpha^2 - \alpha + 1)(\alpha - 1)).$$

The proposition can be restated in terms of projective coordinates as follows, where c is replaced by $-c$:

Corollary 1.4.3 *Let C denote the Selmer curve $aU^3 + bV^3 + cW^3 = 0$ defined over the field K of characteristic $\neq 2$ or 3; assume $abc \neq 0$ and $\theta := -\sqrt[3]{c/b} \in K$; let E denote the homogeneous form of the Weierstrass equation: $Y^2Z = X^3 - 432a^2b^2c^2Z^3$; let $C(K)$ and $E(K)$ denote the set of points in $\mathbf{P}^2(K)$ on C and E respectively. Then mutually inverse bijections $C(K) \longleftrightarrow E(K)$ are defined by*

$$(U, V, W) \mapsto \left(-12ab\theta^2 U, -36abc(V + \theta W), V - \theta W\right),$$

$$(X, Y, Z) \mapsto \left(-6b\theta^2 X, (Y - 36abcZ)\theta, Y + 36abcZ\right)$$

in which $O \in E(K)$ corresponds to $(0, \theta, 1) \in C(K)$. ■

Thus Fermat's last theorem for exponent 3, *i.e.,* Euler's result that $U^3 + V^3 + W^3 = 0$ has only the three solutions in $\mathbf{P}^2(\mathbf{Q})$ in which one of U, V, W is 0, is equivalent to $|E(\mathbf{Q})| = 3$ where E (in affine form) is $y^2 = x^3 - 432$. This will come out as an example of '2-descent' in Corollary 3.7.5.

Selmer curves will serve as important examples of various topics later in these notes. For example, $aU^3 + bV^3 + cW^3 = 0$ will be seen to be a "torsor" of $U^3 + V^3 + abcW^3 = 0$. The latter curve has the rational point $(U, V, W) = (1, -1, 0)$, and so is an elliptic curve in the sense of the second definition of §1.1, and in fact is the Jacobian of the former curve, as will be explained later. For now we mention[‡]

Proposition 1.4.4 *If*

$$au^3 + bv^3 + cw^3 = 0$$

then

$$r^3 + s^3 + abct^3 = 0$$

where

$$\begin{aligned} r &= -6\,bc^2v^3w^6 - c^3w^9 - 3\,b^2cv^6w^3 + b^3v^9, \\ s &= -3\,bc^2v^3w^6 + c^3w^9 - 6\,b^2cv^6w^3 - b^3v^9, \\ t &= -3\,uvw\left(b^2v^6 + bcv^3w^3 + c^2w^6\right). \end{aligned}$$

If $3abcuvw \neq 0$ (the only case of interest) and abc is not a cube, then $t \neq 0$; thus, by the previous corollary, the elliptic curve

$$y^2 = x^3 - 432a^2b^2c^2$$

has the non-O point

$$x = \frac{4(b^2v^6 + bcv^3w^3 + c^2w^6)}{u^2v^2w^2},$$

[‡]See also Proposition 1.4.6 and its corollary in the next section which apply in particular to Selmer curves.

$$y = \frac{4(2b^3v^9 + 3b^2cv^6w^3 - 3bc^2v^3w^6 - 2c^3w^9)}{u^3v^3w^3}.$$

The statement $t \neq 0$ is a consequence of the implication $au^3 + bv^3 + cw^3 = 0$ and $b^2v^6 + bcv^3w^3 + c^2w^6 = 0 \Longrightarrow$

$$a^2u^6 = (-au^3)^2 = (bv^3 + cw^3)^2 = bcv^3w^3.$$

The verification of the equation $r^3 + s^3 + abct^3 = 0$ is a simple computer exercise. However we should indicate how the formulas for r, s, t were obtained; here we are guided by [Cas91, p.86].[§]

To obtain these formulas we work in a field of characteristic $\neq 3$ containing the quantities $a, \ldots, w$ and also a primitive cube root of unity ρ. Let

$$\lambda = au^3 + \rho bv^3 + \rho^2 cw^3,$$

$$\mu = au^3 + \rho^2 bv^3 + \rho cw^3,$$

so that

$$\begin{aligned} \lambda^3 + \mu^3 &= (\lambda + \mu)(\rho^2\lambda + \rho\mu)(\rho\lambda + \rho^2\mu) \\ &= (3au^3)(3bv^3)(3cw^3). \end{aligned}$$

Hence the points $P = (\lambda, \rho\mu, \nu)$ and $P' = (\mu, \rho^2\lambda, \nu)$, where $\nu = -3uvw$, lie on the curve $R^3 + S^3 + abcT^3 = 0$. By Proposition 1.3.2(b), the line joining P and P' meets this curve in a third point Q, and we expect that point to be "rational", *i.e.,* not involving ρ (because if σ denotes the automorphism sending $\rho \mapsto \rho^2$ and leaving $a, \ldots, w$ fixed — we can take the latter as transcendentals subject only to the relation $au^3 + bv^3 + cw^3 = 0$ — then P and $\sigma P = P'$ are conjugate).

Calculation shows that the third point $Q = (r, s, t)$ is given by the formulas in the proposition. Starting with other $P = (\lambda, \mu, \nu)$, $(\rho\lambda, \rho\mu, \nu)$, $\ldots$, and corresponding $P' = \sigma P = (\mu, \lambda, \nu)$, $(\rho^2\mu, \rho^2\lambda, \nu)$, $\ldots$, does not lead to anything essentially new, only to one of Q, (s, r, t), $(1, -1, 0)$.

A famous example of Selmer is that

$$3U^3 + 4V^3 + 5W^3 = 0$$

has no points in $\mathbf{P}^2(\mathbf{Q})$, in other words, the equation has no solution in rational numbers other than $(0, 0, 0)$. For if there were a solution then, by the proposition, the elliptic curve

$$y^2 = x^3 - 432 \cdot 60^2$$

would have a point defined over $\mathbf{Q}$ distinct from O, which is not the case. But the proof of the last statement must wait until Corollary 3.7.8.

[§]A more natural, but more complicated way of obtaining the formulas will be explained in §1.7.2 (using 'multiplication by 3').

1.4.2 Example 2: Desboves curves

By a **Desboves curve** we understand a homogeneous cubic equation of the form

$$aU^3 + bV^3 + cW^3 + dUVW = 0,$$

or an affine version of such an equation. We chose this name for this class of curves because of the historical reference [Des86] brought to our attention in [Cas91, p.130]; references to related work by Cauchy and others are given in [Dic52, vol.2, ch.XXI]. Selmer curves are included as the particular case $d = 0$.

Proposition 1.4.5 *Let the Desboves curve*

$$au^3 + bv^3 + c + duv = 0$$

be defined over the field K of characteristic $\neq 2, 3$, and assume (permuting a, b, c if necessary) that

$$abc\lambda \neq 0 \quad \textit{where} \quad \lambda := 27abc + d^3, \quad \textit{and} \quad \theta := -\sqrt[3]{c/b} \in K.$$

Then, by Nagell's algorithm, this curve is birationally equivalent to

$$y^2 = x^3 - 3d^2x^2 + \frac{8}{3}d\lambda x - \frac{16}{27}\lambda^2 = x^3 - 3\left(dx - \frac{4}{9}\lambda\right)^2.$$

Remark. The transformation equations between u, v and x, y are somewhat lengthy and for that reason are not included in the statement of the proposition.
Proof. The proof proceeds as in the case of Selmer curves, except that now $s_6 = d$ and $s_8 = d\theta$; $e_2 = 0$ again, so no transformation is needed in step 2, and $e_3 = c\lambda$. The rest is calculation. ■

We quote Desboves' formulas. Once again the verification is a computer exercise and, as in the special case of Selmer curves, the underlying idea is that in $\mathbf{P}^2$ a line meets a Desboves curve in three points, provided these points are counted with the appropriate multiplicities: this includes the case of a line tangent to the curve when two of the points are coincident.

Proposition 1.4.6 *Let $P = (x_1, x_2, x_3)$ be a point on the Desboves curve*

$$a_1X_1^3 + a_2X_2^3 + a_3X_3^3 + dX_1X_2X_3 = 0 \tag{D}$$

defined over a field of characteristic $\neq 3$. Then the third point of intersection (t_1, t_2, t_3) of the tangent line at P has coordinates

$$t_j = x_j(a_{j+1}x_{j+1}^3 - a_{j+2}x_{j+2}^3) \quad (\textit{subscripts taken} \bmod 3).$$

If $Q = (y_1, y_2, y_3)$ is another point on the curve then the third point of intersection (z_1, z_2, z_3) of the line joining P and Q has coordinates (again subscripts are taken mod 3*)*

$$z_j = x_j^2y_{j+1}y_{j+2} - y_j^2x_{j+1}x_{j+2}. \quad ■$$

The following corollary is due to Hurwitz [Hur17].

Corollary 1.4.7 *Let S be the set of points in $\mathbf{P}^2(\mathbf{Q})$ on the Desboves curve* (D) *where a_1, a_2, a_3, d are integers and the a_j are positive, distinct and square-free. Then S is either empty or infinite. In fact, if $P_1 \in S$ then all the points in the sequence $P_1, P_2, \ldots$ are distinct, where P_{n+1} is the third point of intersection of the tangent at P_n.*

Remarks. There is no real loss of generality in assuming that the a_i are positive since X_i can be replaced by $-X_i$. (Equations where an $a_i = 0$ are trivially solved.) Hurwitz [Hur17, p.465] and Mordell [Mor69, p.78] make the additional, and apparently unnecessary, assumption that the a_i are coprime.

See Corollary 1.7.2 where the present corollary is re-interpreted.

Proof. The assumptions on the coefficients ensure that $\lambda = 27a_1a_2a_3 + d^3 \neq 0$.

Let $P_1 = (x_1, x_2, x_3)$ where $x_j \in \mathbf{Z}$ and $\gcd\{x_j\} = 1$, and let $P_2 = (t_1, t_2, t_3) = (t'_1, t'_2, t'_3)$ where the t_j are given by the formulas in the proposition, and $t'_j = t_j/k$ where $k = \gcd\{t_j\}$. Thus $\gcd\{t'_j\} = 1$. The result will follow from the strict inequality $|t'_1t'_2t'_3| > |x_1x_2x_3|$.

First we note that the x_j are co-prime; for if the prime p divides x_1 and x_2, say, then $p \nmid x_3$ and (D) implies $p^2|a_3$, contrary to the assumption that the a_j are square-free. Second, the x_j are nonzero; for if $x_1 = 0$, say, then x_2, x_3 being prime to x_1 are ± 1, and (D) implies $a_2 \pm a_3 = 0$, which contradicts the assumptions that a_2, a_3 are positive and distinct. Applying this result to P_2 shows that no $t_j = 0$.

Let us write the formulas as $t_j = x_ju_j$. We wish to prove that for all j, $k|u_j$, so that $t'_j = x_ju'_j$ where $u_j = u'_jk$. For then, since $\sum u_j = 0$, therefore $\sum_{j=1}^3 u'_j = 0$, hence not all u'_j can be ± 1, *i.e.,* at least one $|u'_j| > 1$, which gives the result.

Suppose, then, $k \nmid u_1$. This means that for some prime p, if $v(n)$ denotes the exponent of p in the unique factorization of a nonzero integer n, we have

$$\alpha := v(k) > v(u_1). \tag{1}$$

Since $k\,|t_1 = x_1u_1$, therefore $v(x_1) > 0$ and $v(x_2) = v(x_3) = 0$. It follows that $v(t_2) = v(u_2) = v(a_3x_3^3 - a_1x_1^3) \geq \alpha$. Since a_3 is square-free, this implies

$$v(a_3) = 1 \quad \text{hence} \quad \alpha = 1. \tag{2}$$

Similarly $v(a_2) = 1$ and therefore $v(u_1) = v(a_2x_2^3 - a_3x_3^3) > 0$. Thus (1) and (2) are in conflict. ■

As an exercise, Silverman proposes ([Sil86, p.43]) the determination of those $a_1, \ldots, d$ for which S is not empty. The double asterisk on the exercise means, in this case, that it is a highly unsolved problem!

1.4.3 Example 3: Intersection of quadric surfaces

A **conic** or **conic section** in $\mathbf{P}^2$ is the set of points satisfying an equation $Q = 0$ where Q is a homogeneous quadratic polynomial in the three coordinates. The analogous definition in three dimensions is: a **quadric surface** in $\mathbf{P}^3$ is the set of points satisfying an equation $Q = 0$ where Q is a homogeneous quadratic polynomial in the four coordinates. In this section we assume that the characteristic is different from 2 and 3; the coordinates of a point in $\mathbf{P}^3$ will be denoted (U, V, W, X).

In general, the intersection of two quadric surfaces in $\mathbf{P}^3(K)$ is an elliptic curve, provided the intersection has at least one rational point. There are exceptions of course; for example the intersection of two spheres is a circle. Apart from the exceptions, the intersection can be transformed into a plane cubic with a rational point as we will explain, and then Nagell's algorithm can be applied.

However in certain cases an *ad hoc* approach that avoids Nagell's algorithm can be quicker and easier. Let us begin with such an example.†

Consider the intersection I of the two quadrics Q_1 and Q_2 given by the equations

$$Q_1 : \ U^2 - V^2 + kX^2 = 0, \quad Q_2 : \ W^2 - V^2 - kX^2 = 0,$$

where k is a nonzero parameter. Eliminating the kX^2 term we obtain

$$U^2 + W^2 = 2V^2$$

which can be interpreted as the equation of a conic C in the the plane $\mathbf{P}^2$ coordinatized by U, V, W. The curves C and I *cannot* be identified because for a given point (U, V, W) on the conic there are generally *two* values of X determined by $kX^2 = V^2 - U^2 = W^2 - V^2$. (One says that $(U, V, W, X) \mapsto (U, V, W)$ defines a covering of degree 2.)

The conic C contains the rational point $(U, V, W) = (1, 1, 1)$. Now, as a general remark, a conic with a rational point P can be (rationally) parametrized. The idea is simply this: because the equation of the conic is quadratic, a general line through P will intersect the conic in exactly one other point and that point will also be rational. (The other point will coincide with P in the special case when the line is tangent to the conic.) As a practical matter, one usually reverts to convenient affine coordinates.

In the present case it is natural to dehomogenize at $V = 1$: we define $u = U/V$ and $w = W/V$, so our conic is $u^2 + w^2 = 2$ with rational point $(u, w) = (1, 1)$. The general line through $(1, 1)$ is given by the equation $(u - 1) = t(w - 1)$ where t is a parameter. Substituting $u = 1 + t(w - 1)$ into the equation of the conic, we obtain a quadratic equation for w. One solution is, of course, $w = 1$;

†I am indebted to Peter Russell for help here, and in general for help with algebraic geometry in this section and elsewhere.

the other is

$$w = \frac{t^2 - 2t - 1}{t^2 + 1}, \qquad \text{hence} \quad u = \frac{-t^2 - 2t + 1}{t^2 + 1}.$$

Thus $(U, V, W) = (-t^2 - 2t + 1, t^2 + 1, t^2 - 2t - 1)$ is a parametrization of the points on the conic, and I is given by the equation

$$kX^2 = V^2 - U^2 = W^2 - V^2 = -4t^3 + 4t.$$

We can tidy this up by substituting $X = 2y/k^2$, $t = -x/k$:

$$E : y^2 = x^3 - k^2 x.$$

In terms of these new coordinates, this elliptic curve is the intersection of Q_1 and Q_2.

Exercise Using the transformations above, set up explicit mutually inverse bijections

$$I(K) \longleftrightarrow E(K).$$

Thus $I(K)$ becomes an elliptic curve by "transport of structure". You may find it more convenient to work with projective coordinates: the lines in $\mathbf{P}^2$ that pass through $(1, 1, 1)$ are $s(U - V) = t(W - V)$ where $(s, t) \in \mathbf{P}^1$ is a parameter; the second point of intersection with C is

$$(s^2 - 2st - t^2, s^2 + t^2, -s^2 - 2st + t^2).$$

Then E should be written in homogeneous form $y^2 z = x^3 - k^2 x z^2$. ■

Now let us consider the general case of the intersection of two quadrics. The ideas for this discussion are taken from Cassels [Cas91].

By a translation, we can suppose that the intersection I of the two quadrics Q_1 and Q_2 contains the point $P_0 = (0, 0, 0, 1)$. Then the equations for the quadrics can be written as

$$Q_1 : \ AX + B = 0, \quad Q_2 : \ CX + D = 0$$

where A, C are linear and B, D are quadratic in U, V, W. Eliminating X from the two equations produces

$$AD - BC = 0$$

which is a homogeneous cubic in U, V, W. Let I^* denote I with the point P_0 removed, and let E denote the curve in $\mathbf{P}^2$ defined by the above cubic. Then $(U, V, W, X) \mapsto (U, V, W)$ defines a map $f : I^* \longrightarrow E$.

Let us suppose first that A and C are linearly independent, that is, neither is a constant times the other. Then the two lines in the U, V, W plane described by $A = 0$ and $C = 0$ intersect in a unique point P_1, and this point lies on E. Let E^* denote E with the point P_1 removed. For each point (U, V, W) on E^*,

the equation for either Q_i uniquely determines a value for X, hence a point $f'(U,V,W) = (U,V,W,X)$ on I^*. The map $f' : E^* \longrightarrow I^*$ is inverse to f. (By extending the definitions by $f(P_0) = P_1$ and $f'(P_1) = P_0$, it follows that f (and f') are coverings of degree 1: the curves I and E are identical as abstract algebraic varieties.) E is thus a plane cubic with a rational point P_1 and Nagell can be applied; of course it may still turn out during the algorithm that E is not elliptic.

In the case that A and C are linearly dependent, say $C = cA$, by subtracting c times the equation for Q_1 from that of Q_2, we can suppose that $C = 0$. Then the equations defining I are $AX + B = 0$ and $D = 0$, hence we can suppose that $A \neq 0$ (otherwise I is a union of lines). The equation $AD = 0$ shows that E is a reducible curve: it contains the line $A = 0$ as a component. (Similarly if B and D are linearly dependent.) Also $X = -B/A$, $D = 0$ displays I as a degree 1 cover of the genus 0 curve defined by $D = 0$, hence I is a curve of genus 0 — not an elliptic curve. (The algebraic geometry background needed to flesh out these statements will be given later.)

Example The sphere $U^2 + V^2 + W^2 = 3X^2$ and the ellipsoid $(U - X)^2 + 2V^2 + 3W^2 = 5X^2$ share the point $P = (1,1,1,1)$. The transformation $U = U' + X'$, $V = V' + X'$, $W = W' + X'$, $X = X'$ gives P the coordinates $(0,0,0,1)$. Taking the point $(0,0,1)$ on the cubic (we are not obliged to take $P_1 = (1,-3,2)$ given by $A = C = 0$ — as for the quartic equations in Proposition 1.2.1, we will explain later that starting with different rational points in Nagell's algorithm yields isomorphic Weierstrass equations), Nagell's algorithm yields — we omit the details —

$$E: \ y^2 = x^3 + 44x^2 + 528x.$$

The reader may also wish to verify that the points $(0,0)$ and $(12,120)$ on E correspond to the points $(1,-1,1,1)$ and $(131,-259,-59,171)$ on the intersection.

1.5 Singular points.

Consider a homogeneous polynomial $F = F(X_0,\ldots,X_n) \in K[X_0,\ldots,X_n]$ of degree d. The Taylor expansion can be written as

$$F(X_0 + \lambda_0,\ldots,X_n + \lambda_n) = F_0 + F_1 + \cdots$$

where $F_i = F_i(\lambda_0,\ldots,\lambda_n)$ is homogeneous of degree i in the λ's, each coefficient of which is homogeneous of degree $d - i$ in the X's. Thus $F_0 = F(X_0,\ldots,X_n)$ and

$$F_1 = \sum_{i=0}^{n} a_i\lambda_i, \quad \text{where} \quad a_i = \frac{\partial F_0}{\partial X_i}\cdot$$

There is no problem with 'factorials in the denominators' since the Taylor expansion is the polynomial over K obtained by substituting $X_i + \lambda_i$ for X_i in F. However if $\operatorname{char} K \neq 2$ then one can write as in the classical Taylor expansion

$$F_2 = \frac{1}{2!} \sum_{i,j=1}^{n} a_{ij}\lambda_i\lambda_j, \quad \text{where} \quad a_{ij} = \frac{\partial^2 F_0}{\partial X_i \partial X_j},$$

and analogously for higher F_i.

Recall

Euler's Theorem: *For* $i = 0, 1, \ldots$

$$F_i(X_0, \ldots, X_n) = \binom{d}{i} F(X_0, \ldots, X_n).$$

Remark. If we add up these equations we obtain the identity

$$F(2X_0, \ldots, 2X_n) = 2^d F(X_0, \ldots, X_n) = \left(\sum_i \binom{d}{i} \right) F(X_0, \ldots, X_n).$$

Usually the theorem is stated in the form: for $i = 1, 2, \ldots$

$$\sum X_{s_1} \cdots X_{s_i} \frac{\partial^k F}{\partial X_{s_1} \cdots \partial X_{s_i}} = d(d-1)\cdots(d-i+1)F,$$

where the sum is over all i-tuples $s_1, \ldots s_i$. The sum on the left is $i!F_i(X_0, \ldots, X_n)$; the statement in the text is superior when $0 < \operatorname{char} K \leq i$.

Corollary 1.5.1 *If* $F(c_0, \ldots, c_n) = 0$ *then* $F_i(c_0, \ldots, c_n) = 0$ *for* $i \geq 0$.

Consider the 3-variable case $F = F(X, Y, Z)$ and the corresponding plane projective curve $C = \mathsf{Z}(F)$. (See §1.3.) We write λ, μ, ν for $\lambda_1, \lambda_2, \lambda_3$. The **order** of a point $P = (X_0, Y_0, Z_0) \in C$ is the minimal i such that F_i is not identically 0 as a polynomial in λ, μ, ν. If $i = 1$ then P is an **ordinary** or **nonsingular point**, while if $i > 1$ then P is a **singular point** or a **singularity of order *i***. The polynomial F, or the corresponding curve C is **nonsingular** or **smooth** if it has no singular points defined over an algebraic closure of K, and therefore in fact none defined over any extension of K (by Proposition 1.3.2(c)).

Proposition 1.5.2 *Let* $F = F(X, Y, Z)$ *be a nonzero homogeneous polynomial. If* F *is nonsingular then it is absolutely irreducible*, i.e., *irreducible over* $\overline{K}$.

Proof. Let $F = GH$ where G and H are homogeneous of positive degree defined over $\overline{K}$, and let P be a point of intersection on the curves corresponding to G

and H (Proposition 1.3.2(c)). Then $F_X = GH_X + G_X H$ vanishes at P, and similarly for the other variables. Thus P is a singular point of F. ■

Let $P = (X_0, Y_0, Z_0)$ be a point of order $i \geq 1$ on $F = 0$. The **tangent cone** at P is

$$\mathsf{Z}(F_i) = \{(\lambda, \mu, \nu) \in \mathbf{P}^2 : F_i(\lambda, \mu, \nu) = 0\}.$$

By the previous corollary, the tangent cone contains the point P. It can be shown that over $\overline{K}$, F_i is a product of i linear forms $aX + bY + cZ$, each satisfying $aX_0 + bY_0 + cZ_0 = 0$; thus the tangent cone consists of i lines through P, possibly some coincident, called the **tangent lines at *P***.

It is much easier to calculate these tangent lines in affine coordinates as follows. Effect a linear change of coordinates so that $P = (0, 0, 1)$. Then in terms of $x = X/Z$, $y = Y/Z$,

$$Z^{-d}F(X, Y, Z) = f(x, y) = f_{i'} + f_{i'+1} + \cdots$$

where f_j is homogeneous in x, y of degree j. It can be shown that $i' = i$, the order of P, that f_i is the product of i linear factors of the form $ax + by$, and the tangent lines are $aX + bY = 0$.

In the case of an ordinary point P on C, when $i = 1$, there is a unique tangent line through P, namely

$$a_X X + a_Y Y + a_Z Z = 0 \quad \text{where} \quad a_X = F_X(X_0, Y_0, Z_0) \quad \textit{etc.}$$

A point of order 2 with distinct tangents is called a **node**, while a point of order 2 with coincident tangents is a **cusp**. The appearance of a node and a cusp in the real case are shown on the following interleaf.

Examples

1. The point (0,1,0) at infinity on the curve defined by the Weierstrass equation $F = Y^2Z + \cdots - a_6Z^3 = 0$ is always nonsingular since

$$F_Z = Y^2 + a_1XY + 2a_3YZ - a_2X^2 - 2a_4XZ - 3a_6Z^2$$

has the value $1 \neq 0$ at that point. The other two derivatives are 0 there, so the tangent line is $Z = 0$. Thus to locate any possible singularities on the Weierstrass form we can use the affine version.

2. $y^2 = x^3 + ax^2$ has a singularity of order 2 at $(x, y) = (0, 0)$:

$$f_2 = y^2 - ax^2 = (y - \sqrt{a}x)(y + \sqrt{a}x),$$

and so the tangents there are $X \pm \sqrt{a}Y = 0$. Thus $(0, 0, 1)$ is a node if $2a \neq 0$ (with irrational tangents if $\sqrt{a} \notin K$), and a cusp if $2a = 0$.

3. On $F = Y^2Z - X^3 - XZ^2$, $P = (0, 0, 1)$ is an ordinary point and the tangent there is $X = 0$. If $\operatorname{char} K = 2$ then $(1, 0, 1)$ is singular with $F_2 = (\lambda + \mu + \nu)^2$, hence is a cusp with tangent $X + Y + Z = 0$.

4. Let $K = \mathbf{Q}$ and

$$F = 11X^3 + 12X^2Y - 9XZ^2 - Y^3 + 2Z^3.$$

Substituting $X = 1$, $Y = 2$, $Z = 3$ in the Taylor expansion of F, we find that $F_0 = F_1 = 0$ and

$$F_2 = 57\lambda^2 + 24\lambda\mu - 6\mu^2 - 54\lambda\nu + 9\nu^2 = L_1L_2$$

where $$L_{1,2}(\lambda, \mu, \nu) = (\pm 2\sqrt{6} - 9)\lambda \mp \sqrt{6}\mu + 3\nu.$$

Thus $P = (1, 2, 3)$ is a node on $F = 0$ with tangent lines $L_{1,2}(X, Y, Z) = 0$.

Alternatively, take the affine equation

$$f(x, y) = 11x^3 + 12x^2y - 9x - y^3 + 2.$$

Now P has coordinates $x = 1/3$, $y = 2/3$, and

$$f(1/3 + l, 2/3 + m) = f_2 + f_3$$

where $$f_3 = 11l^3 + 12l^2m - m^3$$

and $$f_2 = 19l^2 + 8lm - 2m^2 = (19l + (4 - 3\sqrt{6})m)(19l + (4 + 3\sqrt{6})m)/19.$$

It is comparatively easier to find the factors of f_2 than F_2. Substituting $x = X/Z$, $y = Y/Z$ in the equations of the tangent lines $19(x - 1/3) + (4 \pm 3\sqrt{6})(y - 2/3) = 0$, a brief calculation shows that they give the same lines as $L_{1,2} = 0$.

Proposition 1.5.3 *For any field K and any $a_1, \ldots, a_6 \in K$,*

$$F = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

is irreducible (even if $\Delta = 0$).

Proof. Suppose $F = GH$ is a nontrivial factorization, say

$$G = aX + bY + cZ.$$

Substituting $Z = 0$ in $F = GH$ yields $-X^3 = (aX + bY)\overline{G}$, hence $a \neq 0$. Now substituting $X = -cZ/a$ yields

$$Y^2Z + dYZ^2 + eZ^3 = 0$$

for certain $d, e \in K$, which is an impossible identity. ■

Proposition 1.5.4 *The Weierstrass equation is singular iff $\Delta = 0$ and then there is a unique singularity of order 2 as follows:*

- *If $c_4 \neq 0$ there is a K-rational node at the point with coordinates*

$$\begin{aligned} x_0 &= (18b_6 - b_2b_4)/c_4 \\ y_0 &= (b_2b_5 + 3b_7)/c_4 \quad = \begin{cases} -(a_1x_0 + a_3)/2 & \text{if } \mathrm{char}\, K \neq 2 \\ (a_3^2 + a_1^2a_4)/a_1^3 & \text{if } \mathrm{char}\, K = 2 \end{cases} \end{aligned}$$

  where

$$\begin{aligned} b_5 &= a_1a_4 - 2a_2a_3 \\ b_7 &= a_1(a_3^2 - 12a_6) + 8a_3a_4. \end{aligned}$$

  The two tangents are given in terms of the parameter t by $x = x_0 + t$, $y = y_0 + \mu t$ for the two distinct roots of the separable polynomial $\mu^2 + a_1\mu - 3x_0 - a_2 = 0$. When char $K \neq 2$, these are

$$\mu = \frac{-a_1c_4 \pm \sqrt{-c_4c_6}}{2c_4}.$$

- *If $c_4 = 0$ there is a cusp at the point with coordinates*

$$\begin{array}{llllll} \mathrm{char}K = 2: & x_0 & = & \sqrt{a_4}, & y_0 & = & \sqrt{a_2a_4 + a_6} \\ \mathrm{char}K = 3: & x_0 & = & -\sqrt[3]{a_3^2 + a_6}, & y_0 & = & a_1x_0 + a_3 \\ \mathrm{char}K \neq 2,3: & x_0 & = & -b_2/12, & y_0 & = & -\frac{1}{2}(a_1x_0 + a_3). \end{array}$$

  (The cusp can be irrational only when K is an imperfect field of characteristic 2 or 3.) The unique tangent line is $x = x_0 + t$, $y = y_0 + \mu t$ where $\mu = \sqrt[4]{a_4} + \sqrt{a_2}$ when char $K = 2$, and $\mu = -a_1/2$ otherwise.

In either case

$$\begin{aligned} f_x &= a_1y_0 - 3x_0^2 - 2a_2x_0 - a_4 = 0, \\ f_y &= 2y_0 + a_1x_0 + a_3 = 0. \end{aligned}$$

A singular Weierstrass equation remains singular over every field extension K'/K; moreover, the nature of the singularity (node or cusp) is constant.

Proof. Since the proof is by straightforward calculation we only give a sketch.

First let char $K \neq 2, 3$. Then as detailed above, a linear change of the (affine) coordinates — which clearly does not affect the occurence of singularities — allows us to take the simple form

$$\begin{aligned} f &= \eta^2 - \xi^3 + \frac{c_4}{48}\xi + \frac{c_6}{864} \\ f_\eta &= 2\eta \\ f_\xi &= -3\xi^2 + \frac{c_4}{48} \end{aligned}$$

If these three quantities are 0 then $\xi = \pm\sqrt{c_4}/12$, $\eta = 0$, $c_6 = \mp\sqrt{c_4}^3$, hence $\Delta = 0$ and the Taylor expansion reduces to

$$f(\pm\sqrt{c_4}/12 + \lambda, \mu) = (\mu^2 \mp \sqrt{c_4}\lambda^2/4) - \lambda^3.$$

Thus the singularity is of order 2 and the number of tangents is 2 or 1 according as $c_4 \neq 0$ or $c_4 = 0$.

Secondly let $\operatorname{char} K = 2$. Then $b_2 = a_1^2$, $b_4 = a_1a_3$, $c_4 = a_1^4$ so $c_4 = 0$ iff $a_1 = 0$. A singularity will be at a common zero of

$$\begin{aligned} f &= y^2 + a_1xy + a_3y + x^3 + a_2x^2 + a_4x + a_6, \\ f_x &= a_1y + x^2 + a_4, \\ f_y &= a_1x + a_3. \end{aligned}$$

If $a_1 = 0$ then, in order that $f_y = 0$, we have $a_3 = 0$ hence $\Delta = 0$, and we find $x_0 = \sqrt{a_4}$, $y_0 = \sqrt{a_2a_4 + a_6}$. The Taylor expansion of $f(x_0 + \lambda, y_0 + \mu)$ works out to

$$(\sqrt[4]{a_2} + \sqrt{a_4}\lambda + \mu)^2 + \lambda^3$$

so the singularity is a cusp.

If $a_1 \neq 0$ then $x = a_3/a_1$ (so that $f_y = 0$), which is the value in characteristic 2 stated by the proposition for x_0 in the node case, and $y = (a_3^2 + a_1^2a_4)/a_1^3$ (so that $f_x = 0$). The condition that $f = 0$ works out to $\Delta = 0$, and the Taylor expansion is

$$\left(\frac{a_3}{a_1} + a_2\right)\lambda^2 + a_1\lambda\mu + \mu^2 + \lambda^3.$$

Thus the tangent slopes are the roots of $\mu^2 + a_1\mu + (a_3/a_1 + a_2) = 0$, and $a_1 \neq 0$ guarantees that they are distinct, *i.e.,* the equation is separable.

The case of characteristic 3 is just as straightforward. ■

1.5.1 Example: No $E_{/\mathbf{Z}}$ has $\Delta = 1$ or -1

Let E be defined over $\mathbf{Z}$, *i.e.,* all the Weierstrass coefficients $a_i \in \mathbf{Z}$; this is indicated notationally by $E_{/\mathbf{Z}}$. Since Δ is a polynomial in the a_i with coefficients in $\mathbf{Z}$, therefore $\Delta \in \mathbf{Z}$. When we interpret the a_i mod p to obtain a Weierstrass equation over the p-element field $\mathbf{F}_p$, the discriminant is Δ mod p. Thus by the previous proposition, the mod p equation gives an elliptic curve when p is not a divisor of Δ. We now prove that this fails for at least one p:

Proposition 1.5.5 (Tate, *cf.* [Ogg66])

Let the elliptic curve E be defined over $\mathbf{Z}$. Then Δ is neither 1 nor -1. More generally, Δ does not have the form δ^3 where δ is a nonzero integer all of whose prime divisors are $\equiv 1 \bmod 8$.

Proof. Suppose $E_{/\mathbf{Z}}$ has $\Delta = \delta^3$, with δ as described in the proposition; in particular, $\delta \equiv \pm 1 \bmod 8$. Let $v_p(n)$ denote the exponent of a prime p in the unique factorization of a nonzero integer n.

If a_1 is even, then, by the formulas in §1.1, $v_2(b_2) \geq 2$, $v_2(b_4) \geq 1$, $v_2(c_4) \geq 4$, hence from

$$c_6^2 = c_4^3 + 1728\delta^3, \qquad \P$$

since δ is odd, we have $v_2(c_6) = 3$, say $c_6 = 8c$. Then ¶ implies the impossibility $c^2 \equiv 27\delta^3 \equiv \pm 3 \bmod 8$.

Therefore a_1 is odd, hence b_2 is odd and $c_4 = b_2^2 - 24b_4 \equiv 1 \bmod 8$. Substituting $x = c_4 + 12\delta$ and $y = c_6$ in ¶ gives

$$y^2 = x(x^2 - 36\delta x + 432\delta^2) = xQ, \quad \text{say,}$$

where $x \equiv 5 \bmod 8$, in particular, $x \neq 0$. Since $Q = (x - 18\delta)^2 + 108\delta^2 > 0$, it follows that $x = y^2/Q > 0$. Thus

$$x = 3^\alpha \prod p^{\beta_p} \prod q^{\gamma_q},$$

where p runs through the prime divisors of $\gcd(x, \delta)$, and q through any remaining prime divisors > 3 of x. Since $v_q(Q) = 0$, each $\gamma_q = v_q(y^2)$ is even, and by assumption each $p \equiv 1 \bmod 8$. Hence $x \equiv 3^\alpha \equiv 1$ or $3 \bmod 8$, which contradicts $x \equiv 5 \bmod 8$. ■

The following examples show the need for the assumption on the divisors of δ.

$$y^2 + y = x^3 + x^2 - 9x - 15, \quad \Delta = -19^3, \qquad \textbf{B19}$$

$$y^2 + y = x^3, \quad \Delta = -3^3, \qquad \textbf{A27}$$

$$y^2 = x^3 - x, \quad \Delta = 2^6, \qquad \textbf{A32}$$

$$y^2 + y = x^3 + x^2 - 23x - 50, \quad \Delta = 37^3, \qquad \textbf{C37}$$

$$y^2 + xy = x^3 - x^2 - 2x - 1, \quad \Delta = -7^3, \qquad \textbf{A49}$$

$$y^2 + y = x^3 + x^2 - 121x - 64, \quad \Delta = 5^3 97^3. \qquad \textbf{A}_{\text{ı}}\textbf{485}$$

For a given number field K, a natural question is whether there exist E defined over the ring of integers of K with Δ a unit. Stroeker [Str83] has proved that this does not occur when K is imaginary quadratic; but we must postpone the proof. Unit Δ do occur over real quadratic fields: Tate gave the example (*cf.* [Ser72, p.320])

$$y^2 + xy + \epsilon^2 y = x^3, \quad \epsilon = (5 + \sqrt{29})/2, \quad \Delta = -\epsilon^{10},$$

(ϵ is in fact the fundamental unit of $\mathbf{Q}(\sqrt{29})$) and several others occur in the table in §4.4.

It is a triviality to find E defined over the ring of integers of a number field with $\Delta = 1$. For example, $y^2 + a_1xy + a_3 = x^3$ has $\Delta = (a_1^3 - 27a_3)a_3^3$; choosing $a_3 = 1$ and $a_1 = \sqrt[3]{28}$ yields $\Delta = 1$. However I do not know of an example of $\Delta = 1$ or -1 over a quadratic field. Here is an example over the biquadratic field $\mathbf{Q}(\sqrt{2}, \sqrt{7})$, which contains $\sqrt{8 - 3\sqrt{7}} = (3 - \sqrt{7})/\sqrt{2}$:

$$y^2 + \sqrt{8 - 3\sqrt{7}}\, xy = x^3 - 8x^2 + (8 + 3\sqrt{7})x, \quad \Delta = 1, \quad j = 255^3.$$

1.6 Affine coord. ring, function field, generic points

We use the abbreviation UFD for unique factorization domain. Recall ([BAC7], p.36) that if A is a UFD then so is the polynomial ring $A[x]$. It follows that $\mathbf{Z}[\{x_i\}]$ and $K[\{x_i\}]$ (K any field) are UFD's for an arbitrary set of indeterminates, *i.e.,* independent transcendentals.

Let S and T be independent transcendentals over the field K, let $a_1, \ldots, a_6 \in K$ and let

$$f(S,T) = T^2 + a_1ST + a_3T - S^3 - a_2S^2 - a_4S - a_6.$$

Lemma 1.6.1 *The principal ideal* $(f(S,T))$ *in the polynomial ring* $K[S,T]$ *is prime.*

Proof. We must prove that f is irreducible. If $f = gh$, then by substituting $S = X/Z$, $T = Y/Z$ and multiplying by Z^3 we get a factorization $F = GH$ of homogeneous polynomials. The result follows by Proposition 1.5.3. ■

Thus

$$A = K[S,T]/(f(S,T))$$

is an integral domain (even if $\Delta = 0$). Writing x and y for the residue classes of S and T mod $f(S,T)$, we have

$$A = K[x,y].$$

The equation $f(S,T) = 0$ defines a curve E in the S,T-plane; but it is customary to replace S and T by x and y, and say that E is given by $f(x,y) = 0$ in the x,y-plane. That is, x and y stand for a pair of independent transcendentals, and also for a pair of variables related by the equation $f(x,y) = 0$. This mild ambiguity causes no problems in practice.

The integral domain A is the **affine coordinate ring** of E, and its quotient field $L = K(x,y)$ is the **function field** of E. The field L can also be described as the quadratic extension $K(x)(y)$ of the rational function field $K(x)$ defined by the polynomial $f(x,y)$, which is quadratic in y; alternatively, $L = K(y)(x)$ is the cubic extension of the simple transcendental extension $K(y)$ of K. When

$\Delta \neq 0$, both the quadratic and cubic extensions are separable (though in general the cubic extension is not Galois). For if L is an inseparable extension of $K(x)$, then char $K = 2$ and $f_y = 2y + a_1x + a_3 = 0$, *i.e.*, $a_1x + a_3 = 0$ which implies $a_1 = 0$ and $a_3 = 0$ and then one calculates $b_2 = 0, \ldots$ leading to $\Delta = 0$; similarly for the cubic extension.

The subfield K of L is called either the **ground field**, which emphasizes that K is the field containing $a_1, \ldots, a_6$ that we started with, or the **constant field** (or *field of constants*), which emphasizes the fact that K is algebraically closed in L.

Let $E(K)$ denote the set of points (a, b) on E defined over K (that is, $a, b \in K$ and $f(a, b) = 0$) together with the one point O at infinity. As explained in Proposition 1.5.4, if $\Delta = 0$ there is exactly one singular point, which is never O, while if $\Delta \neq 0$ then E is nonsingular and is, by definition, an elliptic curve. If K' is any extension field of K, then we can regard E as being defined over K' and so $E(K')$ is defined. In particular, $(x, y) \in E(L)$ since the point (x, y) satisfies $f(x, y) = 0$ by definition.

Now, for each nonzero point $(a, b) \in E(K)$, we have a K-algebra homomorphism $A \longrightarrow K$ defined by $x \mapsto a$ and $y \mapsto b$. Thus every nonzero point of $E(K)$ is obtained by specializing the values of x and y, and for this reason (x, y) is called a **generic point**. (We could include O by taking a projective generic point (X, Y, Z) satisfying the projectivized Weierstrass equation $F(X, Y, Z) = 0$, should the need arise.)

When several generic points (x_1, y_1), (x_2, y_2), ... are needed, take the field $K(x_1, y_1, x_2, y_2, \ldots)$ where x_1, x_2, ... are independent transcendentals and each y_i defines a quadratic extension by the equation $f(x_i, y_i) = 0$.

1.7 The group law: nonsingular case

The set of points $E(K)$ on an elliptic curve has a natural structure of an abelian group. This has a simple geometric description when E is a nonsingular plane cubic with a rational point O, for example when E is given by a Weierstrass equation with $\Delta \neq 0$, and O is the point at infinity; a non-Weierstrass example is the Fermat curve $X^3 + Y^3 + Z^3 = 0$ with $O = (1, -1, 0)$. The description depends on the fact that a line in $\mathbf{P}^2$ meets a cubic in 3 points when the points of intersection are properly counted, as described in §1.3. In this section the details will become clear for the Weierstrass equation by direct algebraic calculation. But first we describe the geometric construction of the group operations for the general nonsingular cubic.

Let O be the chosen point in $E(K)$ and let the tangent at O meet E in the third point O'. Note that $O' = O$ iff O is a flex; this is the case for the Weierstrass equation since the line at ∞ meets E only at O. Now let $P, Q \in E(K)$ and let the line joining P and Q meet the cubic in the third point R; 2 or even 3 of these 3 points may coincide. The third point of intersection

of the line joining R and O is defined to be $P+Q$; the third point on the line joining P and O' (not O unless O is a flex!) is $-P$; and O is the zero of the group. These constructions are illustrated in a real example on the following interleaf.

As an exercise the reader may note that when O is a flex every flex F satisfies $F+F+F=0$. It is a fact that a nonsingular cubic over an algebraically closed field of characteristic $\neq 3$ has exactly 9 flexes.

Proposition 1.7.1 *Let C be a nonsingular cubic defined over the field K and let $O \in C(K)$.*

(a) *With $+$ and $-$ as described above, $C(K)$ is an abelian group with neutral element O.*

(b) *If $O_1, O_2 \in C(K)$ and for $i=1,2$, $C(K)_i$ denotes the group determined by choosing O_i as neutral element, then a group isomorphism $C(K)_1 \longrightarrow C(K)_2$ is defined by*

$$P \longmapsto P + O_2$$

where $+$ denotes addition in $C(K)_1$.

The associative law and statement (b) are not obvious from the geometric definitions. Since they will become "transparent" after we discuss divisors in Chapter 6, for now we leave the proof "to the reader" as an arduous computer exercise. For a direct proof see [Kna92,p.67].

As an example we reconsider the curves of Corollary 1.4.7.

Corollary 1.7.2 *Let C denote the plane cubic curve*

$$a_1X_1^3 + a_2X_2^3 + a_3X_3^3 + dX_1X_2X_3 = 0$$

where a_1, a_2, a_3, d are integers and the a_j are positive, distinct and square-free. Then C is nonsingular, hence absolutely irreducible.

Suppose the set $C(\mathbf{Q})$ of rational points on C in $\mathbf{P}^2(\mathbf{Q})$ is nonempty, say $O \in C(\mathbf{Q})$. With O as neutral element, the group $C(\mathbf{Q})$ contains at least one point O' of infinite order, namely the third point of intersection with C of the tangent at O. In particular, $O' \neq O$, and it follows that none of the flexes is rational over $\mathbf{Q}$. †

Proof. Suppose $P=(X_1,X_2,X_3)$ is a singular point (defined over $\overline{\mathbf{Q}}$). Then

$$dX_1X_2X_3 = -3a_iX_i^3, \quad i=1,2,3.$$

It follows that $X_1X_2X_3 \neq 0$, hence (within a common factor) $X_i = 1/\sqrt[3]{a_i}$ from which one obtains $27a_1a_2a_3 + d^3 = 0$. But the last equation is not allowed by the assumptions.

†This is also obvious by direct calculation: if H denotes the 3×3 Hessian determinant of $F = a_1X_1^3 + \cdots + dX_1X_2X_3$, then the flexes are the points of intersection of the curves $F=0$ and $H=0$. They are $(0, \sqrt[3]{a_3}, -\sqrt[3]{a_2})$, *etc.* (9 points in all).

Now let P_1 be any point in $C(\mathbf{Q})$ and let $P_2, \ldots$ be the sequence described in Corollary 1.4.7. The geometric construction of addition shows that

$$P_2 + [2]P_1 = O', \quad \text{or} \quad P_2 = O' - [2]P_1,$$

hence
$$P_3 = O' - [2]P_2 = -O' + [4]P_1, \ \textit{etc.}$$

Solving the recurrence we find

$$P_n = \left[\frac{1-(-2)^{n-1}}{3}\right] O' + \left[(-2)^{n-1}\right] P_1.$$

In particular, by Proposition 1.4.7, the sequence

$$O_n = \left[\frac{1-(-2)^{n-1}}{3}\right] O'$$

consists of distinct points, and therefore O' has infinite order. ■

With $O \in C(\mathbf{Q})$ as in the corollary, one might jump to the false conclusion that the group $C(\mathbf{Q})$ is torsion-free (as did Selmer at the beginning of [Sel54] and Cassels [Cas66,p.264] — but none of their subsequent statements are invalidated). An example is the curve $u^3 + 2v^3 + 5 - 8uv = 0$ with $O = (1,1)$ and point $(3,2)$ of order 2 (alternatively with $O = (3,2)$ and $(1,1)$ of order 2). This example is plotted on an interleaving sheet.† Some similar examples are $u^3+3v^3+6+4uv = 0$ with points $(1,-1)$ and $(-3/2, 1/2)$; $u^3+2v^3+6-14uv = 0$ with $(-4,1)$, $(2/3, 5/3)$; $u^3 + 2v^3 + 7 - 12uv = 0$ with $(1,2)$,$(3,1)$. In Chapter 6 we will see that for elliptic curves as in the corollary and with a rational point, the order of the torsion subgroup is one of 1,2,3,4,6,9,12 (and is 1 in the Selmer case $d = 0$). However I have been able to find examples only of orders 1 and 2.

We now describe algebraically the group operations for a Weierstrass equation. Since O is going to be the group 0 and since it is the only point at ∞, we can confine our description of $-P_1$ and $P_1 + P_2$ to affine coordinates: let $P_i = (x_i, y_i)$. The line $x = x_1$ contains the point P_1 and, considering its projective version $X = x_1 Z$, it also contains O. Thus $-P_1$ is the third point of intersection, which therefore has x-coordinate x_1 and it remains to calculate the y-coordinate. When we substitute x_1 for x in the Weierstrass equation we obtain a quadratic equation for y:

$$y^2 + (a_1x_1 + a_3)y - (x_1^3 + a_2x_1^2 + a_4x_1 + a_6) = 0.$$

The sum of the roots is $-(a_1x_1 + a_3)$, and one root is y_1, hence the other root, which is the y-coordinate of $-P_1$, is $-a_1x_1 - a_3 - y_1$.

†We note that the locus of a real projective cubic curve is never contained in an affine part of $\mathbf{P}^2(\mathbf{R})$, *i.e.,* the graph is never finite, as is the case for example with ellipses, since a cubic polynomial with real coefficients has a real root, and therefore the line at infinity always intersects the cubic in a real point.

Next let us calculate $P_1+P_2 = P_3 = (x_3, y_3)$. If $x_1 \neq x_2$, *i.e.*, $P_1 \neq \pm P_2$ then the line joining P_1 and P_2 is $y = y_1 + \lambda(x - x_1)$ where $\lambda = (y_2 - y_1)/(x_2 - x_1)$. Substituting this expression for y into the Weierstrass equation gives a cubic equation for x whose three roots are x_1, x_2, x_3. Identifying the sum of the roots with the negative of the coefficient of x^2 yields $x_3 = -x_1 - x_2 - a_2 + a_1\lambda + \lambda^2$ and putting this into the equation of the line gives the y-coordinate of $-P_3$ from which we find $y_3 = -[y_1 + (x_3 - x_1)\lambda + a_1x_3 + a_3]$.

There remains the case $P_1 = P_2$ which is treated similarly where now $y = y_1 + \lambda(x - x_1)$ is the tangent line. We leave to the reader the calculation of λ as well as a few other details in the following proposition.

Notation: For any abelian group A and $m \in \mathbf{Z}$, $[m]$ denotes the endomorphism multiplication by m and $A[m]$ denotes $\ker[m]$; if m' is a divisor of m then $A[m']$ is a subgroup of $A[m]$. When $m > 0$ the elements of $A[m]$ not in $A[m']$ for any proper divisor m' of m are called ***m*-division points**. For example for $P \in E(K)$ we have $[-1]P = -P$ and the 2-division points defined over K are those $P \neq O$ satisfying $[2]P = P + P = O$. As will be explained in detail in §1.7.2, there are only finitely many m-division points defined over any extension field of K and adjoining the x and y coordinates of all these points gives a finite extension of K called the ***m*-division field** of E. The usual Weierstrass coordinates of a point $P \in E(K)$ are denoted $x(P)$ and $y(P)$. This notation is extended to any function f of x and y: $f(P)$ simply means the value of f when the coordinates of P are substituted for x and y. Thus, maintaining the notation introduced in §1.1, when $\operatorname{char} K \neq 2$:

$$\eta(P) = y(P) + (a_1x(P) + a_3)/2.$$

Proposition 1.7.3 *For points on an elliptic curve in Weierstrass form we have*

$$-(x_1, y_1) = (x_1, -y_1 - a_1x_1 - a_3).$$

Hence the points of order 2 *in the group are as follows:*

> $\operatorname{char} K = 2$*: if* $a_1 = 0$*, equivalently* $j = 0$*, there are no points of order* 2*; if* $a_1 \neq 0$ *there is a unique point of order* 2*, possibly quadratic over* K*:*
>
> $$[2]\left(a_3/a_1, \left[a_4 + \sqrt{b_8 + (a_3^3/a_1)}\right]/a_1\right) = O$$
>
> $\operatorname{char} K \neq 2$*: there are exactly* 3 *points of order* 2*, possibly some irrational over* K*:* $(x, \eta) = (x_i, 0)$ *where* x_i *runs through the three roots of*
>
> $$x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4} = 0.$$

For $(x_2, y_2) \neq -(x_1, y_1)$ *we have the addition law*

$$(x_1, y_1) + (x_2, y_2) = (x_3, y_3)$$

where

$$\begin{aligned} x_3 &= -x_1 - x_2 - a_2 + a_1\lambda + \lambda^2 \\ y_3 &= -y_1 - (x_3 - x_1)\lambda - a_1 x_3 - a_3 \end{aligned}$$

and

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} & \text{if } x_1 \neq x_2 \\ \dfrac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} & \text{if } x_1 = x_2 \end{cases}$$

Hence

$$x([2](x,y)) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6}.$$

When $\operatorname{char} K = 2$,

$$y([2](x,y)) = (c_1y + c_2)/(a_1x + a_3)^3 \qquad \textit{where}$$

$$\begin{aligned} c_1 &= a_1x^4 + a_1b_8 + a_3b_6, \\ c_2 &= x^6 + a_4x^4 + (b_6 + a_4b_2)x^3 + (b_8 + a_4b_4)x^2 \\ &\quad + (b_2b_8 + (a_4 + b_4)b_6)x + (b_4b_8 + b_6^2 + a_4b_8). \end{aligned}$$

When $\operatorname{char} K \neq 2$

$$\eta([2](x,\eta)) = f(x)/16\eta^3$$

where $f(x) =$

$$2x^6 + b_2x^5 + 5b_4x^4 + 10b_6x^3 + 10b_8x^2 + (b_2b_8 - b_4b_6)x + (b_4b_8 - b_6^2).$$

■

There is a special case of the duplication formula that we record in a corollary for future reference:

Corollary 1.7.4 *If* $\operatorname{char} K \neq 2$ *and*

$$y^2 = x(x^2 + ax + b)$$

then

$$[2](x,y) = \left(\left(\frac{x^2 - b}{2y} \right)^2, \frac{(x^2 - b)(x^4 + 2ax^3 + 6bx^2 + 2abx + b^2)}{8y^3} \right).$$

■

Many numerical examples of adding points are given in the standard texts. We content ourselves with the following four.

Example 1. Let $K = \mathbf{Q}(t)$ be a simple transcendental extension of the rational field. Then on

$$y^2 = x^3 + tx^2 - tx$$

one calculates†

$$\Delta = 16t^3(t+4), \qquad j = 16^2(t+3)^3/(t+4),$$

$$(0,0) + (1,1) = (-t, t),$$

$$[2](0,0) = O,$$

$$[2](1,1) = \left(\left(\frac{t+1}{2}\right)^2, -\frac{1}{8}(t^3 + 5t^2 + 3t - 1)\right),$$

$$[3](1,1) = (a^2/d^2, ab/d^3)$$

where $a = t^2 + 6t + 1$, $b = 3t^4 + 16t^3 + 22t^2 + 24t - 1$, and $d = (t+2)(t+3)$.

Example 2. For the twisted Fermat curve $u^3 + v^3 = a$, $y^2 = x^3 - 432a^2$, $x = 12a/(u+v)$, etc., introduced in Corollary 1.4.2, we find

$$-(u,v) = (v,u)$$

by transforming to (x,y) coordinates, doing the calculation, then transforming the result back to u, v coordinates. Similarly one can give (rather complicated) formulas for $[2](u,v)$ and the addition of two points. Alternatively one can work directly in u, v coordinates using the geometric constructions. The plot on the following interleaf shows

$$(1,12) + (9,10) + (-37/3, 46/3) = O$$

on the 'taxicab curve' — the case $a = 1729$.

Example 3. The "generic $\mathbf{R}$-example" is depicted in the diagram. The equation of the horizontal line is $\eta = y + (a_1x + a_3)/2 = 0$ and the line joining a point P with O is the vertical line through P. Let us denote the connected component of O by C_1; it is the part on the right passing through P_1. The second real component C_2, the dotted oval part, is present when $\Delta > 0$; then C_1 is a subgroup of index 2 in $E(\mathbf{R})$ and C_2 is a coset. Thus $P \in C_2 \Rightarrow [2]P \in C_1$. The real points of order 2 are P_1 and, if $\Delta > 0$, P_2 and P_3. The points of order 3 as indicated in the diagram are Q and $-Q$; the real flexes are $O, \pm Q$. As we will see in Proposition 1.7.8, the point $P = (x,y)$ satisfies $[3]P = O$ iff x is a root of a certain 4-th degree polynomial $\psi_3(x)$. An easy Sturm's theorem calculation (*cf.* [Con82], p.273) shows that ψ_3 always has exactly 2 real roots. One of

†Actually **aPecs** made these calculations — see the appendix to this chapter.

these roots gives two corresponding real values of y, hence the points $\pm Q$, but the values of y corresponding to the other real root x are always nonreal.

Example 4. On the next interleaf a particular real case is plotted, actually one defined over $\mathbf{Q}$, which we have deliberately chosen with $a_1 \neq 0$ to illustrate the fact that the change from (x, y) to (x, η) coordinates given by $\eta = y + (a_1 x + a_3)/2$ is not orthogonal. Therefore the x-axis symmetry illustrated in the previous figure is now skewed. But notice that a point P and its negative are still joined by a vertical line and, in particular, the tangents at points of order 2 are vertical.

When $\operatorname{char} K \neq 2$ the x-coordinates of the 2-division points are the roots of

$$f(x) = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4}.$$

Since $\operatorname{char} K \neq 2$, this polynomial is always separable over K; for it could be inseparable only if $\operatorname{char} K = 3$ and $b_2 = b_4 = 0$, but then $\Delta = 0$.

Let e_i, $i = 1, 2, 3$ denote the roots and let K_2 denote the 2-division field $K(e_1, e_2, e_3)$. Since Δ is 16 times the polynomial discriminant of $f(x)$ and $\Delta \neq 0$, by standard field theory the possibilities are as follows.
— all three $e_i \in K$: $K_2 = K$;
— just one $e_i \in K$: K_2 is quadratic over K;
— no $e_i \in K$ and Δ is a square in K: K_2 is Galois (cyclic) order 3 over K;
— no $e_i \in K$ and $\Delta \notin K^{*2}$: K_2 is Galois over K with group S_3, the symmetric group of order $3! = 6$.

The possibilities are illustrated by the following three examples over $\mathbf{Q}$.

$$y^2 + xy + y = x^3 + x^2 - 135x - 660 \qquad \Delta = 3^8 5^2 \tag{\textbf{E15}}$$

$$y^2 + xy + y = x^3 + x^2 + 35x - 28 \qquad \Delta = -3^2 5^8 \tag{\textbf{F15}}$$

$$y^2 = x^3 - x^2 - 2x + 1 \qquad \Delta = 2^4 7^2 \tag{\textbf{A196}}$$

The number of 2-division points defined over $\mathbf{Q}$ is respectively, 3,1,0. In fact one can determine (by methods to be described later) that the group of rational points in these cases is as follows (C_n denotes the cyclic group of order n and the coordinates are (x, y)):

$$\begin{aligned}
\mathbf{E15}(\mathbf{Q}) &= C_2 \times C_2 = \{O, (13, -7), (-7, 3), (-29/4, 25/8)\}, \\
\mathbf{F15}(\mathbf{Q}) &= C_8 = \{P = (2, 6), [2]P = (7, -29), [3]P = (32, 171), \\
&\quad [4]P = (3/4, -7/8), [5]P = (32, -204), [6]P = (7, 21), \\
&\quad [7]P = (2, -9), [8]P = O\}, \\
\mathbf{A196}(\mathbf{Q}) &= C_\infty = \langle (0, 1) \rangle.
\end{aligned}$$

The group orders $|\mathbf{E15}(\mathbf{Q})|$ and $|\mathbf{F15}(\mathbf{Q})|$, namely 4 and 8, are interchanged in table 1 of [AntIV]; remarkably this is the only misprint that has come to light in this manually typed catalog!

1.7.1 Halving points

Division by 2 is naturally a tad more complicated than multiplication by 2:

Proposition 1.7.5 *Let E be an elliptic curve defined over the field K, let* $\operatorname{char} K \neq 2$ *and let the x-coordinates of the 2-division points be e_i, $i = 1, 2, 3$, in a separable algebraic closure $\overline{K}^s$ of K.*

(a) *Let $Q \in E(K)$, $Q \neq O$. Then there exists $P \in E(K)$ such that $[2]P = Q$ iff $\forall i$, $x(Q) - e_i$ is a square in $K(e_i)$. When these three conditions are satisfied, let $x(Q) - e_i = \rho_i^2$ where the ρ_i are chosen so that*

(i) *they are algebraically compatible, i.e., $\sigma \in \operatorname{Gal}(\overline{K}^s/K)$, $\sigma e_i = e_j \Longrightarrow \sigma\rho_i = \rho_j$, and*

(ii) *so that $\eta(Q) = \rho_1\rho_2\rho_3$.* ‡

Then all the solutions P are given by

$$x(P) = x(Q) + \rho_1\rho_2 + \rho_1\rho_3 + \rho_2\rho_3,$$

$$\eta(P) = m(x(P) - x(Q)) - \eta(Q),$$

where $m = \rho_1 + \rho_2 + \rho_3$. Thus the equation of the line in the (x, η)-plane that is tangent to E at P and passes through $-Q = (x(Q), -\eta(Q))$ is

$$\eta = m(x - x(Q)) - \eta(Q). \tag{T}$$

‡If all $e_i \in K$ then condition (i) imposes no condition while if Q itself is a point of order 2 then one of the ρ is 0 and condition (ii) imposes no condition.

(a′) [Was87, Prop.4] *An alternative, rational criterion: the solutions of* $[2]P = Q$ *as in* (a) *are in* 1-1 *correspondence with the roots in* K *of the polynomial* $\mathrm{Quar}_Q(m)$ (defined in §1.2). *For each root* m*, the corresponding point* P *has coordinates*

$$x(P) = (m^2 - b_2/4 - x(Q))/2, \quad \eta(P) = m(x(P) - x(Q)) - \eta(Q),$$

and (T) *is the equation of the tangent line at* P.

(b) *In the quadratic case, that is, when one* $e_i \in K$ *and the other two are conjugate quadratic over* K*, there are simpler rational criteria for the existence of* P *as follows.*† *Replacing* x *by* $x + e_i$*, the equation takes the form*

$$y^2 = x(x^2 + ax + b) \qquad \textit{where} \quad d = a^2 - 4b \notin K^{*2},$$

so
$$e_1 = 0, \quad e_2 = (-a + \sqrt{d})/2, \quad e_3 = \overline{e_2} = (-a - \sqrt{d})/2.$$

Then P *exists iff*

- *when* $Q = (0, 0)$,
  - (i) $b \in K^{*2}$*, say* $b = r^2$*, and*
  - (ii) *one of* $a \pm 2r \in K^{*2}$ *;*

  choosing the sign of r *so that* $a + 2r = p^2$*, the two solutions are*

  $$[2](r, \pm rp) = (0, 0);$$

- *when* $Q = (s, t)$*,* $s \neq 0$,
  - (i) $s \in K^{*2}$*, say* $s = r^2$*, and*
  - (ii) *one of* $q_\pm = 2s + a \pm 2t/r \in K^{*2}$*;*

  choosing the sign of r *so that* $q_+ = p^2$*, the two solutions are*

  $$[2](S, \pm pS) = (s, t) \quad \textit{where} \quad S = s \pm pr + \frac{t}{r}.$$

(c) *If one of the* 2*-division points is defined over* K*, say* $e_1 \in K$*, then all three are defined over* K *iff* $\Delta \in K^{*2}$*. (Since* $j - 1728 = c_6^2/\Delta$*, when* $j \neq 1728$ *this is equivalent to* $j - 1728 \in K^{*2}$*.) And then*

$$e_2, e_3 = -\frac{1}{2}\left(e_1 + \frac{b_2}{4} \pm \sqrt{\Delta}/(2(6e_1^2 + b_2 e_1 + b_4))\right).$$

When this is the case, and the Weierstrass equation is written $y^2 = x(x^2 + ax + b) = x(x - e_2)(x - e_3)$*, the criteria for* $[2]P = Q$ *are the same as in* (b):

†as pointed out to me by John Cremona.

- *when* $Q = (0,0)$, *iff* $b = r^2$ *and* $a + 2r = p_+^2$, *hence* $a - 2r = p_-^2$ *where* $p_- = (e_2 - e_3)/p_+$; *then the four solutions are given by the four variations of sign in*
$$[2](ur, wrp_u) = (0,0), \quad u, w \in \{\pm 1\};$$
- $Q = (s,t)$, $s \neq 0$, *iff* $s = r^2$ *and* $q_+ = p_+^2$, *hence* $q_- = p_-^2$ *where* $p_- = \sqrt{d}/p_+$; *then the four solutions are given by the four choices of* $u, w \in \{\pm 1\}$ *in*
$$[2](S, wp_u S) = (s,t) \quad \textit{where} \quad S = s + wup_u r + ut/r.$$

Remarks. (a) is the general case, and so (b) and (c) are in a sense repetetive, but the formulas can be more convenient.

Referring to (a), since $x(P)$ and $\eta(P)$ are symmetric functions in the ρ_i they lie in K. When there is a solution $[2]P = Q$ for a given Q the number of solutions is the cardinality $|E(K)[2]|$ since any two solutions P differ by an element of $E(K)[2]$. This is all borne out by the formulas in (a). For example if all $e_i \in K$, so $|E(K)[2]| = 4$, then P exists iff all $\rho_i \in K$ and then

— if $\eta(Q) \neq 0$, among the 8 variations of the signs of the ρ_i only 4 qualify because of the requirement $\eta(Q) = \rho_1\rho_2\rho_3$;

— if $\eta(Q) = 0$ then one $\rho_i = 0$ and the 4 variations of the signs of the other two ρ_i all qualify.

Proof. (a) Suppose first that $[2]P = Q$. For each i make the coordinate shift $x = x' + e_i$, so the equation takes the form
$$\eta^2 = x'^3 + ax'^2 + bx',$$
and let $x'(P) = p \in K(e_i)$, $\eta(P) = q \in K$. By Corollary 1.7.4,
$$x'([2]P) = x([2]P) - e_i = \left(\frac{p^2 - b}{2q}\right)^2 \in K(e_i)^{*2},$$
so the conditions are necessary. The converse is once again a computational verification — a simple computer exercise.

(a′) Let
$$f(x) = x^3 + \frac{b_2}{4}x^2 + \frac{b_2}{2}x + \frac{b_6}{4}$$
so the b-form of E is $\eta^2 = f(x)$. If we write
$$f(x + x(Q)) = x^3 + bx^2 + cx + \eta(Q)^2,$$
then
$$\begin{aligned} b &= \frac{b_2}{4} + 3x(Q), \\ c &= \frac{b_4}{2} + \frac{b_2}{2}x(Q) + 3x(Q)^2. \end{aligned}$$

Assuming $[2]P = Q$, the tangent line at P can be written in the form (T), and

$$(mx - \eta(Q))^2 = f(x + x(Q)) = x^3 + bx^2 + cx + \eta(Q)^2$$

has the three roots $x = x(P) - x(Q)$, $x(P) - x(Q)$, 0. Their sum is $2(x(P) - x(Q)) = m^2 - b$, hence

$$x(P) = \frac{m^2 - b}{2} + x(Q), \quad \text{and} \quad \eta(P) = \frac{m(m^2 - b)}{2} - \eta(Q).$$

The 2-nd symmetric function of the roots is $(x(P) - x(Q))^2 = c + 2m\eta(Q)$, or

$$\left(\frac{m^2 - b}{2}\right)^2 = \frac{b_4}{2} + \frac{b_2}{2}x(Q) + 3x(Q)^2 + 2m\eta(Q),$$

which works out to $\mathrm{Quar}_Q(m) = 0$.

(b) First, suppose that $Q = (0, 0)$. In order to have $[2](x, y) = (0, 0)$, by Corollary 1.7.4 we need $b = r^2$ for some $r \in K^*$ and then with $x = \pm r$ we need

$$y^2 = \pm r(r^2 \pm ar + r^2) = r^2(a \pm 2r) \in K^{*2}.$$

Choosing the sign of r so that $a + 2r = p^2$, we have $[2](r, rp) = (0, 0)$ by the formula of that corollary.

Second, suppose $Q = (s, t)$, $s = r^2 \neq 0$. On the one hand if $s - e_2 = \rho^2$ where $\rho \in K(\sqrt{d})$, so $s - e_3 = \overline{\rho}^2$, then $\rho^2\overline{\rho}^2 = t^2/r^2$. Choose the sign of r so that $\rho\overline{\rho} = t/r$ and define $p = \rho + \overline{\rho}$. Then

$$p^2 = \rho^2 + \overline{\rho}^2 + 2\rho\overline{\rho} = s - e_2 + s - e_3 + 2t/r = 2s + a + 2t/r \in K^{*2}.$$

Conversely if (i) and (ii) are satisfied with $q_+ = p^2$ then the roots of $z^2 - pz + t/r$ are $\rho, \overline{\rho} = (p \pm \sqrt{q_-})/2$, and these are in $K(\sqrt{d})$ since $q_+q_- = d$. Hence $\sqrt{q_-} = \pm\sqrt{d}/p$ and $\rho^2, \overline{\rho}^2 = s - e_2, s - e_3$. The formula for P now follows from the general formula in (a).

(c) The e_i are the roots of

$$f(x) = x^3 + (b_2/4)x^2 + (b_4/2)x + b_6/4$$

$$= (x - e_1)(x^2 + (e_1 + b_2/4)x + (e_1^2 + b_2e_1/4 + b_4/2)).$$

The roots of the quadratic factor are

$$(-e_1 - b_2/4 \pm \tau)/2 \quad \text{where} \quad \tau = \sqrt{\Delta}/(4f'(e_1)).$$

To prove the statements concerning $[2]P = Q$, one just follows the proof of (b), noting that when $Q = (0, 0)$ then $(a + 2r)(a - 2r) = (e_2 - e_3)^2$, and when $Q = (s, t)$, $s \neq 0$, then $q_+q_- = d$ is a square. ■

Examples

On $E: \ y^2 = x(x^2 + 4x + 13)$ the point $Q = (1/4, 15/8)$ is not of the form $[2]P$, $P \in E(\mathbf{Q})$, since the e_j are $0, -2 \pm 3i$ and

$$\frac{1}{4} + 2 + 3i = \frac{3(2+i)^2}{4}$$

is not a square in $\mathbf{Q}(i)$. Note that when $x = 1/4$, both factors x and $x^2+4x+13$ of the right side of the Weierstrass equation are squares in $\mathbf{Q}$, so in general it is not sufficient that the K-irreducible factors evaluated at $x(Q)$ are squares in K.

On the other hand, for $Q = (0, 0)$ on

$$y^2 = x(x^2 - x + 1) \tag{\textbf{A24}}$$

we have $b = r^2$, $r = 1$ and $a + 2r = p^2$, $p = \pm 1$, hence

$$[2](1, \pm 1) = (0, 0),$$

and $(1, \pm 1)$ are 4-division points, that is, elements of order 4 in the group $E(\mathbf{Q})$.

The 2-division points of

$$y^2 + xy + y = x^3 - x^2 - x - 14 \tag{\textbf{C17}}$$

are $(11/4, -15/8)$ and $(-1 \pm 2i, \mp i)$. Since $11/4 - (-1 \pm 2i) = ((4 \mp i)/2)^2$, we have $\rho_1 = 0$ and $\rho_2, \rho_3 = (4 \pm i)/2$, so $x(P) = (11/4) + (17/4) = 7$,

and
$$[2](7, 13) = [2](7, -21) = \left(\frac{11}{4}, -\frac{15}{8}\right).$$

The final example is stated in a corollary for future reference.

Corollary 1.7.6 *Let K be a field of characteristic $\neq 2$ that contains neither $\sqrt{-1}$ nor $\sqrt{2}$* (e.g. $\mathbf{Q}$ *or a finite field* $\mathbf{F}_q$ *where* $q = p^n \equiv 3 \bmod 8$)*, and let the elliptic curve E be given by an equation of the form*

$$y^2 = x^3 + bx, \quad b \in K^*.$$

Then the 2-torsion subgroup of $E(K)$ (i.e. *points of order 2^n, some n*) *is as follows, where C_m denotes a cyclic group of order m :*

- $C_2 \oplus C_2$ *if $-b$ is a square*;
- C_4 *if $b = 4u^4$ for some $u \in K^*$*;
- C_2 *otherwise.*

Remark. It will be explained later (Proposition 4.2.2) that when char$K \neq 2$ the Weierstrass equations $y^2 = x^3 + bx$, $b \in K^*$, represent, up to isomorphism, precisely the curves with a point of order 2 and $j = 1728$. Actually $j = 1728$ guarantees a rational point of order 2 except when $\operatorname{char} K = 3$ (or $\operatorname{char} K = 2$). See also Corollary 7.2.1.

Proof. First suppose that $-b$ is not a square. Then $Q = (0, 0)$ is the only point of order 2 in $E(K)$, and therefore the 2-torsion subgroup is C_2 unless $[2]P = Q$ has a solution. By part (b) of the proposition, this is the case precisely when b has the form $4u^4$. Then $P = (2u^2, \pm 4u^3)$, and since $\sqrt{2} \notin K$, again by part (b), $[2]R = P$ has no solution, hence the 2-torsion subgroup is C_4.

When $-b$ is a square then $e_1 = 0$, $e_2 = \sqrt{-b}$ and $e_3 = -\sqrt{-b}$ are all in K. Now, $Q = (0, 0)$ cannot be halved since, by part (c), that would require also $\sqrt{b} \in K$, hence $\sqrt{-1} \in K$. Next, $Q' = (e_2, 0)$ cannot be halved since that requires $e_2 \in K^{*2}$ and $q_+ = 2e_2 \in K^{*2}$, hence $\sqrt{2} \in K$. Similarly $Q'' = (e_3, 0)$ cannot be halved. Thus the 2-torsion subgroup is $\{O, Q, Q', Q''\}$ which has the form $C_2 \oplus C_2$. ■

1.7.2 The division polynomials

Inductively one can obtain formulas for the coordinates of $[m](x, y)$ for every positive integer m.† The case $m = 2$ occurs in Proposition 1.7.3 above; however we will not take the trouble to carry through with a formula for the y-coordinate in characteristic 2 for general m.

From §1.1 we recall the definition

$$\kappa = 2y + a_1 x + a_3$$

and the relation

$$\kappa^2 = 4x^3 + b_2 x^2 + 2b_4 x + b_6 \qquad (\#)$$

which is valid in all characteristics. It follows that if $P \neq O$ then

$$[2]P = O \Longleftrightarrow \kappa(P) = 0.$$

Next we define the **division polynomials** inductively:

$$\begin{aligned}
\psi_0 &= 0, \qquad \psi_1 = 1, \qquad \psi_2 = \kappa, \\
\psi_3 &= 3x^4 + b_2 x^3 + 3b_4 x^2 + 3b_6 x + b_8, \\
\psi_4 &= \kappa\left(2x^6 + b_2 x^5 + 5b_4 x^4 + 10b_6 x^3 + 10b_8 x^2 + \right.
\end{aligned}$$

†We are treating (x, y) as a generic point on E: if $(a, b) \in E(K')$ for any field $K' \supset K$ then $x \mapsto a$, $y \mapsto b$ defines a K-algebra homomorphism $K[x, y] \longrightarrow K'$, and the formulas for $[m](x, y)$ can be specialized to evaluate $[m](a, b)$. But usually one can be more informal and refer to a 'general' point (x, y) with the understanding that the coordinates may have 'specific' values.

$$
\begin{aligned}
& (b_2b_8 - b_4b_6)x + b_4b_8 - b_6^2\big)\,, \\
\psi_{2m+1} &= \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3, \\
\psi_{2m} &= (\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2)\psi_m/\kappa.
\end{aligned}
$$

By taking $m = 2, 3, \ldots$ these recurrence relations define ψ_n for all $n \geq 0$. In [Cas49], Cassels defines

$$\psi_{-n} = -\psi_n$$

and it is easily seen that now the recurrences are valid $\forall m \in \mathbf{Z}$. In fact both recurrences are subsumed by

$$\psi_{m+n}\psi_{m-n} = \psi_{m-1}\psi_{m+1}\psi_n^2 - \psi_{n-1}\psi_{n+1}\psi_m^2 \quad \forall m, n \in \mathbf{Z}.$$

Cassels [*ibid.*] also introduces the following useful notation, not present in the classical texts, such as [Web08], vol.3, p.196; [Jor13], vol.3, p.190; [Fri22], vol.2, p.184 (and at the same time multiplies the old ψ_m by $(-1)^{m+1}$, giving our ψ_m, so that the leading coefficient is always positive):

$$\phi_m = x\psi_m^2 - \psi_{m-1}\psi_{m+1},$$

$$\text{hence} \qquad \phi_{-m} = \phi_m, \quad \phi_0 = 1;$$

$$\text{and when } \operatorname{char} K \neq 2, \qquad \omega_m = \psi_{2m}/2\psi_m \text{ for } m \neq 0, \text{ and } \omega_0 = 1,$$

$$\text{hence} \qquad \omega_{-m} = \omega_m.$$

Proposition 1.7.7 (a) *The functions* ψ_{odd}, $\psi_{\text{even}}/\kappa$, ϕ_{all}, ω_{even}, $\omega_{\text{odd}}/\kappa$ *are all polynomials in* x*; more precisely, they are in the ring*

$$\overline{\mathbf{Z}}[a_1, a_2, a_3, a_4, a_6, x]$$

where $\overline{\mathbf{Z}}$ *denotes* $\mathbf{Z}$ *modulo* char K *(so* $\overline{\mathbf{Z}}$ *is* $\mathbf{Z}$ *or* $\mathbf{F}_p$*). Their leading terms are*

$$
\begin{aligned}
\psi_m &= mx^{(m^2-1)/2} + \cdots, \quad \textit{(m odd)} \\
\psi_m/\kappa &= (m/2)x^{(m^2-4)/2} + \cdots, \quad \textit{(m even)} \\
\phi_m &= x^{m^2} + \cdots, \\
\omega_m &= x^{3m^2/2} + \cdots, \quad \textit{(m even)} \\
\omega_m/\kappa &= (1/2)x^{3(m^2-1)/2} + \cdots, \quad \textit{(m odd)}
\end{aligned}
$$

Thus for all m*,*

$$\psi_m^2 = m^2x^{m^2-1} + \cdots, \qquad \omega_m^2 = x^{3m^2} + \cdots.$$

(b) *If the* a_i *are independent transcendentals and if we assign the weights*

$$w(a_i) = i, \quad w(x) = 2, \quad w(y) = 3,$$

i.e., *we make* $B = \overline{\mathbf{Z}}[a_1, \ldots, a_6, x, y]$ *into a graded ring, then the functions* ψ_m, ϕ_m, ω_m *are homogeneous of weights* m^2-1, $2m^2$, $3m^2$ *respectively — when they are expressed as elements of* B, *all terms have equal weight.* (It is immaterial whether or not y^2 is replaced by $x^3+\cdots$ according to the Weierstrass equation.)

The easy proof by induction for ψ and then deduction for ϕ and ω, which is rather long when written out in detail, is left to the reader. ■

For $P \in E(K)$, $P \neq O$, we let $\psi_m(P)$, $\phi_m(P)$ and (when char $K \neq 2$) $\omega_m(P)$ denote the values of these functions when the coordinates of P are substituted for x and y in the above.

Proposition 1.7.8 (a) *Let* $P \in E(K)$, $P \neq O$ *and let* $m \in \mathbf{Z}$. *Then* $[m]P = O$ *iff* $\psi_m(P) = 0$, *and when this is not the case,*

$$x([m]P) = \phi_m(P)/\psi_m(P)^2;$$

this is a rational function of $x(P)$. *When* $\operatorname{char} K \neq 2$,

$$\eta([m]P) = \omega_m(P)/\psi_m(P)^3, \quad \textit{hence}$$

$$y([m]P) = \frac{\omega_m(P)}{\psi_m(P)^3} - \frac{1}{2}\left[a_1\frac{\phi_m(P)}{\psi_m(P)^2} + a_3\right].$$

(b) *For positive integers* m *and* n,

$$\begin{aligned}
\psi_{mn} &= \psi_n^{m^2}\psi_m([n](x,y)),\\
\phi_{mn} &= \psi_n^{2m^2}\phi_m([n](x,y)),\\
\omega_{mn} &= \psi_n^{3m^2}\omega_m([n](x,y)).
\end{aligned}$$

Proof. (a) A somewhat more challenging exercise for the computer.

(b) These three formulas are merely an elaboration of

$$[mn](x,y) = [m][n](x,y).$$

Note that the ω_m in [Aya92] is not the same as ours; nevertheless, both ω satisfy the third equation. ■

Example The map in Proposition 1.4.4 is essentially multiplication by 3; the details are as follows. Let K be a field of characteristic $\neq 2$ or 3, let a, b, c be nonzero elements of K and suppose $\theta := -\sqrt[3]{c/b} \in K^*$. Consider the curves

$$\begin{tikzcd}
C_1 \arrow[r, "\tau"] \arrow[d, "f_1"'] & C_2 \arrow[d, "f_2"] \\
E \arrow[r, "{[3]}"'] & E
\end{tikzcd}$$

$C_1:\ aU^3 + bV^3 + cW^3 = 0,$
$C_2:\ abcR^3 + S^3 + T^3 = 0,$
$E:\ Y^2Z = X^3 - 432a^2b^2c^2Z^3.$
By Corollary 1.4.3 there are bijections $f_i : C_i(K) \longrightarrow E(K)$. The map $(u, v, w) \mapsto (r, s, t)$

of Proposition 1.4.4, with obvious minor changes in notation, induces a map $\tau : C_1 \longrightarrow C_2$. Calculation shows that the accompanying diagram is commutative. Thus an alternative way to introduce τ is to make the definition $\tau := f_2^{-1}[3]f_1$.

In the following corollary we record the most immediate deductions from the proposition concerning division points. These results will be refined later; see the remarks in the next section.

C_m denotes a cyclic group of order m and $\overline{K}$ an algebraic closure of the field K. Since $[-m](P) = -[m](P)$, $E(K)[-m] = E(K)[m]$.

Corollary 1.7.9 *Let m be a positive integer.*

(a) $\psi_m \neq 0$*, in other words,* $[m] \neq [0]$ *as endomorphisms of the abelian group* $E(\overline{K})$.

(b) *The subgroup $E(K)[m]$ of $E(K)$ is isomorphic to a subgroup of $C_m \oplus C_m$. If* $\mathrm{char} K = p > 0$ *then $E(\overline{K})[p]$ is either* 0 *or C_p.*

(c) *If $P = (x, y) \in E(K)[m]$ then $mx \in \overline{\mathbf{Z}}[a_1, \ldots, a_6]$.*

Proof. (a) By Proposition 1.7.7(a), if $\psi_m = 0$ then $\mathrm{char}\, K = p > 0$ and $p|m$. Let q be any prime not dividing m and choose $P \in E(\overline{K})[q]$, $P \neq O$. Then $[m]P = [q]P = O$ where $\gcd(m, q) = 1$, say $sm + tq = 1$, which gives the contradiction $[1]P = O$.

(b) First let m be odd. $\psi_m(x)$ has at most $(m^2 - 1)/2$ distinct roots $x \in K$. For each such x there are at most 2 values of $y \in K$ satisfying the Weierstrass equation, and therefore $E(K)[m]$ contains at most $m^2 - 1$ nonzero points, hence at most m^2 points in all.

Similarly when m is even, $E(K)[m]$ contains O, at most three 2-division points and at most two points for each of the distinct roots of ψ_m/κ in K, hence at most $1 + 3 + 2(m^2 - 4)/2 = m^2$ points in all.

Since $[m]P = O$ for each $P \in E(K)[m]$, the first statement follows for both odd and even m.

Now let $\mathrm{char} K = p > 0$ so that $E(\overline{K})[p] = V$, say, is a vector space over $\mathbf{F}_p$. When $p = 2$, we saw in Proposition 1.7.3 that $\dim V = 0$ or 1. When p is odd, $\psi_p = px^{(p^2-1)/2} + \cdots$ has fewer than $(p^2 - 1)/2$ roots since $p = 0$ in K. This implies that $\dim V \leq 1$.

(c) reflects the fact that the coefficients of ψ_m are in $\overline{\mathbf{Z}}[a_1, \ldots, a_6]$ and the leading coefficient is m. ■

Definition: When $\mathrm{char} K = p > 0$, the elliptic curve E is **supersingular** resp. **ordinary** when $E(\overline{K})[p]$ is 0 resp. C_p. (The terminology is a trifle perverse: a supersingular E is not singular.)

Proposition 1.7.10 *Let E be an elliptic curve defined over the field K.*

- *If* $\mathrm{char}\, K = 2$ *then E is supersingular iff $a_1 = 0$ iff $j = 0$.*
- *If* $\mathrm{char}\, K = 3$ *then E is supersingular iff $b_2 = 0$ iff $c_4 = 0$ iff $j = 0$.*

Proof. When $\operatorname{char} K = 2$, then $b_2 = a_1^2$ and $b_6 = a_3^2$ are not both 0 in order that $\Delta \neq 0$. The statement follows from relation (#) at the beginning of this section, or from Proposition 1.7.3.

When $\operatorname{char} K = 3$, the statement follows from $\psi_3 = b_2 x^3 + b_8$ and $c_4 = b_2^2$. ■

Many equivalent conditions for supersingularity will emerge as the theory is developed. We will see, *e.g.*, that there are nonzero supersingular $j \in \mathbf{F}_p$ for all $p \geq 7$.

Lemma 1.7.11 *Let a_1, a_2, a_3, a_4, a_6 be independent transcendentals over $\mathbf{Q}$ and let A denote the* UFD $\mathbf{Z}[a_1, a_2, a_3, a_4, a_6]$. *Then*

(a) *Δ is an irreducible element of A;*

(b) *as polynomials in x over the quotient field of A, the* $\gcd(\phi_m, \psi_m^2) = 1$; *in fact*

$$\operatorname{Resultant}(\phi_m, \psi_m^2) = \pm \Delta^{m^2(m^2-1)/6}.$$

Remarks. (a) We will use this result only in characteristic 0, but we note that the proof is easily adapted to $A/pA = \mathbf{F}_p[a_1 \ldots]$ for all primes $p \geq 5$. Special proofs would be needed for $p = 2$ and 3, as is often the case in matters having to do with elliptic curves, in order to establish the irreducibility of the generic discriminant in these characteristics.

(b) In [Aya92] it is stated that the sign is always +, and I have no reason to doubt this. But I have found a reasonably simple proof only for m not divisible by 4. In any case we will not need this refinement.

(When m is odd the positiveness follows from $\operatorname{Res}(\phi_m, \psi_m^2) = \operatorname{Res}(\phi_m, \psi_m)^2$; but this no longer applies when m is even since then ψ_m contains the factor κ which is not a polynomial in x.)

Proof. The irreducible elements of A are the prime numbers p (and $-p$; note that $A^* = \{\pm 1\}$) and the irreducible polynomials with content (the gcd of the integer coefficients) = 1.

(a) We note that the factorization of Δ contains no p. For if $\Delta = p\Delta'$, this factor p would remain when we specialize the a_i to values in $\mathbf{Z}$; but a pair of examples such as **A24** with $\Delta = -2^4 3$ and **C17** with $\Delta = -17^4$ shows that there is no such common factor p. (We would have a problem if we restricted ourselves to a family such as $y^2 = x^3 + bx + c$ for which $\Delta = -16(4b^3 + 27c^2)$.)

The proof of (a) is completed by looking at Δ as a polynomial in a_6:

$$\Delta = -432a_6^2 + \alpha a_6 + \beta, \quad \text{where} \quad \alpha, \beta \in \mathbf{Z}[a_1, a_2, a_3, a_4].$$

Suppose there were a nontrivial factorization $\Delta = fg$. If the degree of f as a polynomial in a_6 is 0, then f could only an integer divisor of 432; but we have just seen that Δ has no integer factors. Thus f and g must both be of degree 1 which means that the above quadratic has roots in $\mathbf{Z}[a_1, \ldots, a_4]$. These roots

have the form, for a certain $\gamma \in \mathbf{Z}[a_1, \ldots, a_4]$,

$$a_6 = \gamma \pm \frac{1}{864}\delta^{3/2} \quad \text{where} \quad \delta = a_1^4 + 8a_1^2 a_2 + 16a_2^2 - 24a_1 a_3 - 48a_4.$$

But in fact the required square root does not exist in $\mathbf{Z}[a_1, \ldots, a_4]$ since δ is a linear polynomial in a_4.

(b) Let K denote the quotient field of A. Suppose p is an irreducible polynomial factor of ϕ_m and ψ_m^2 of positive degree in $K[x]$. The formula defining ϕ_m shows that p divides one of

$$\psi_{m-1}, \quad \text{or} \quad \psi_{m+1} \quad \text{when } m \text{ is even,}$$

$$\psi_{m-1}/\kappa, \quad \psi_{m+1}/\kappa, \quad \text{or} \quad \kappa^2 \quad \text{when } m \text{ is odd.}$$

Extend K to a field K' containing a root θ of p, and a point $P = (\theta, \beta) \in E(K')$. Since $\psi_m(P) = 0$, by the previous proposition P is a point of order dividing m. Hence neither $\psi_{m-1}(P)$ nor $\psi_{m+1}(P)$, nor (in the case of odd m) $\kappa^2(P) = \psi_2^2(P)$ can be 0, and we have a contradiction.

Since the vanishing of the resultant is a necessary and sufficient condition for nontrivial gcd, we now know at least that the resultant in question — call it R — is nonzero. Thus we must show that the only irreducible element dividing R is Δ, and that the exponent is right.

First, R cannot have a prime number factor p since that would make R identically 0 in characteristic p. Second, if π is a polynomial factor of R then we can specialize the a_i, extending $\mathbf{Q}$ to a number field if necessary, to make $\pi = 0$. Then $R = 0$ which implies that the Weierstrass equation cannot define an elliptic curve: Δ is now 0. Conclusion: $\pi = \pm\Delta$.

Finally, the exponent is right since R is a homogeneous polynomial of weight $2m^2(m^2 - 1)$ (each term in the determinant expansion has that weight, as can be seen by an inspection of the entries in R) and Δ is homogeneous of weight 12. ■

Now, a basic property of the resultant R of two polynomials f and g is that $sf - tg = R$ for some polynomials s, t with coefficients in the domain generated by the coefficients of f and g. Thus

$$s\psi_m^2 - t\phi_m = \Delta^{m^2(m^2-1)} \quad \text{for some} \quad s, t \in \mathbf{Z}[b_2, b_4, , b_6, b_8][x].$$

We wish to make two comments.

First, this equation will survive any specialization of the Weierstrass coefficients to specific values in a field K of any characteristic (although when char$K$ divides m the right side is no longer the resultant since the degree of ψ_m^2 is reduced). This observation gives part (a) of the following proposition.

Second, at least for the cases $m = 2$ and 3, the Euclidean algorithm gives a better exponent on Δ:

$$m = 2, 3: \ \sigma_m \psi_m^2 - \tau_m \phi_m = \Delta^{(m-1)^2}, \quad \sigma_m, \tau_m \in \overline{\mathbf{Z}}[b_2, b_4, b_6, b_8, x].$$

Actually, a direct application of the Euclidean algorithm gets out of hand already for $m = 3$; one can use the algorithm instead to find $\alpha, \beta \in \overline{\mathbf{Z}}[b_2, \ldots, x]$ such that $\alpha\psi_3 + \beta\phi_3 = \Delta^2$, and then square. We record the computer results for $m = 2$ in part (b); this identity will be a basic ingredient in a number of results: the generalized Nagell–Lutz theorem presented in §2.10, an estimate of Siksek in Chapter 7, and Olson's theorem in Chapter 8.

Proposition 1.7.12 *Let E be an elliptic curve over the field K (of any characteristic).*

(a) *As polynomials in x over K, the* $\gcd(\phi_m, \psi_m^2) = 1$*; in fact*

$$s\psi_m^2 - t\phi_m = \Delta^{m^2(m^2-1)} \quad \textit{for some} \quad s, t \in \overline{\mathbf{Z}}[b_2, b_4, b_6, b_8][x].$$

(b) *We have the polynomial identity*

$$\sigma_2\psi_2^2 - \tau_2\phi_2 = \Delta$$

where

$$\begin{aligned} \sigma_2 &= 12x^3 - b_2x^2 - 10b_4x + b_2b_4 - 27b_6, \\ \psi_2^2 &= 4x^3 + b_2x^2 + 2b_4x + b_6, \\ \tau_2 &= 48x^2 + 8b_2x + 32b_4 - b_2^2, \\ \phi_2 &= x^4 - b_4x^2 - 2b_6x - b_8. \end{aligned}$$

Equivalently, since $x([2](x, y)) = \phi_2/\psi_2^2$ and $\psi_2 = \kappa$,

$$(\sigma_2 - x([2](x, y))\tau_2)\kappa^2 = \Delta.$$

1.7.3 Remarks on the group of division points

In Chapter 6 we will prove the following statement:

> *Let E be an elliptic curve defined over the field K, let m be a positive integer and let K' be any overfield of K that contains the x and y coordinates of all the m-division points* (e.g. $K' = \overline{K}$)*. Suppose that* char K *is either* 0 *or a positive prime that does not divide m. Then*
>
> $$E(K')[m] \approx C_m \oplus C_m. \qquad (\%)$$

The proof will be an elegant application of basic properties of isogenies; but let us consider how a computational proof might go at this stage.

The case $m = 1$ is trivial and from earlier work we know the case $m = 2$ to be true. Thus we can assume $m \geq 3$.

Inspecting the proof of Corollary 1.7.9, the proof of the above statement would seem to be almost there: all we need to know (with the stated assumption on char K) is that ψ_m doesn't have a repeated root. For convenience define

$$\psi_m^* = \begin{cases} \psi_m & \text{if } m \text{ is odd,} \\ \psi_m/\kappa & \text{if } m \text{ is even,} \end{cases}$$

so that ψ_m^* is a polynomial in x for all m, and let d_m denote its discriminant. Thus ψ_m^* has a repeated root iff $d_m = 0$. Thus (%) would follow from the next statement, which I believe is true:

$$d_m = \begin{cases} \pm m' \Delta^{(m^2-1)(m^2-3)/24} & m \text{ odd}, \\ \pm m' \Delta^{(m^2-4)(m^2-6)/24} & m \text{ even}, \end{cases}$$

where m' is a positive integer whose prime divisors are precisely those of m.

On the computer one finds, for example, that

$$d_3 = -27\Delta^2.$$

But one can't go much further with a Weierstrass equation with general coefficients — the calculations take too long.

Now any torsion abelian group is the direct sum of its p-primary components. For example, let $\overline{\mathcal{T}}$ denote the torsion subgroup of $E(\overline{K})$ so that

$$\overline{\mathcal{T}} = \bigoplus_{p \text{ prime}} \overline{\mathcal{T}}_p$$

where
$$\overline{\mathcal{T}}_p = \bigcup_{n \geq 1} E(\overline{K})[p^n].$$

(In fancier terms, the natural inclusions $E(\overline{K})[p^n] \hookrightarrow E(\overline{K})[p^{n+1}]$ form a direct system and

$$\overline{\mathcal{T}}_p = \varinjlim E(\overline{K})[p^n].\Big)$$

Thus it is sufficient to prove the proposition when m is a prime power, and this would follow from (again I believe these statements to be true):

- *if $m = p^n$ is a power of the odd prime p, then*

$$d_m = (-1)^{(m-1)/2} m^{(m^2-3)/2} \Delta^{(m^2-1)(m^2-3)/24};$$

- *if $m = 2^n$ then*

$$d_m = -2^{\alpha_n} \Delta^{(m^2-4)(m^2-6)/24}$$

  where α_n is a positive integer ($\alpha_2 = 8$, $\alpha_3 = 82$, $\alpha_4 = 492$, $\alpha_5 = 2534$, "etc." — I do not know the values of any higher α).

1.8 The group law: singular case

We now consider a Weierstrass equation with $\Delta = 0$. Let $E_{\rm ns}(K)$ denote the set of nonsingular points, that is, the point O at ∞ together with all affine points (x, y) satisfying the equation except for the unique singular point as described in Proposition 1.5.4. We will prove that the same geometric construction used to define the group operations for nonsingular Weierstrass equations gives $E_{\rm ns}(K)$ a group structure. This is expected for reasons of "continuity": think of the real case $K = \mathbf{R}$ — a small change in the a_i makes $\Delta \neq 0$. The operations $-, +$ are defined by rational functions which are continuous (away from 0 denominators) in all the quantities involved; and the associative law, for instance, is identically satisfied. It is not surprising that the singular point must be excluded. But of course these heuristics must be replaced by algebraic proofs paying due regard to $\operatorname{char} K$.

The analysis of the singular case will have application to elliptic curves. For example the curve $y^2 = x(x+2)(x+6)$ defined over $\mathbf{Q}$ can be interpreted as a curve over $\mathbf{F}_p$ for each prime p. Over $\mathbf{F}_2$ it has a cusp; over $\mathbf{F}_3$ it has a node (with tangents defined only over $\mathbf{F}_9$); and since $\Delta = 2^{12}3^2$, over $\mathbf{F}_p$ for $p > 3$ it is an elliptic curve. This reduction mod p information conveys significant information about the original curve; indeed by the famous Birch, Swinnerton-Dyer conjecture which we will discuss later, it conveys a huge amount of information.

Three points in $E_{\rm ns}(K)$ are defined to be **colinear** if they are the three points of intersection of a line in $\mathbf{P}^2$ with $E_{\rm ns}(K)$; if two of the points are the same this means that the line is tangent there, and if all three points are the same this means that the point is a flex.

Proposition 1.8.1 *Let E be given by a singular Weierstrass equation over K.*

(a) *The formulas of Proposition 1.7.3 for $-$ and $+$ endow $E_{\rm ns}$ with the structure of an abelian group with O as* 0*-element.*

(b) *Suppose E has a node with tangents $x = x_0 + t,\ y = y_0 + \mu_i t,\ i = 1, 2$ as described in Proposition 1.5.4.*

(b1) *If the tangents are rational*, i.e., $\mu_i \in K$, *then*

$$\begin{cases} O & \longmapsto & 1 \\ (x, y) & \longmapsto & \dfrac{\mu_1(x - x_0) - (y - y_0)}{\mu_2(x - x_0) - (y - y_0)} \end{cases}$$

defines a group isomorphism $\phi : E_{\rm ns}(K) \longrightarrow K^*$. Hence three points $P_1, P_2, P_3 \in E_{\rm ns}(K)$ are colinear iff $\phi(P_1)\phi(P_2)\phi(P_3) = 1$.

(b2) *If the tangents are irrational let K_2 denote the quadratic extension $K(\mu_1, \mu_2)$, let N denote the norm homomorphism $K_2^* \longrightarrow K^*$ and let $\phi : E_{\rm ns}(K_2) \xrightarrow{\sim} K_2^*$ be the isomorphism of* (b1). *Then $\phi(K_{\rm ns}(K)) = \ker N$.*

(c) *Suppose E has a cusp ($c_4 = 0$) with unique tangent $x = x_0 + t$, $y = y_0 + \mu t$, as in Proposition 1.5.4, and let $K_2 = K(x_0, y_0, \mu)$ ($= K$ except in certain inseparable cases in characteristics 2 and 3). Then*

$$\begin{cases} O & \longmapsto & 0 \\ (x,y) & \longmapsto & \dfrac{x - x_0}{(y - y_0) - \mu(x - x_0)} \end{cases}$$

defines a group isomorphism† $\phi : E_{\mathrm{ns}}(K_2) \longrightarrow K_2^+$ *to the additive group. Hence three points $P_1, P_2, P_3 \in E_{\mathrm{ns}}(K)$ are colinear iff $\phi(P_1) + \phi(P_2) + \phi(P_3) = 0$.*

Proof. The geometric definition of negative is

$$-(x,y) = (x, -y - a_1 x - a_3).$$

In the multiplicative case ϕ sends this to

$$\frac{\mu_1(x - x_0) - (-y - a_1 x - a_3 - y_0)}{\mu_2(x - x_0) - (-y - a_1 x - a_3 - y_0)}.$$

We wish to show that this coincides with

$$(\phi(x,y))^{-1} = \frac{\mu_2(x - x_0) - (y - y_0)}{\mu_1(x - x_0) - (y - y_0)}.$$

This follows from the relations $\mu_1 + \mu_2 = -a_1$, $2y_0 + a_1 x_0 + a_3 = 0$ of Proposition 1.5.4.

Similarly in the additive case we find $\phi(-(x,y)) = -\phi((x,y))$ using $\mu = -a_1/2$ when $\operatorname{char} K \neq 2$, or $\Delta = c_4 = 0 \Rightarrow a_1 = a_3 = 0$ when $\operatorname{char} K = 2$.

Assuming that (b1) is proved we deduce (b2) as follows. Let σ denote the nontrivial element of $\mathrm{Gal}(K_2/K)$. If $P = (x,y) \in E_{\mathrm{ns}}(K_2)$ then $P^\sigma = (x^\sigma, y^\sigma) \in E_{\mathrm{ns}}(K_2)$ and since σ interchanges the μ's

$$\phi(P)^\sigma = \frac{\mu_2(x^\sigma - x_0) - (y^\sigma - y_0)}{\mu_1(x^\sigma - x_0) - (y^\sigma - y_0)} = \phi(P^\sigma)^{-1}.$$

Hence

$$\begin{aligned} P \in \ker N \quad &\Leftrightarrow \quad \phi(P)/\phi(P^\sigma) = 1 \\ &\Leftrightarrow \quad \phi(P) = \phi(P^\sigma) \\ &\Leftrightarrow \quad P = P^\sigma \quad \text{since } \phi \text{ is injective} \\ &\Leftrightarrow \quad P \in E_{\mathrm{ns}}(K). \end{aligned}$$

Since ϕ is surjective it follows that $\ker N = E_{\mathrm{ns}}(K)$.

†with the sign chosen the same way as in [Sil86]; in the multiplicative case the switch from ϕ to ϕ^{-1} comes from interchanging the μ's.

We now assume that $K_2 = K$ in the two cases and all will follow if we prove that ϕ is bijective and three points P_i are colinear iff the product (resp. sum) of the $\phi(P_i)$ is 1 (resp. 0).

In the multiplicative case we make the linear projective transformation $(X, Y, Z) \mapsto (X', Y', Z')$ where

$$\begin{aligned} X' &= \mu_1(X - x_0 Z) - (Y - y_0 Z), \\ Y' &= (\mu_1 - \mu_2)^3 Z, \\ Z' &= \mu_2(X - x_0 Z) - (Y - y_0 Z). \end{aligned}$$

This transformation[‡] is invertible since its determinant $= (\mu_1 - \mu_2)^4 \neq 0$. Using the various relations explained in Proposition 1.5.4, calculation shows that

$$X'Y'Z' = (X' - Z')^3.$$

The singular point $(X, Y, Z) = (x_0, y_0, 1)$ in the new coordinates is $(0, 1, 0)$, the unique point at infinity on the curve. Thus lines not passing through this point are described in terms of affine coordinates $x' = X'/Z', y' = Y'/Z'$ by equations of the form $y' = ax' + b$. The transformation and its inverse are described affinely as follows:

$$x' = \phi(x, y), \qquad y' = \frac{(\mu_1 - \mu_2)^3}{\mu_2(x - x_0) - (y - y_0)},$$

$$\begin{aligned} x &= x_0 + (\mu_1 - \mu_2)^2 (x' - 1)/y' \\ y &= y_0 + (\mu_1 - \mu_2)^2 (\mu_2 x' - \mu_1)/y', \end{aligned}$$

and we have bijections

$$P = (x, y) \longleftrightarrow (x', y') = (x', (x' - 1)^3/x') \longleftrightarrow x' = \phi(P).$$

The line $y' = ax' + b$ meets the curve in the three points whose x'-coordinates are the three roots x'_1, x'_2, x'_3 of the cubic $(x'-1)^3 - x'(ax'+b) = 0$. The constant term shows that $x'_1 x'_2 x'_3 = 1$. Conversely if three points $P_i = (x'_i, y'_i) \in E_{\mathrm{ns}}(K)$ are such that $x'_1 x'_2 x'_3 = 1$ then these x'_i are the roots of the above cubic where we define $a = x'_1 + x'_2 + x'_3 - 3$ and $b = 3 - (x'_1 x'_2 + x'_1 x'_3 + x'_2 x'_3)$, and consequently the points lie on the line $y' = ax' + b$.

In the additive case μ has just one value and we make the transformation

$$\begin{aligned} X' &= X - x_0 Z, \\ Y' &= Z, \\ Z' &= Y - y_0 Z - \mu(X - x_0 Z); \end{aligned}$$

[‡]For a step by step explanation of how one is led to this transformation see [Sil86], p.61.

the determinant is -1, so the transformation is invertible. Again the line $Z' = 0$ meets the curve only in the singular point. If we dehomogenize by $x' = X'/Z', y' = Y'/Z'$, the equation is

$$y' = x'^3$$

and we have bijections

$$P = (x, y) \longleftrightarrow (x', y') = (x', x'^3) \longleftrightarrow x' = \phi(P).$$

Three point $P_i \in E_{\mathrm{ns}}(K)$ are colinear iff their x'-coordinates x'_i are the three roots of $x'^3 - (ax' + b) = 0$ for some $a, b \in K$. Since the x'^2 term is absent, this is the case iff $x'_1 + x'_2 + x'_3 = 0$. ■

It is natural to adopt the definitions:

A singular Weierstrass equation is called **split multiplicative**, **nonsplit multiplicative** or **additive** according as the conditions of paragraph (b1), (b2) or (c) of the foregoing proposition are met. By the final statement in Proposition 1.5.4, the only possible change in the 'type' of a singular curve is from non-split multiplicative to split multiplicative.

1.8.1 Examples over finite fields

An example of the previous proposition that will be important when we study the number theory of elliptic curves is the case of finite fields:

Corollary 1.8.2 *Let E be given by a singular Weierstrass equation over the finite field $\mathbf{F}_q$. Then*

$$|E_{\mathrm{ns}}(\mathbf{F}_q)| = \begin{cases} q-1 & \textit{if } E \textit{ is split multiplicative,} \\ q+1 & \textit{if } E \textit{ is nonsplit multiplicative,} \\ q & \textit{if } E \textit{ is additive.} \end{cases}$$

The nonsplit case follows from the fact that the norm map $N : \mathbf{F}_{q^2}^* \longrightarrow \mathbf{F}_q^*$ is surjective ([Ire-Ro82], p.159) hence $|\ker N| = (q^2 - 1)/(q - 1) = q + 1$.

Example: We can interpret

$$y^2 + xy = x^3 - 18x + 27 \qquad (\mathbf{D_1 525})$$

as a curve over any $\mathbf{F}_q$. As a curve over $\mathbf{Q}$ it has covariants and invariant

$$c_4 = 5 * 173, \quad c_6 = -5^3 197, \quad \Delta = 3^3 5^3 7, \quad j = \frac{173^3}{3^3 7}.$$

Thus the curve is singular mod p for $p = 3, 5, 7$ and nonsingular mod p for all other p. The following results were obtained by **aPecs**. The notation $(x, y; n)$ stands for a point (x, y) of order n, and C_n denotes a cyclic group of order n. We denote the finite abelian group $E_{\mathrm{ns}}(\mathbf{F}_p)$ by G.

p	singular point	nonsingular points	G	$\lvert G\rvert$
2		$O, (1,1;4), (0,1;2), (1,0;4)$	C_4	4
3	$(2,0)$ split node	$O, (2,2;2)$	C_2	2
5	$(2,4)$ cusp	$O, (1,4;5), (3,0;5),$ $(3,2;5), (1,0;5)$	C_5	5
7	$(5,1)$ nonsplit node	$O, (2,6;2), (6,2;4), (4,2;8),$ $(4,1;8), (6,6;4), (3,0;8), (3,4;8)$	C_8	8
11		$O, (9,0;18), \ldots$	C_{18}	18
13		$O, (0,1;8), (2,12;2), \ldots$	$C_2 \times C_8$	16

Since additive and split multiplicative types remain the same in extensions, we can predict, for example, that

$$|E(\mathbf{F}_{3^2})| = 8, \quad |E(\mathbf{F}_{5^2})| = 25$$

without determining the actual points. We will see later how to determine $|E(\mathbf{F}_{p^n})|$ on the basis of $|E(\mathbf{F}_p)|$ for all p and n.

We conclude with a tabulation of the $2^5 = 32$ Weierstrass equations over $\mathbf{F}_2$. It happens that in all 32 cases $E_{\mathrm{ns}}(\mathbf{F}_2)$ is cyclic. (Not all $E_{\mathrm{ns}}(\mathbf{F}_3)$ are cyclic.) In the nonsingular cases, $\Delta = 1$ and $j = c_4 = 0$ or 1. Other special relations implied by $2 = 0$ and $a^2 = a$ are $b_2 = c_4 = c_6 = a_1$.

Additive cases ($\Delta = 0$, $c_4 = 0$):

a_1	a_2	a_3	a_4	a_6	b_2	b_4	b_6	b_8	c_4	c_6	Δ	$\lvert E_{\mathrm{ns}}(\mathbf{F}_2)\rvert$
0	0	0	0	0	0	0	0	0	0	0	0	2
0	0	0	0	1	0	0	0	0	0	0	0	2
0	0	0	1	0	0	0	0	1	0	0	0	2
0	0	0	1	1	0	0	0	1	0	0	0	2
0	1	0	0	0	0	0	0	0	0	0	0	2
0	1	0	0	1	0	0	0	0	0	0	0	2
0	1	0	1	0	0	0	0	1	0	0	0	2
0	1	0	1	1	0	0	0	1	0	0	0	2

Split multiplicative cases ($\Delta = 0$, $c_4 = 1$, $a_2 = a_3$):

1	0	0	0	0	1	0	0	0	1	1	0	1
1	0	0	1	1	1	0	0	0	1	1	0	1
1	1	1	0	1	1	1	1	0	1	1	0	1
1	1	1	1	1	1	1	1	0	1	1	0	1

Nonsplit multiplicative cases ($\Delta = 0$, $c_4 = 1$, $a_2 \neq a_3$):

1	0	1	0	0	1	1	1	0	1	1	0	3
1	0	1	1	0	1	1	1	0	1	1	0	3
1	1	0	0	0	1	0	0	0	1	1	0	3
1	1	0	1	1	1	0	0	0	1	1	0	3

Supersingular cases ($\Delta = 1$, $a_1 = 0$):

0	0	1	0	0	0	0	1	0	0	0	1	3
0	0	1	0	1	0	0	1	0	0	0	1	3
0	0	1	1	0	0	0	1	1	0	0	1	5
0	0	1	1	1	0	0	1	1	0	0	1	1
0	1	1	0	0	0	0	1	1	0	0	1	5
0	1	1	0	1	0	0	1	1	0	0	1	1
0	1	1	1	0	0	0	1	0	0	0	1	3
0	1	1	1	1	0	0	1	0	0	0	1	3

Ordinary cases ($\Delta = 1$, $a_1 = 1$):

1	0	0	0	1	1	0	0	1	1	1	1	4
1	0	0	1	0	1	0	0	1	1	1	1	4
1	0	1	0	1	1	1	1	1	1	1	1	2
1	0	1	1	1	1	1	1	1	1	1	1	2
1	1	0	0	1	1	0	0	1	1	1	1	2
1	1	0	1	0	1	0	0	1	1	1	1	2
1	1	1	0	0	1	1	1	1	1	1	1	4
1	1	1	1	0	1	1	1	1	1	1	1	4

Student project

Let $F \in \mathbf{Z}[x, y, \ldots]$ be a polynomial in several variables with integer coefficients, let p be a prime, and for $n = 1, 2, \ldots$ let ν_n denote the number of solutions $(x, y, \ldots)$ of the congruence

$$F \equiv 0 \bmod p^n.$$

I believe it was Igusa who proved that the power series

$$f = \nu_1 T + \nu_2 T^2 + \cdots$$

is a rational function of T.

Elaborate Igusa's result in the case

$$F = y^2 + a_1 xy + a_3 y - x^3 - a_2 x^2 - a_4 x - a_6, \quad a_i \in \mathbf{Z}.$$

For example, when $F = y^2 - (x^3 + 2x - 28)$ and $p = 5$, verify that

$$f = \frac{6T + 20T^2 + 25T^3}{1 - 5T}.$$

Appendix: introduction to apecs

apecs (arithmetic of plane elliptic curves) is a program written in Maple available at an anonymous ftp site. Here is the procedure to obtain a copy:

	ftp math.mcgill.ca	(internet 132.206.150.3)
< login >	anonymous	
< password >	your e-mail address	
	cd pub/apecs	
	get README	

The file README contains all the info. on how to get **apecs** (and the ubasic version *upecs*).
If there's a problem please let me know:

connell@math.mcgill.ca (internet 132.206.150.3)
or
Ian Connell
Mathematics Dept.
McGill University
805 Sherbrooke W.
Montreal, Quebec
Canada H3A 2K6

We describe some of the **apecs** procedures (or commands) that relate to the topics discussed in this chapter. The descriptions will be simplified and rather terse; for more information, in an **apecs** session use the commands `Nota();` and `menu();`, and for information on a particular command *xxx* use `Menu(`*xxx*`);`.

We present the descriptions here in a series of Problems followed by solutions in the format `input in typewriter font` $\Longrightarrow$ *output, normally to the computer screen*, where $\Longrightarrow$ is an abbreviation for 'results in the output'.

Problem 1. (p.102) Calculate $b_2, \ldots, c_6, \Delta, j$ for $y^2+2xy-3y = x^3+4x^2-5x+6$.
Solution. `Ell(2,4,-3,-5,6);` $\Longrightarrow$

$$b's = 20, -16, 33, 101$$

$$c's = 784, -26648$$

$$DD = -132075 = -3^2 5^2 587$$

$$jay = -481890304/132075, \quad \text{denom}(jay) = 3^2 5^2 587$$

PROBLEM 2. (p.110,141) Check that the point $Q = (3, 6)$ is on this curve and factor the quartic polynomial $\mathrm{Quar}_Q(m)$.
Solution. `x:=3:y:=6:on(3,6),factor(QuarQ);` $\Longrightarrow$

$$true, M(-60 - 28M + M^3)$$

PROBLEM 3. (p.102) As in PROBLEM 1 for $y^2 + \sqrt{2}y = x^3 - 1/2x^2 + 1/t$ where t is an indeterminate (transcendental).
Solution. `t:='t':Ell(0,-1/2,sqrt(2),0,1/t);` $\Longrightarrow$

$$b's = -2, 0, 2\frac{t+2}{t}, -\frac{t+2}{t} \quad \text{(etc.)}$$

PROBLEM 4. (p.102) What are Δ and j when $t = 2/3$?
Solution. `t:=2/3:DD,jay;` $\Longrightarrow$

$$-1712, -4/107$$

PROBLEM 5. (p.129) For what t is this curve singular?
Solution. `t:='t':solve(DD,t);` $\Longrightarrow$

$$-\frac{27}{13}, -2$$

PROBLEM 6. (p.104) Find an E with $j = 27$.
Solution. `Genj(27);` $\Longrightarrow$

several lines of data and then $\qquad A1323 = [1, -1, 1, 64, -1592]$

This procedure first calculates the "generic j" curve and then — since the argument 27 is rational — converts that Weierstrass equation to **Z**-minimal form by the *Laska-Kraus algorithm* (see §5.6.1):

$$y^2 + xy + y = x^3 - x^2 + 64x - 1592$$

and adds the curve to the **aPecs** catalog with name A1323 (1323 is the *conductor*; this term will be defined in a later chapter.)

PROBLEM 7. (p.104) Calculate the real root(s) of the cubic right side of $y^2 =$

$x^3 - 3x + 3$.
Solution. `ell(0,0,0,-3,3):Raf();` $\Longrightarrow$

$$\textit{x-coord.'s of real point(s) of order } 2:$$

$$raF = [-2.103803\ldots]$$

Capitalized commands such as `Ell` are "verbose", whereas their uncapitalized companions output only bottom line or abbreviated results; in this case `ell` has in fact no output.

PROBLEM 8. (p.102) For $y^2 + 2xy + 4y = x^3 - 6x^2$ display the cubic right side of $\eta^2 = x^3 + (b_2/4)x^2 + (b_4/2)x + b_6/4$ and its real root(s).
Solution. `ell(2,-6,4,0,0):x:='x':prac,raf();` $\Longrightarrow$

$$x^3 - 5x^2 + 4x + 4, \quad [-.56\ldots,\ \ 2,\ \ 3.56\ldots]$$

PROBLEM 9. (p.105) Put $v^2 = 5u^4 + 4$ in Weierstrass form, taking $(u, v) = (0, 2)$ as O, and find the point on the Weier. curve that (1,3) maps to.
Solution. `Quar(5,0,0,0,4,0,2);Trcw(1,3);` $\Longrightarrow$

$$\text{a few lines, then} \qquad \textit{curve is } A800 = [0, 0, 0, -5, 0]$$

$$U = 2X/Y, \qquad V = \frac{-2Y^2 + 4X^3}{Y^2}$$

$$X = \frac{V + 2}{U^2}, \qquad Y = \frac{2V + 4}{U^3}$$

$$[1, 3] \longmapsto [5, 10]$$

`Quar(a,b,c,d,e,f,g);` finds the Weierstrass equation of $V^2 = aU^4 + bU^3 + cU^2 + dU + e$ with point $(U, V) = (f, g)$ (*e.g.,* when $e = q^2$, one can take $f = 0$, $g = q$) serving as O according to Proposition 1.1.1 and then converts to minimal form, in the same manner as `Genj` in PROBLEM 6. If only 5 arguments are present `Quar` attempts to find a rational point, possibly at infinity.

The formulas of the birational correspondence between the quartic and the Weierstrass equation are displayed (when `Quar` is capitalized), and `Trcw` evaluates X and Y for the given $[U, V] = [1, 3]$. The map in the opposite direction is `Trwc(5,10)` $\Longrightarrow$ $[5, 10] \longmapsto [1, 3]$. (The letters `cw` and `wc` stand for 'curve to weierstrass' and vice versa.)

PROBLEM 10. (p.105) (Continuing the previous problem), assign to the variable `Pt` the image of (1,3) under the birational map from $v^2 = 5u^4 + 4$ to $y^2 = x^3 - 80x$.
Solution. `ein(0,0,0,-80,0):Pt:=trwc(trcw([1,3],1),2);` $\Longrightarrow$

$$\texttt{Pt} = [20, 80]$$

`ein` (the letters stand for E input) takes 5 *rational* arguments and, unlike `ell`, finds the minimal Weierstrass equation — in this case $A800$ — and does a lot more, including keeping track of E and the transformation equations between E and $A800$.

We now have two curves birat. equiv. to $A800$ — the quartic Q of the previous problem and the (non-$\mathbf{Z}$-minimal) curve $E : y^2 = x^3 - 80x$, and they are distinguished by the code numbers 1 and 2 respectively. We obtain the map $Q \to E$ as the composition $Q \to A800 \to E$. Thus `trcw([1,3],1)` returns the value $[5, 10]$ which is used as the first argument in the `trwc` command, and the latter returns the value $[20, 80]$. (`bec;bic;` $\Longrightarrow$ a display of these two variables which contain, respectively, the equations of Q and E and the details of the birational transformations.)

As a general rule, uncapitalized commands return the expected value, while capitalized commands display the result(s) and return NULL; see Menu(xxx); for the details about command xxx.

The point (or points) at infinity are denoted $[\,]$. `Trcw([ ])` and `Trwc([ ])` are valid.

Problem 11. (p.129) Find the location and nature of the singularity of $y^2 + y = x^3 - 3/4x$
Solution. `Ein(0,0,1,-3/4,0);` $\Longrightarrow$

$$\textit{curve is singular, } b's = \text{ (etc.)}$$

$$\textit{The singularity is at } [1/2, -1/2] \textit{ and is a node}$$

$$\textit{Tangent slopes there } = \sqrt{6}\Big/2, \quad -\sqrt{6}\Big/2$$

In the singular case there is no minimal Weierstrass equation and so no `Trwc` to worry about.

Problem 12. (p.–) We wish to recall the Weierstrass equation of the present curve.
Solution. `We();` $\Longrightarrow$

Apecs is not pointing to a catalog curve ...

$$Y^2 + Y = X^3 - 3/4X$$

More compactly, `we();` $\Longrightarrow$ $[0, 0, 1, -3/4, 0]$. Only minimal Weierstrass curves (and no elliptic curves introduced by `Ell`/`ell`) are entered into the **apecs** catalog and given a catalog name, *e.g.* `cur=`$A800$, and a catalog number denoted `ncur`. On the other hand, all curves, however introduced to apecs, are added to the *stack* = list of such curves considered in the present **apecs** session. Incidentally, `Go();` displays the contents of the stack and `Go(n);` transfers attention back to curve number n in the stack: **apecs** is now 'pointing to' that

curve.

PROBLEM 13. (p.115) Transform to minimal Weierstrass form using $(u, v) = (-10, 11)$ as O:

$$u^2v - 2uv^2 - 3v^3 - 4u^2 + 5uv - 6v^2 + 7u - 8v + 2307 = 0$$

Solution. `gcub(0,1,-2,-3,-4,5,-6,7,-8,2307,-10,11):we();` $\Longrightarrow$

$$[1, 1, 1, -122401, 24417674]$$

It is also possible to specify O as a point at infinity by homogeneous coordinates in the form $U, V, 0$; then `gcub` is given a total of 13 arguments. On the other hand if only 10 arguments are given then `gcub` attempts to find a rational point (possibly at infinity). Again `Trcw`, `Trwc`, `bec` and `bic` are available.

PROBLEM 14. (p.138) Verify the addition of points on $y^2 = x^3 + tx^2 - tx$ given in example 1, p.138.
Solution. `t:='t':ell(0,t,0,-t,0):P0:=[0,0]:P1:=[1,1]:Eadd(P0,P1);` `Mult(2,P0);Mult(2,P1);` $\Longrightarrow$

$$[0, 0] + [1, 1] = [-t, t]$$

$$[2][0, 0] = [\,]$$

$$[2][1, 1] = [1/4(t + 1)^2, -1/8(t + 1)(t^2 + 4t - 1)]$$

[ ] denotes O in **aPecs**.

Three other basic arithmetic operations that should be mentioned:

`Neg(P1);` (equivalently `Mult(-1,P1);`) $\Longrightarrow -P_1$;

`Sub(P1,P2);` (equivalently `Eadd(P1,neg(P2));`) $\Longrightarrow P_1 - P_2$;

`Comb(n1,P1,n2,P2,...);` $\Longrightarrow n_1P_1 + n_2P_2 + \cdots$ (any $n_i \in \mathbf{Z}$, any number of points).

The last procedure has a kind of inverse: `Lin(P1,P2,...)` finds the least i for which there exist integers $n_1, n_2, \ldots, n_i$ with $n_i \neq 0$ such that $n_1P_1 + \cdots + n_iP_i =$ a torsion point, *i.e.,* a point of finite order, possibly O.

PROBLEM 15. (p.–) On the Buhler-Gross-Zagier curve $y^2 + y = x^3 - 7x + 6$ ([BGZ85]), test [0,2], [1,0], [26/25,-101/125] for linear dependence mod torsion. (Actually the torsion subgroup is trivial as we see by the command `Tor`; see the next two examples.)
Solution. `Lin([0,2], [1,0], [26/25,-101/125]);` $\Longrightarrow$

$$\textit{Grammian height-pairing det } = 1.66857\ldots\ .$$

The non-vanishing of *det* implies the independence of the points.

Extending this example,
`Lin([0,2], [1,0], [26/25,-101/125], [2,0]);` $\Longrightarrow$
The points satisfy the relation

$$0[0,2] - [1,0] + [26/25, -101/125] - [2][2,0] = O.$$

See also `Menu(Expr);`. The subject of the (canonical or Néron-Tate) *height* of points will be treated in Chapter 3.

Problem 16. (p.136) Find all the rational points of finite order on $F15$.
Solution. `ein(F15);PP;` $\Longrightarrow$

$$[2,6,8], [7,-29,4], [32,171,8], [3/4,-7/8,2], [32,-204,8], [7,21,4], [2,-9,8]$$

`[a,b,n]` is alternative notation for a point `[a,b]` of finite order `n` — **apecs** understands both forms. **apecs** comes from the factory with all the [AntIV] curves from $A11$ to $F126$ plus some more it needs to know about (involving isogenies or complex multiplication — topics to be discussed in later chapters) in its starter catalog. It was originally intended to go all the way to $J200$ and perhaps even further following Cremona's tables [Cre92], but it was felt that the catalog was becoming too bulky. Each of these catalog entries contains the sequence `PP` of points of finite order (without O), along with other information, and so is available as soon as **apecs** is pointed to the curve.

For a new curve it is necessary to calculate `PP` — `ein` or the other input commands do not do this automatically.

Problem 17. (p.132, [Cre92]) Find all points of finite order on B_1354.
Solution. `ein(1,0,1,9,-8):Tor();` $\Longrightarrow$

(various comments as the calculation progresses, then)

Actually $NN = 3$ *and the points are*

$$O, [2,3,3], [2,-6,3]$$

The method of determining these points depends on an improved Nagell-Lutz theorem to be discussed in §2.10.1.

Problem 18. (p.140) Find all solutions of $[2]P = [4,-10]$ on $A65 : y^2 + xy = x^3 - x$.
Solution. `ein(A65):half(4,-10);` $\Longrightarrow$

$$[[1,0], [-1,1]]$$

If any ψ_n (p.145) are needed they can be calculated by the command `rec` — see `Menu(rec);`.

Finally we mention a few of the procedures for elliptic curves over $\mathbf{F}_p$.

PROBLEM 19. (p.156) Taking the coefficients of $C20 : y^2 = x^3 + x^2 - 41x - 116 \bmod p$, find $|E_{\rm ns}(\mathbf{F}_p)|$ for the first few primes; in particular, verify Corollary 1.8.2 for $p = 2$ and 5.
Solution. `ein(C20):Emods(2,17);` $\Longrightarrow$

p	N_p	a_p	θ	$\Delta \bmod p$	$j \bmod p$	comments
2	3	0	90.00	0	$--$	singular—additive
3	6	-2	54.74	2	2	anomalous
5	7	-1	77.08	0	$--$	singular—nonsplit multiplicative
7	6	2	112.2	5	4	
11	12	0	90.00	9	1	supersingular
13	12	2	106.1	11	4	
17	24	-6	43.31	11	1	

$N_p = p + 1 - a_p$ is the total number of points on $E \bmod p$ including the singular point if there is one. Thus $N_p = |E_{\rm ns}(\mathbf{F}_p)| + 1$ in the singular cases $p = 2, 5$. One writes $a_p = \cos\theta_p$; that this makes sense depends on Hasse's theorem, the 'Riemann Hypothesis for $E \bmod p$' — the explanation of this and the other undefined terms will come in due course.

PROBLEM 20. (p.156) Find the group $E \bmod p$ for $p = 2, 3, 5, 11$
Solution. `Allp(2);Allp(3);Allp(5);Allp(11)` $\Longrightarrow$

$[1, 1]$ *is a cusp on* $C20 \bmod 2$

group of non-singular points on $C20 \bmod 2 = O$, [0, 0, 2]

group order = 2 , *type: cyclic*

group of points on $C20 \bmod 3 = O$, [0, 2, 3], [1, 1, 6], [0, 1, 3], [1, 2, 6], [2, 0, 2]
group order = 6 , *type: cyclic*

$[4, 0]$ *is a node on* $C20 \bmod 5$

group of non-singular points on $C20 \bmod 5 = O$, [0, 2, 3], [0, 3, 3],

[1, 0, 2], [2, 2, 6], [2, 3, 6]

group order = 6 , *type: cyclic*

group of points on $C20 \bmod 11 = O$, [0, 4, 3], [0, 7, 3],

[6, 0, 2], [2, 1, 6], [2, 10, 6], [4, 3, 6], [7, 0, 2],

[4, 8, 6], [5, 4, 6], [8, 0, 2], [5, 7, 6]

group order = 12 , *type: non-cyclic* 6×2

There are also commands for doing arithmetic mod p: `Eadp` and `Mulp` are the analogs of `Eadd` and `Mult`. For these commands p must be preset. *E.g.* referring to the last example ($C20 \bmod 11$),
`p:=11:Mulp(2,[2,1,6]);` $\Longrightarrow [0, 7]$.

Chapter 2

Formal Groups

This chapter will involve calculations with formal power series. We summarize here the basic notation.

Let A be a commutative ring. Then $A[[T]]$ denotes the **ring of formal power series**, a typical element being $\alpha = a_0 + a_1T + a_2T^2 + \cdots$ where the $a_i \in A$ and T is an indeterminate (transcendental) over A. The ring operations are as follows: if $\beta = \sum b_iT^i$ is also in $A[[T]]$ then $\alpha + \beta = \sum(a_i + b_i)T^i$ and $\alpha\beta = \sum c_iT^i$ where $c_i = \sum_{j=0}^{i} a_jb_{i-j}$.

$A((T))$ denotes the localization $S^{-1}A[[T]]$ where $S = \{1, T, T^2, \ldots\}$ is the multiplicative set generated by T. The elements of $A((T))$ have the form $\sum_{i=N}^{\infty} a_iT^i$, for some $N \in \mathbf{Z}$, and are called **formal Laurent series**.

If A is an integral domain then, obviously, so is $A[[T]]$. If K is the quotient field of A, then the quotient field of $A[[T]]$ is $K((T))$.

This construction can be iterated. Because of the A-algebra isomorphism $f : A[[T_1]][[T_2]] \xrightarrow{\sim} A[[T_2]][[T_1]]$ determined by $f(T_1) = T_1$ and $f(T_2) = T_2$, these rings are identified and denoted $A[[T_1, T_2]]$. (Of course there is another isomorphism which switches T_1 and T_2, but that does not give the desired identification.) By induction, $A[[T_1, \ldots, T_n]]$ can be regarded as $B[[T_i]]$ for any i between 1 and n, where B is the power series ring in the set of variables $T_1, \ldots, T_n$ with T_i omitted.

The pitfall to avoid here is that the isomorphic rings $A((T_1))((T_2))$ and $A((T_2))((T_1))$ cannot be identified in the same way as $A[[T_1]][[T_2]] = A[[T_2]][[T_1]]$ since, for example, $\sum_{i=0}^{\infty} T_1^{-i}T_2^i$ is a member of the first ring, but not the second.

2.1 Discrete valuations

Let V be an integral domain with quotient field K. Then V is a **generalized valuation ring** if it has the property

(V1) $$x \in K,\ x \notin V \implies 1/x \in V.$$

This axiom implies that the **value group** $\Gamma := K^*/V^*$ is a totally ordered abelian group with the definition $xV^* \geq yV^*$ iff $x/y \in V$. It is customary to write this group additively, and to adjoin to it the symbol ∞ with the properties that for all $\gamma \in \Gamma \cup \{\infty\}$, $\infty \geq \gamma$ and $\infty + \gamma = \gamma + \infty = \infty$. Then one defines the **generalized valuation**

$$v : K \longrightarrow \Gamma \cup \{\infty\} \quad \text{where} \quad v(x) = xV^* \ \text{ for } \ x \in K^*, \ \text{ and } \ v(0) = \infty.$$

It follows easily from (V1) that $\forall\, x, y \in K$,
(V1a) $v(xy) = v(x) + v(y)$, and
(V1b) $v(x + y) \geq \min\{(v(x), v(y)\}$.
Conversely, if Γ is a totally ordered abelian group and $v : K^* \longrightarrow \Gamma$ a surjective map of the nonzero elements of a field K satisfying (V1a) and (V1b), then

$$V = \{x \in K : v(x) \geq 0\}$$

is a generalized valuation ring with quotient field K.

Proposition 2.1.1 (i) $v(x_1 + \cdots + x_n) \geq \min\{v(x_1), \ldots, v(x_n)\}$ *with equality if the minimum is attained by just one of the* $v(x_i)$*; hence for* $n \geq 2$*,*

$$x_1 + \cdots + x_n = 0 \Longrightarrow \textit{ at least two } x_i \textit{ have the minimum value.}$$

(ii) $V = \{x \in K : v(x) \geq 0\}$ *is a local ring with maximal ideal* $M = \{x \in V : v(x) > 0\}$ *and group of units (invertible elements)* $V^* = V - M = \{x \in V : v(x) = 0\}$.

(iii) V *is integrally closed: for* $x \in K$ *and* $a_i \in V$

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0 \Longrightarrow a_n x \in V;$$

Proof. (i) and (ii) are simple deductions from (V1a) and (V1b).

To prove (iii), suppose $v(a_n x) < 0$. Then $v((a_n x)^n) < v(a_i a_n^{n-1-i}(a_n x)^i)$ for $i = 0, 1, \ldots, n-1$, which precludes

$$(a_n x)^n + a_{n-1}(a_n x)^{n-1} + \cdots + a_0 a_n^{n-1} = 0$$

by (i). ■

$k = V/M$ is the **residue field** of v. Associated to v is the canonical surjective ring homomorphism

$$P : V \longrightarrow k$$

called the (generalized) **place** of v. This is extended to

$$P : K \longrightarrow k \cup \{\infty\}$$

by defining $P(x) = \infty$ for $x \notin V$, and then $P(1/x) = 0$. It is easy to check that any one of v, V, P determines the other two uniquely.

We quote without proof

Proposition 2.1.2 *Let A be an integral domain contained in the field L.*

(a) *If P is a prime ideal in A then there exists a valuation w on L such that $A \subset V_w \subset L$ and $M_w \cap A = P$.*

(b) *The integral closure of A in L is the intersection of all valuation rings containing A and with quotient field L.*

See e.g. [BAC6,p.91–92]. Note that Bourbaki does not allow the trivial valuation with $V = K$, $M = 0$, $\Gamma = 0$, and so they require $P \neq 0$ in (a) and $A \neq K$ in (b).

Let v be a valuation on the field K with ring V_v and maximal ideal M_v, and let L be an overfield of K. A valuation w on L is an **extension** of v if

$$V_w \cap K = V_v \quad \text{and} \quad M_w \cap K = M_v.$$

Then the natural map

$$\gamma : \Gamma_v = K^*/V_v^* \longrightarrow L^*/V_w^* = \Gamma_w$$

is an order preserving injective homomorphism, and

$$\gamma(v(x)) = w(x) \quad \forall x \in K^*.$$

Applying part (a) of the proposition with $A = V_v$ and $P = M_v$, we have the

Corollary 2.1.3 (Chevalley) *A valuation v on a field K has at least one extension to an arbitrary overfield L.*

The notation $w|v$ means that w is an extension of v.

For now we will not need this level of generality: we use only V satisfying the three axioms

(V1) $$x \in K,\ x \notin V \implies 1/x \in V,$$

(V2) $$V \neq K \quad (V \text{ is not a field}), \quad \text{and}$$

(**V3**) $$V \quad \text{is noetherian.}$$

These assumptions imply, as we will see in a moment, that $\Gamma = \mathbf{Z}$ with the usual $\geq$. (The converse $\Gamma = \mathbf{Z} \implies V$ is noetherian can also be proved.) Then an element $\pi \in V$ with $v(\pi) = 1$ is called a **uniformizer**.

We define **discrete valuation ring** or simply **valuation ring** to mean an integral domain satisfying (V1) – (V3).†

Among the myriad characterizations of valuation ring, we mention the following. PID stands for principal ideal domain.

Proposition 2.1.4 *Let V be an integral domain with quotient field K. Then the following are equivalent:*

(a) *V is a discrete valuation ring.*

(b) *V is a local PID distinct from K.*

(c) *V is a noetherian local ring with maximal ideal $M \neq 0$ such that*

$$\dim_k M/M^2 = 1$$

where $k = V/M$ denotes the residue field.

Proof. (a)$\Longrightarrow$(b) If $x, y \in V$ then (V1) implies that either $x/y \in V$ or $y/x \in V$. Hence, given two principal ideals Vx and Vy, one contains the other. Every ideal I is finitely generated, say $I = Vx_1 + \cdots Vx_n$, and pairwise comparison of Vx_i shows that $I = Vx_i$ for some i. Thus V is a PID. It is local since given two maximal ideals one contains the other, hence they are equal.

(b)$\Longrightarrow$(c) Let the maximal ideal be $M = V\pi$. If x is any nonzero element of V we can write $x = u\pi^n$ with n maximal, so $u \notin M$ hence $u \in V - M = V^*$. It follows that the nonzero ideals of V are precisely $M^n = V\pi^n$, and $u\pi^n + M^n \mapsto u + M$ induces an abelian group isomorphism

$$M^n/M^{n+1} \longrightarrow k^+$$

to the additive group of the residue field. Thus $\dim_k M^n/M^{n+1} = 1$ for $n = 0, 1 \ldots$.

(c)$\Longrightarrow$(a) Choose an element $\pi \in M$ that maps to a basis of the vector space M/M^2, and consider the V-module $N = M/V\pi$. Recall

> Nakayama's Lemma Let V be a commutative ring, N a finitely generated V-module, J an ideal contained in all maximal ideals of V, and suppose $JN = N$. Then $N = 0$.

Thus we will prove that $M = V\pi$ by showing that $MN = N$: given $m + V\pi \in N$, we can write $m = r\pi + s$ where $r \in V$ and $s \in M^2$, say $s = \sum m_i' m_i''$. Then

$$m + V\pi = \sum m_i'(m_i'' + V\pi) \in MN.$$

†When we need to refer to a generalized valuation, *e.g.* with value group $\mathbf{Q}$ or $\mathbf{R}$ with its natural order, or $\mathbf{Z} \times \mathbf{Z}$ with 'lexicographical' order, it will be clearly labelled as *generalized* and *nondiscrete* or *rank* > 1, as the case may be.

By the noetherian property, every nonzero element $r \in V$ can be written as $r = s\pi^n$ where $u \in V$ is not a multiple of π. Thus $u \in V - M = V^*$. Hence every $x \in K^*$ can be written in the form $x = u\pi^n$ where $n \in \mathbf{Z}$, and (V1) follows. ■

Corollary 2.1.5 *Let V be a discrete valuation ring with maximal ideal $M = V\pi$ and quotient field K. Then*

(1) *Each $x \in K^*$ has the unique factorization $x = u\pi^n$ where $u \in V^*$ and $n \in \mathbf{Z}$;*

(2) *the value group $\Gamma = \mathbf{Z}$ and the valuation map $v : K^* \longrightarrow \mathbf{Z}$ is defined by $v(u\pi^n) = n$; if another uniformizer $w\pi$ is chosen, where $w \in V^*$, then v is unchanged since $u\pi^n = uw^{-n}(w\pi)^n$.*

(3) *V is a maximal subring of K: if A is a subring of K containing V, then $A = V$ or $A = K$.*

Proof. (1) The existence of the factorization $u\pi^n$ was explained in the last part of the proof of the proposition. It is unique since n is characterized as the only integer such that $x\pi^{-n} \in V^*$.

(2) By (1), for $x \in K^*$, $xV^* = \pi^n V^* \mapsto n$ defines an isomorphism $\Gamma \longrightarrow \mathbf{Z}$, and $\geq$ is the usual ordering.

(3) If $A \neq V$ then $u\pi^n \in A$ for some $n < 0$. It follows that $\pi^{-1} \in A$, hence $A = K$. ■

Example Let A be a UFD (unique factorization domain) with quotient field K. Thus A contains a designated set $\mathcal{P}$ of irreducible elements such that every element $x \in K^*$ has a unique factorization in the form

$$x = u \prod_{p \in \mathcal{P}} p^{v_p(x)}$$

where $u \in A^*$. Once $\mathcal{P}$ has been chosen, the **gcd** (greatest common divisor) of a set of nonzero $a_i \in A$ is uniquely defined by

$$\gcd\{a_i\} = \prod_{p \in \mathcal{P}} p^{\min_i\{v_p(a_i)\}}.$$

It is an immediate consequence of unique factorization that each v_p is a discrete valuation on K whose ring is the localization A_{Ap}. In fact all discrete valuations occur in this way since, as we have seen, a valuation ring is a UFD with a single irreducible element.

We now present some additional basic information about discrete valuations, treated as background material, *i.e.,* without proof. We retain the same notation: V, K, M, π; also we sometimes denote the residue field by $\widetilde{v}$.

• if we choose any real constant c satisfying $0 < c < 1$ and define the **v-adic absolute value** by $|0| = 0$ and for $x = u\pi^n \in K^*$, $|x| = c^n$ so that

$$\begin{aligned} V &= \{x \in K : |x| \leq 1\}, \\ V^* &= \{x \in K : |x| = 1\}, \\ M &= \{x \in K : |x| < 1\}, \end{aligned}$$

then $\forall x, y \in K$,

$$|xy| = |x||y|, \qquad |x + y| \leq \max\{|x|, |y|\}$$

(the latter being known as the **strong** or **ultrametric** triangle inequality, which implies the ordinary or **archimedean** triangle inequality $|x + y| \leq |x| + |y|$), hence K becomes a metric space if we take $|x - y|$ as the distance between x and y; note that v and $|\ |$, which are often called the π-adic valuation and π-adic absolute value, do not change if π is replaced by another irreducible element $u\pi$; another choice of c replaces $|\ |$ with an equivalent metric (*i.e.* same topology);

• the completion of K as a metric space, denoted $\widehat{K}$ (or K_v, especially if more than one v is under consideration) is a field and can be constructed in the same way as $\mathbf{R}$ is constructed from $\mathbf{Q}$: the set of cauchy sequences $(a_1, a_2, \ldots)$ (**cauchy** meaning as usual that $|a_i - a_j|$ is arbitrarily small for i and j sufficiently big) form a subring of the product ring $K^{\mathbf{N}}$, the null sequences (**null** meaning $|a_i|$ is arbitrarily small for i sufficiently big) comprise a maximal ideal, and $\widehat{K}$ is the quotient ring of cauchy sequences modulo null sequences; the canonical identification of K as a dense subspace of its completion amounts to identifying $a \in K$ with the class of the constant cauchy sequence $(a, a, \ldots)$; K is thus a subfield of $\widehat{K}$; moreover $\widehat{K}$ has a discrete valuation, again denoted v, extending that of K: if $(a_1, \ldots)$ is a non-null cauchy sequence representing $\alpha \in \widehat{K}^*$ then $v(a_i)$ remains constant for large i, this value does not depend on the representative, and is taken as $v(\alpha)$;

• the valuation ring $\widehat{V}$ is the completion of V and admits the alternative description as the inverse limit

$$\widehat{V} = \lim_{\leftarrow} V/M^n$$

with respect to the canonical maps $V/M^j \to V/M^i$ where $x + M^j \mapsto x + M^i$ for $i < j$; then $\widehat{K}$ can be defined as the field of quotients of $\widehat{V}$;

• the residue field remains the same under completion: since the valuation v on $\widehat{K}$ extends the original v on K, therefore $\widehat{V} \cap K = V$, and $\widehat{M} \cap K = M$; in fact the map $V/M \to \widehat{V}/\widehat{M}$ is an isomorphism.

2.1.1 Examples

Example 1. The UFD $\mathbf{Z}$ gives p-adic valuations on $\mathbf{Q}$, one for each prime p. The valuation ring V is the localization $\mathbf{Z}_{(p)}$ of $\mathbf{Z}$ at the prime ideal $p\mathbf{Z}$

and consists of rational numbers of the form m/n where m and n are coprime integers with n not divisible by p. The completion is the ***p*-adic field $\mathbf{Q}_p$** with valuation ring the ***p*-adic integers**

$$\mathbf{Z}_p = \lim_{\leftarrow} \mathbf{Z}/p^n\mathbf{Z}.$$

The residue field is the p-element field

$$\mathbf{Z}/p\mathbf{Z} = \mathbf{Z}_{(p)}/p\mathbf{Z}_{(p)} = \mathbf{F}_p.$$

The nonzero elements of $\mathbf{Q}_p$ are uniquely representable as series $a_np^n + a_{n+1}p^{n+1} + \cdots$ where $a_i \in S$ and S is a set of representatives of the residue field, usually taken to be $S = \{0, 1, \ldots, p-1\}$. If $a_n \neq 0$ then $v(a_np^n + \cdots) = n$. The valuation topology on $\mathbf{Z}_p$ coincides with the inverse limit topology, hence $\mathbf{Z}_p$ is a **profinite** ring, *i.e.,* a topological ring whose topology is compact and totally disconnected, the latter meaning that given two distinct points there is an open and closed set containing the first point and excluding the second. (Recall from general topology that an inverse limit of compact totally disconnected spaces, *e.g.* finite discrete spaces, is again a compact totally disconnected space.)

Example 2. If k is a field then the polynomial ring $k[T]$ is a UFD with quotient field $k(T)$, the field of rational functions in the variable T over the field of constants k. The usual choice for $\mathcal{P}$ is the set of monic irreducible polynomials of degree ≥ 1. Of course $k(T_1) = k(T)$ when

$$T_1 = \frac{aT+b}{cT+d}, \quad \text{for} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(k)$$

but it turns out that only one new valuation shows up besides those we already have from $\mathcal{P}$, namely the one attached to the irreducible polynomial $1/T$ in $k[1/T]$. We denote it by v_∞ rather than $v_{1/T}$, when that is not ambiguous; thus for a polynomial in T of degree n we have

$$v_\infty(a_0 + \cdots + a_nT^n) = v_\infty((1/T)^{-n}f) = -n$$

since $f = a_0(1/T)^n + \cdots + a_n$ is not divisible by $1/T$ in $k[1/T]$. In general for an element of K written as a quotient of polynomials in T, $v_\infty(s/t) = \deg t - \deg s$. The residue field of v_∞ is $k_\infty = k$, and for the corresponding place we have $P_\infty(T) = \infty$, $P_\infty(1/T) = 0$.

For $p \in \mathcal{P}$, the valuation ring V is the local ring $k[T]_{(p)} = \{s/t \in K : p \nmid t\}$ and the residue field is

$$V/pV = k[T]/pk[T],$$

which we denote by k_p. This is an algebraic extension of k of degree $\deg p$ obtained by adjoining a root of p. In fact the image of T under the place P_w is a root of p in k_p:

$$k_p = k(P_w(T)) \quad \text{and} \quad [k_p : k] = \deg p.$$

In particular, for $c \in k$, $k_{(T-c)} = k$, and $P_{T-c}(T) = c$.

For any $p \in \mathcal{P}$ the completion of V is

$$\widehat{V} = \lim_{\leftarrow} V/p^n V = \lim_{\leftarrow} k[T]/p^n k[T] = k_p[[T']],$$

the ring of formal power series $a_0 + a_1 T' + \cdots$ with $a_i \in k_p$, and the quotient field $\widehat{K}$ consists of Laurent series $a_n T'^n + a_{n+1} T'^{n+1} + \cdots$, any $n \in \mathbf{Z}$. The canonical embedding $\lambda : K \longrightarrow \widehat{K}$ is defined as follows. By repeated divisions by p, every polynomial $a \in k[T]$ has a finite p-adic expansion:

$$a = a_0 + a_1 p + \cdots + a_n p^n \quad \text{where} \quad a_i \in k[T] \ \text{ and } \ a_i = 0 \ \text{ or } \ \deg a_i < \deg p.$$

Then $\lambda(a) = \overline{a_0} + \overline{a_1} T' + \cdots + \overline{a_n} T'^n$, where $\overline{a_i}$ denotes the image of a_i in $k_p = V/pV$. This is extended to all of K by $\lambda(f/g) = \lambda(f)/\lambda(g)$.

For $p = \infty$ we replace T by $1/T$ and proceed as in the case $p = T$. Thus $\lambda : k(T) \longrightarrow \widehat{K} = k((T'))$ is defined first for polynomials by $\lambda(a_n T^n + a_{n-1} T^{n-1} + \cdots) = a_n T'^{-n} + a_{n-1} T'^{-(n-1)} + \cdots$, and then extended to all of K as before.

In all cases the extension of v to $\widehat{K}$ is defined by $v(a_n T'^n + a_{n+1} T'^{n+1} + \cdots) = n$, assuming $a_n \neq 0$.

All the valuations v_p on $k(T)$ are **trivial** on k, *i.e.,* $v(c) = 0 \ \forall c \in k^*$, equivalently, $k \subset V$. And in fact these are all the k-trivial valuations on $k(T)$.

2.1.2 The filtration $E_m(K)$

Let v be a discrete valuation on the field K with ring V and uniformizer π. For any point $P = (X_0, \ldots, X_n) \in \mathbf{P}^n(K)$, we can multiply the projective coordinates by a power of π to ensure that $\min\{v(X_i)\} = 0$; let us call such coordinates ***v*-proper**. Note that the v-proper coordinates $(X_0, \ldots, X_n) = (uX_0, \ldots, uX_n)$ of P are unique up to homothety by a unit $u \in V^*$, and therefore $v(X_0), \ldots, v(X_n)$ are well-defined.

Let E be the elliptic curve defined by the nonsingular Weierstrass equation in projective form

$$F = Y^2 Z + a_1 XYZ + a_3 YZ^2 - X^3 - a_2 X^2 Z - a_4 XZ^2 - a_6 Z^3 = 0,$$

which we assume is **defined over** V, *i.e.,* all $a_i \in V$. For $P \in E(K)$, this allows us to define $v(P)$ in terms of v-proper coordinates $P = (X, Y, Z)$ as follows.

- If $v(Z) = 0$ then P is a ***v*-integral point** and we define $v(P) = 0$;
- if $v(Z) > 0$ then, since $\min\{v(X), v(Y)\} = 0$ and at least two terms in the polynomial F must have the minimum value, therefore $v(Y) = 0$ and $3v(X) = v(Z)$; we define $v(P) = v(X)$.

Thus $v(P) \geq 0$ for all P, and in particular

$$v(O) = \infty. \tag{1}$$

In fact, $v(P) = \infty \Leftrightarrow P = O$. Note that P is v-integral iff it is a nonzero point with v-integral affine coordinates $x = X/Z$ and $y = Y/Z$. In the contrary case when $P \neq O$, denoting $v(Z) = m > 0$, we have

$$v(x) = -2m, \quad v(y) = -3m.$$

Thus if $P = (x, y) \neq O$ then

$$v(P) > 0 \Leftrightarrow v(x) < 0 \ \text{ and then } \ v(P) = -v(x)/2.$$

In particular,

$$v(-P) = v(P) \ \text{ for all } \ P. \tag{2}$$

In fact if the coordinates in $P = (X, Y, Z)$ are v-proper, then so are those in $-P = (X, -Y - a_1X - a_3Z, Z)$.

Here are some numerical examples over $K = \mathbf{Q}$. The points $P_1 = (1, 1)$ and $P_2 = (4, 13)$ lie on $y^2 + y = x^3 + x^2 + 34x - 34$, and

$$[2]P_1 = (2 \cdot 83, -19 \cdot 113), \quad [2]P_2 = \left(\frac{19}{3^2}, -\frac{2^4 13}{3^3}\right),$$

$$[-2]P_2 = \left(\frac{19}{3^2}, \frac{181}{3^3}\right), \quad [2]P_1 + P_2 = \left(\frac{61}{3^2}, \frac{2 \cdot 311}{3^3}\right),$$

$$[2]P_1 + [3]P_2 = \left(\frac{2^4 8311}{3^4 7^2}, -\frac{2 \cdot 13 \cdot 1924289}{3^6 7^3}\right), \quad [2]P_1 - P_2 = \left(\frac{5 \cdot 17}{2^2 3^2}, -\frac{17 \cdot 109}{2^3 3^3}\right).$$

With $v = v_3$, these illustrate the following rule:

$$v(P + Q) \geq \min\{v(P), v(Q)\} \quad \text{with equality when } \ v(P) \neq v(Q). \tag{3}$$

The proof is trivial if either of the points is v-integral or is O, or if their sum is O. But the proof is distinctly nontrivial when both $0 < v(P) < \infty$ and $0 < v(Q) < \infty$, at least if we approach it using the formulas of Proposition 1.4.1. The verification seems to be particularly troublesome when $v(P) = v(Q) > 0$, treating the various possibilities in $0 \leq v(2) \leq \infty$.

Our proof will use the formal group $\widehat{E}$ attached to E, which is defined in §2.5.2, and will be the first illustration of the power and utility of formal groups in the theory of elliptic curves. Another illustration will occur in the analysis of torsion in the group $E(K)$; see §2.10. Before the introduction of formal groups that analysis was performed by rather arduous calculations with the division polynomials.

For positive integers m define

$$E_m(K) = \{P \in E(K) : v(P) \geq m\},$$

so that we have a filtration

$$E(K) \supset E_1(K) \supset E_2(K) \supset \cdots \supset \bigcap_{i=1}^{\infty} E_i(K) = \{O\}.$$

Caution In §6.2, using quite different considerations, we will define a subgroup $E_0(K)$ lying between $E(K)$ and $E_1(K)$. Thus $E_0(K)$ does *not* signify simply $\{P : v(P) \geq 0\} = E(K)$.

It is convenient to record here the essential point of the above discussion, but the proof will only be completed in Proposition 2.6.7.

Proposition 2.1.6 *Let v be a discrete valuation on the field K with ring V, and let E be an elliptic curve defined over V, that is, given by a nonsingular Weierstrass equation with $a_1, \ldots, a_6 \in V$. Then $E_m(K)$ for $m \geq 1$ are subgroups of $E(K)$.*

Remark. Later in this chapter we will use the 'parameter' $z = -x/y$ for a nonzero point $P \in E_1(K)$ with affine coordinates (x, y). Note that

$$P \in E_m(K) \Longleftrightarrow v(z) \geq m \text{ and } v(x) < 0.$$

One must not forget the last condition since examples of $v(x) > v(y) \geq 0$, and therefore $P \notin E_1$, are easily found.

Proof. This is immediate from the three statements (1), (2) and (3). (And conversely, the statement that the E_m are all subgroups implies (1), (2) and (3)). For the proof of (3) see Proposition 2.6.7.

2.1.3 Finite extensions

We now give a brief résumé of the facts concerning extensions w of a discrete valuation v on K to an overfield L of finite degree $n = [L : K]$; we omit proofs, regarding the subject as background material.

First of all, an extension w does exist by Corollary 2.1.3. The residue field $\widetilde{w}$ of w is an extension of $\widetilde{v}$ of finite degree denoted

$$f = f(w, v) = [\widetilde{w} : \widetilde{v}].$$

Moreover, any such w is discrete, and the order preserving injection

$$\gamma : \Gamma_v = \mathbf{Z} \longrightarrow \mathbf{Z} = \Gamma_w$$

is multiplication by an integer $e = e(w, v) \geq 1$ called the **ramification index**. Thus

$$w(x) = ev(x) \quad \forall x \in K.$$

This definition of the index agrees with that in [BAC6]; but some authors, *e.g.* [Zar-Sa58,p.284] define the index to be e times $[\widetilde{w} : \widetilde{v}]_i$, the inseparable part of the residue field degree. All agree with the following definition. The extension is **unramified** if

$$e(w, v)[\widetilde{w} : \widetilde{v}]_i = 1,$$

which means of course that $e = 1$ and the residue field extension is separable.

In general there are at most $n = [L : K]$ extensions w of v. There is exactly one extension when either

(i) $K = K_v$ is complete, and then $L = L_w$ is also complete, or

(ii) the field extension L/K is purely inseparable.

When the extension is unique it is given by the formula

$$w(x) = \frac{e}{n} v(N(x))$$

where N denotes the norm from L to K.

If $w_1, \ldots, w_g$ are the distinct extensions of v to L then

$$\sum_{i=1}^{g} e(w_i, v) f(w_i, v) \leq n \tag{¶}$$

with equality in most "reasonable" cases: there is equality if either

• the extension L/K is separable (hence all cases of $\operatorname{char} K = 0$), or

• if V contains a Dedekind domain† R (*e.g.* a PID) whose quotient field is K and which is a finitely generated algebra over some field.

With $w|v$ as above, the injection $K \hookrightarrow L$ induces a canonical injection of the completions $K_v \hookrightarrow L_w$ (by the universal property of the completion) and the **local degree** satisfies — in all cases —

$$[L_w : K_v] = e(w, v) f(w, v).$$

Moreover, if $\hat{w}$ and $\hat{v}$ denote the extensions of w and v to L_w and K_v respectively, then $e(\hat{w}, \hat{v}) = e(w, v)$ and $f(\hat{w}, \hat{v}) = f(w, v)$. Of course the analog of g is 1 since $\hat{v}$ has the unique extension $\hat{w}$ from K_v to L_w.

When the extension L/K is Galois, the Galois group G acts transitively on the extensions of v: if w is any extension of v and $\sigma \in G$ then σw defined by $(\sigma w)(x) = w(x^\sigma)$ also extends v, and all extensions are obtained in this way. Moreover the $e(\sigma w, v)$ have a common value e and similarly all the $f(\sigma w, v)$ have the same value f. Thus

$$efg = n.$$

†We recall the definition of Dedekind domain and basic properties in §2.2.1.

Consider now a tower of arbitrary finite extensions $K \subset L \subset M$; let v be a valuation on K, let w extend v to L and let W extend w to M. The following multiplicative properties are immediate from the definitions:

$$e(W,v) = e(W,w)e(w,v), \qquad f(W,v) = f(W,w)f(w,v).$$

2.1.4 Gauss's lemma

Proposition 2.1.7 *Let v be a valuation on the field K with ring V. Then v can be extended to a valuation v' on the power series field $K((T))$ as follows.*

$$v'\left(\sum_{i=N}^{\infty} a_i T^i\right) = \min\{v(a_i)\}.$$

The valuation ring of v' is $V((T))$.

Remark. That v can be extended from K to the field of rational functions $K(T)$ is one of at least three results known as Gauss's lemma. This is a corollary of the proposition since $K(T)$ is a subfield of $K((T))$.

Proof. If $f = \sum a_i T^i$ and $g = \sum b_j T^j$ are nonzero members of $K((T))$, let $fg = \sum c_k T^k$, let $v'(f) = \alpha$, say $v(a_i) = \alpha$ with i minimal, and similarly $v'(g) = \beta = v(b_j)$ with j minimal. Then $v(c_{i+j}) = \alpha + \beta$ since $v(a_i b_j) < v(a_h b_{i+j-h})$ for all $h \neq i$. Also $v(c_k) \geq \alpha + \beta \, \forall k$, and therefore

$$v'(fg) = \alpha + \beta = v'(f) + v'(g).$$

It is also clear that

$$v'(f+g) \geq \min\{v'(f), v'(g)\}. \quad \blacksquare$$

2.2 Krull domains

The study of the arithmetic properties of elliptic curves leads one to consider families of discrete valuations. For example if the elliptic curve E is given by a Weierstrass equation with coefficients in $\mathbf{Z}$ or more generally in the ring of integers $\mathcal{O}$ of a number field, then the valuations associated to the Dedekind domain $\mathcal{O}$ (we recall definitions below) will help us to investigate E. Similarly we may wish to consider E with Weierstrass coefficients in, say, a two variable polynomial ring $k[s,t]$ over a field k; this gives what one calls a two parameter family of elliptic curves. The ring $k[s,t]$ is a UFD but is not Dedekind.

Consider first the *globalization* of discrete valuation rings, that is, the class of integral domains A such that for every nonzero prime ideal P the local ring A_P is a discrete valuation ring. This class consists of all Dedekind domains plus a certain class of non-noetherian Prufer domains (known as *almost Dedekind;*

cf. [Hei67]). However the only UFD's included in this class are the PID's, so this class of rings would be too restrictive for our purposes.

The class of rings that contains both Dedekind domains and UFD's and is closely tied to a family of discrete valuations is the following: †

A **Krull domain** is an integral domain A such that ([BAC7], p.14)

(i) for every $P \in \mathcal{P}$, where $\mathcal{P}$ denotes the set of all minimal prime ideals, the local ring A_P is a discrete valuation ring;

(ii) $A = \bigcap\{A_P : P \in \mathcal{P}\}$;

(iii) each nonzero $a \in A$ is contained in at most finitely many members of $\mathcal{P}$.

The valuations associated to the $P \in \mathcal{P}$ are called the **essential** valuations of A; these are precisely the valuations on the quotient field of A that are non-negative on A. (For, from [BAC7, Cor.2, p.10], if v is non-negative on A, then v together with the essential valuations of A satisfy axioms $\mathrm{AK_I}$ to $\mathrm{AK_{III}}$, and so v must be essential for A.)

We also use $\mathcal{P}$ to denote the set of essential valuations (the context should make it clear whether we mean the prime ideals or the valuations), and we write $\mathcal{P}_A$ when it is necessary to identify A.

We quote from [BAC7]:

Proposition 2.2.1 *Let A be a Krull domain with quotient field K. Then the following extensions B of A are Krull:*

(i) *The localization $B = S^{-1}A$ for a multiplicative subset $S \subset A$, and*

$$\mathcal{P}_B = \{v \in \mathcal{P}_A : v(s) = 0\, \forall s \in S\}.$$

(ii) *The integral closure B of A in a finite extension L of K, and $\mathcal{P}_B$ consists of the extensions to L of all $v \in \mathcal{P}_A$.*

(iii) *The polynomial ring $B = A[\{x_i\}]$ in an arbitrary set of indeterminates; for simplicity in the case $B = A[x]$ of a single variable, $\mathcal{P}_B$ consists of two types of valuations:*

(a) *the extensions of $v \in \mathcal{P}_A$ to $L = K(x)$ as given by Gauss's lemma* (*cf.* the remark following Proposition 2.1.7)*, and*

(b) *the standard valuations on $K[x]$ that are trivial on K* (*cf.* Example 2 in §2.1.1. ■

The **divisor group** $\mathrm{Div}(A)$ of the Krull domain A is the free abelian group on the set $\mathcal{P}$ written additively, and an element of $\mathrm{Div}(A)$ is a **divisor**. The **support** of a divisor

$$D = \sum_{P \in \mathcal{P}} n_P P$$

†An authoratative reference for these concepts is [BAC7], except for the homological aspects for which one should consult [Car-Ei56].

is the finite set $\{P : n_P \neq 0\}$. For nonzero $x \in K$, where K denotes the quotient field of A, one defines the **principal divisor**

$$\mathrm{div}(x) = \sum v_P(x)P \in \mathrm{Div}(A).$$

Thus $\mathrm{div}(xy) = \mathrm{div}(x) + \mathrm{div}(y)$, $\mathrm{div}(1/x) = -\mathrm{div}(x)$, for $x, y \in K^*$, and the principal divisors form a subgroup H of $\mathrm{Div}(A)$. The quotient group $\mathrm{Div}(A)/H$ is denoted $\mathrm{Cl}(A)$, and is called the **divisor class group**. For $D \in \mathrm{Div}(A)$, the class of D in $\mathrm{Cl}(A)$ is denoted $\mathrm{cl}(D)$.

By $\sum m_P P \leq \sum n_P P$ one means $m_P \leq n_P$ $\forall P$. Clearly by (ii) in the definition of Krull domain, for $x, y \in K^*$

$$x \in A \Leftrightarrow \mathrm{div}(x) \geq 0; \quad \mathrm{div}(x) = \mathrm{div}(y) \Leftrightarrow x = uy \text{ for some } u \in A^*.$$

An integral domain A is a UFD *iff it is Krull with* $\mathrm{Cl}(A) = 0$.

Proof. If A is a UFD, then the minimal prime ideals are $A\pi$ where π is irreducible and the three axioms in the definition are obviously satisfied. $\mathrm{Div}(A)$ is isomorphic with the multiplicative group K^*/A^* and all divisors are principal. Conversely, if A is Krull with all divisors principal, then the minimal prime ideals are principal, say $P = A\pi_P$ where the π_P are irreducible. Then $\forall\, x \in K^*$, $x = u \prod \pi_P^{v_P(x)}$ for some $u \in A^*$, and A is a UFD.

Recall that an integral domain A with quotient field K is **completely integrally closed** if for all $x \in K$ and nonzero $d \in A$ we have the implication

$$dx^n \in A \,\forall n \in \mathbf{N} \Longrightarrow x \in A.$$

This implies that A is integrally closed and the converse is true when A is noetherian.

Every Krull domain is completely integrally closed.

Proof: If x, d are as above then for every essential valuation v we have $v(d) + nv(x) \geq 0$, so $v(x) \geq 0$, hence $x \in A$ by (ii) of the definition.

A noetherian domain is Krull iff it is integrally closed ([BAC7], p.9).

From [BAC7], p.12:

The approximation theorem for Krull domains: *let A be a Krull domain with quotient field K, let $v_1, \ldots, v_r$ be the valuation maps corresponding to distinct minimal prime ideals and let $n_1, \ldots, n_r$ be integers. Then there exists $x \in K$ such that $v_i(x) = n_i$ for $i = 1, \ldots, r$ and $v(x) \geq 0$ for all other essential v.*

Here is a simple application of the approximation theorem. For convenience we call the minimal prime ideals of a Krull domain A the **primes** of A.

Corollary 2.2.2 *Let A be a Krull domain with only finitely many primes $P_1, \ldots, P_n$. Then* $\mathrm{Cl}(A) = 0$*, and consequently A is a* UFD.

Proof. Since $\mathrm{Div}(A)$ is generated by the P_i, it is sufficient to show that each P_i is a principal divisor. Choose $a \in A$ so that $v_{P_i}(a) = 1$ and $v_{P_j}(a) = 0$ for all $j \neq i$. Then $\mathrm{div}(a) = P_i$. ■

For the next proposition, which will be used in §5.7, we make two definitions. A divisor $D = \sum n_P P$ is **squarefree** if $n_P \neq 0 \Longrightarrow n_P = 1$; a divisor class $c \in \mathrm{Cl}(A)$ is **prime-full** if it contains infinitely many primes.

Proposition 2.2.3 *Let A be a Krull domain with* $\mathrm{Cl}(A)$ *finite. Then*

(a) $\mathrm{Cl}(A)$ *is generated by the prime-full classes;*

(b) *for any given finite set S of primes and divisor class c, there exists a squarefree divisor in c whose support is disjoint from S.*

Remark. When A has infinitely many primes perhaps it is true that every class is prime-full. This is so when A is the ring of integers in a number field K: by the Čebotarev density theorem applied to the Hilbert class field of K, the prime ideals of A are equi-distributed among the classes.

Proof. Let the prime-full classes be denoted $c_1, \ldots, c_m$ $(m \geq 0)$. First let us deduce (b) from (a). A typical class c can be written as a sum of c_i since $\mathrm{Cl}(A)$ is finite. We can choose distinct prime representatives for these summands outside of any given S. Then the sum of these primes is a squarefree divisor in c with support disjoint from S.

Let H be the subgroup generated by the c_i and let T denote the set of primes whose classes fall in $\mathrm{Cl}(A) - H$. Since $\mathrm{Cl}(A) - H$ consists of a finite number of non-prime-full classes, T is finite, say $T = \{P_1, \ldots, P_n\}$. Suppose $n > 0$. By the approximation theorem choose $a \in A$ such that $v_{P_1}(a) = 1$ and $v_{P_2}(a) = \cdots = v_{P_n}(a) = 0$. Then $\mathrm{div}(a) = P_1 + D$ where the support of D consists of primes whose classes are in H. But this gives the contradiction $\mathrm{cl}(P_1) = -\mathrm{cl}(D) \in H$. Hence $T = \emptyset$, *i.e.,* all primes have their classes in H, so by definition $\mathrm{Cl}(A) = H$. ■

The discussion in §2.1.2 of the v-adic value of points $P \in E(K)$ can be globalized immediately:

Proposition 2.2.4 *Let A be a Krull domain with quotient field K, and let E be an elliptic curve defined over A. Then, for each nonzero $P = (x, y) \in E(K)$ and each essential valuation v of A we have*

$$v(x) < 0 \Longrightarrow v(x) = -2m \quad \textit{and} \quad v(y) = -3m,$$

where $m = v(P)$ is a positive integer.

In particular, if A is a UFD*, then we can write the coordinates of P in the form*

$$(x, y) = \left(\frac{m}{e^2}, \frac{n}{e^3}\right),$$

where $m, n, e \in A$ and $\gcd(m, e) = \gcd(n, e) = 1$.

Also when A is a UFD, the projective coordinates (a, b, c) of a point can be chosen to be ***A*-proper**, i.e., *v-proper $\forall$ essential v, equivalently, $a, b, c \in A$ and* $\gcd(a, b, c) = 1$.

2.2.1 Dedekind domains

The following definitions apply to any integral domain A with quotient field K. A **fractional ideal** is an A-submodule I of K for which there exists a "common denominator" — a nonzero element $d \in A$ such that $dI \subset K$. Examples are the ideals of A, and finitely generated submodules of K. A **principal fractional ideal** is one of the form Ax for some $x \in K$. The **product** of two fractional ideals I and J is defined as the set of all finite sum $\sum ij$ of products of elements from I and J and is denoted IJ. This is again a fractional ideal since if common denominators for I and J are d and d', then dd' serves as a common denominator for IJ. Of course the product of ideals is an ideal. A fractional ideal I is **invertible** if there exists a fractional ideal J such that $IJ = A$. In that case J is unique and is given by

$$J = \{x \in K : xI \subset A\}.$$

A **Dedekind domain** is an integral domain satisfying the following equivalent conditions

— Krull and all nonzero prime ideals are maximal;
— hereditary, *i.e.,* every nonzero ideal is invertible;
— every submodule of a projective module is projective;
— every quotient module of an injective module is injective;
— every divisible module is injective;
— every ideal is a product of prime ideals;
— every nonzero ideal is uniquely a product of prime ideals;
— noetherian, integrally closed and every nonzero prime ideal is maximal.

An integral domain is a PID iff it is both Dedekind and a UFD. Examples of PID's are: discrete valuation rings, $\mathbf{Z}$ and $k[X]$ (k any field), and for that matter k itself. However polynomial rings $k[X, Y, \ldots]$ in two or more variables are UFD's that are not Dedekind.

We quote the analog of Proposition 2.2.1(i)–(ii) for Dedekind domains from [Zar-Sa58,ch.V]:

Proposition 2.2.5 *Let A be a Dedekind domain with quotient field K. Then the following extensions B of A are Dedekind:*

(i) *The localization $B = S^{-1}A$ for a multiplicative subset $S \subset A$.*

(ii) *The integral closure B of A in a finite extension L of K. Moreover, when the extension L/K is a separable, then there are only finitely many $w \in \mathcal{P}_B$ for which the ramification index $e(w, v) > 1$.* ■

Now let A be a Dedekind domain with quotient field K, let its set of maximal ideals be $\mathcal{P} = \{P\}$, and for an ideal I let

$$I = \prod_{P \in \mathcal{P}} P^{v_P(I)}$$

denote the unique prime ideal factorization. Here are some basic facts.

- The nonzero fractional ideals are all invertible and form an abelian group, with the product defined above and with identity element A. This group is free on the set $\mathcal{P}$ and thus the unique factorization is extended to all nonzero fractional ideals:

$$J = \prod P^{v_P(J)}, \qquad v_P(J) \in \mathbf{Z}.$$

- The fact that the Dedekind domain A is a Krull domain can be expressed this way: for $a \in K^*$, $v_P(a) := v_P(Aa)$ defines a discrete valuation on K, the corresponding valuation ring is the localization A_P, and

$$A = \bigcap_{P \in \mathcal{P}} A_P.$$

- $\mathrm{Div}(A)$ is isomorphic with the group of nonzero fractional ideals via

$$\sum n_P P \longmapsto \prod P^{n_P}$$

  and in this bijection principal divisors correspond exactly with principal fractional ideals.

  The divisor class group $\mathrm{Cl}(A)$ is also known as the **ideal class group**.

- The approximation theorem can be sharpened for Dedekind domains in the following way ([BAC7],p.26): given finitely many distinct prime ideals P_i in A, positive integers n_i and elements $x_i \in K$ there exists $x \in K$ such that $v_{P_i}(x - x_i) = n_i$ for all i in the finite set and $v_P(x) \geq 0$ for all other P. The special case where the $x_i = a_i \in A$ is known as **the Chinese Remainder Theorem**: the system of simultaneous congruences

$$x \equiv a_i \bmod P_i^{n_i}$$

  has a solution $x \in A$. Note that $x \equiv a \bmod P^n$ is equivalent to $v_P(x - a) \geq n$. The Chinese remainder theorem is equivalent to the statement that the natural map of A to the product of $A/P_i^{n_i}$ induces a ring isomorphism

$$A/\prod P_i^{n_i} \cong \prod A/P_i^{n_i}.$$

  Here are some simple applications of the approximation theorem.

  (i) Every fractional ideal can be generated by two elements: $J = Ax + Ay$. This is not true for Krull rings in general; for example the 3-variable polynomial ring $A = k[x, y, z]$ is a UFD, and the ideal $Ax + Ay + Az$ cannot be generated by two elements.

(ii) Let P a maximal ideal. Then a uniformizer π for v_P can be chosen in A, and for any positive integer e, the natural map

$$A/P^e \longrightarrow A_P/\pi^e A_P$$

is a ring isomorphism.

2.2.2 One variable function fields

A **one variable function field with constant field k** is a field L containing k as a subfield such that

FF1 k is algebraically closed in L;
FF2 L is of transcendence degree 1 over k;
FF3 if $x \in L$ is transcendental over k then the extension $L/k(x)$ is finite.

k is called the **constant field** or **field of constants**. Two immediate examples are the field of rational functions $L = k(x)$, and the function field $K(x, y)$ of an elliptic curve as defined in §1.6 where now the constant field is denoted K.

If another transcendental x' is chosen in **FF3** then, by **FF2**, $L/k(x')$ is also finite but the degree may be different from that of $L/k(x)$. We do not assume that L is separably generated over k, *i.e.,* that x can be chosen so that L is separable over $k(x)$. However this is always so for $L = K(x, y)$ where x and y are related by a Weierstrass equation since the extensions $L/K(x)$ and $L/K(y)$ are of degrees 2 and 3 repectively, and so cannot both be inseparable. In fact both extensions are separable in the case of an elliptic curve ($\Delta \neq 0$), as we noted in §1.6.

When L/k satisfies only **FF2** and **FF3**, we say that L is a one variable function field **over** k. Then the constant field is the algebraic closure of k in L; this is a finite extension of k.

The following notation is convenient.†

Let L be a one variable function field over k. Then $$\mathrm{gam}_k(L)$$ denotes the set of all k-trivial valuations on L.

Thus from Example 2 of §2.1.1, $\mathrm{gam}_k(k(x)) = \{v_p : p \in \mathcal{P} \cup \{\infty\}\}$. In general, $\mathrm{gam}_k(L)$ consists of all extensions w to L of these v. Using the PID $R = k[x]$, the second equality case of (¶) in §2.1.3 applies:

$$\sum_{w|v} e(w, v) f(w, v) = [L : k(x)]. \qquad (\#)$$

†For the meaning of gam in the context of general commutative rings see [Con70].

(See (##) below for a generalization.) As also explained in Example 2 in §2.1.1, the residue field $\widetilde{v_p}$ is an extension of k of degree $\deg p$, where $\deg \infty := 1$. For $w|v_p$ we define the **degree** by

$$\deg w = f(w, v_p) \deg p = [\widetilde{w} : k].$$

In particular, taking $L = k(x)$, this defines $\deg v_p = \deg p$ and $\deg v_\infty = 1$.

This notation is carried over to the place P_w corresponding to w: we define $\deg P_w = \deg w$. The place P_w is called a **rational place** when $\deg P_w = 1$. The degree is now extended linearly to the **divisor group** $\mathsf{Div}_k(L)$, which is the additive free abelian group based on the set $\{P_w : w \in \mathrm{gam}_k L\}$:

$$\deg \sum n_w P_w = \sum n_w \deg P_w, \quad \text{where} \quad n_w \in \mathbf{Z}.$$

Clearly $$\mathsf{Div}_k^0(L) := \{D \in \mathsf{Div}_k(L) : \deg(D) = 0\}$$

is a subgroup. For $D = \sum n_w P_w \in \mathsf{Div}_k(L)$ we define $w(D) = n_w$, and the **divisor of zeros** and the **divisor of poles** are

$$\mathsf{Zer}(D) = \sum \{w(D)P_w : w(D) > 0\}, \quad \mathsf{Pol}(D) = \sum \{-w(D)P_w : w(D) < 0\}.$$

Thus $$D = \mathsf{Zer}(D) - \mathsf{Pol}(D) \quad \text{and} \quad \deg D = \deg \mathsf{Zer}(D) - \deg \mathsf{Pol}(D).$$

Proposition 2.2.6 (a) *For $f \in L^*$, $w(f) \neq 0$ for only finitely many $w \in \mathrm{gam}_k L$. Consequently*

$$\mathsf{div}(f) := \sum_w w(f) P_w \in \mathsf{Div}_k(L).$$

A divisor of the form $\mathsf{div}(f)$ is called a **principal divisor**. $\mathsf{Zer}(\mathsf{div}(f))$ and $\mathsf{Pol}(\mathsf{div}(f))$ are shortened to $\mathsf{Zer}(f)$ and $\mathsf{Pol}(f)$.

(b) $\mathsf{div}(f) = 0$ *iff f is a constant: $f \in k^*$.*

(c) *Every principal divisor has degree* 0. *More precisely,*

$$\deg \mathsf{Zer}(f) = \deg \mathsf{Pol}(f) = [L : k(f)].$$

(d) $\mathsf{div}(f^{-1}) = -\mathsf{div}(f)$, $\mathsf{div}(fg) = \mathsf{div}(f) + \mathsf{div}(g)$, *hence the principal divisors form a subgroup of* $\mathsf{Div}_k^0(L)$.

Proof. (d) is clear once (a) is proved. We prove (a), (b) and (c) together.

If $f \in k^*$ then $w(f) = 0$ for all w, so let f be nonconstant. Then, by **FF1**, $k[f]$ is a polynomial ring, f playing the role of x now, and L is a finite extension of $k(f)$. We have $v_f(f) = 1$, $v_\infty(f) = -1$, and $v(f) = 0$ for all other

$v \in \mathrm{gam}_k(k(f))$. The finitely many extensions w of v_f and v_∞ to L are the only valuations on L for which $w(f) \neq 0$. Finally, using (#),

$$\deg \mathsf{Zer}(f) = \sum_{w|v_f} w(f) = [L : k(f)],$$

and similarly

$$\deg \mathsf{Pol}(f) = \deg \mathsf{Zer}(1/f) = \sum_{w|v_\infty} w(f) = [L : k(f)]. \quad \blacksquare$$

Proposition 2.2.7 *Let L be a one variable function field over k with constant field k_L and let M be a finite extension of L.*

(a) *M is a one variable function field whose constant field is an extension k_M of k_L satisfying $k_M \cap L = k_L$, hence $[k_M : k_L] \leq [M : L]$.*

(b) *If $v \in \mathrm{gam}_k L$ and w runs through the extensions of v to M, then*

$$\sum_{w|v} e(w,v)f(w,v) = [M : L]. \qquad (\#\#)$$

(c) *If N denotes the norm from M to L then for $x \in M$,*

$$v(Nx) = \sum_{w|v} f(w,v)w(x).$$

Proof. (a) is obvious.

(b) Let x be a transcendental in L, so we have the tower of fields $k(x) \subset L \subset M$, and let A be the integral closure of $k[x]$ in L. Then

(i) A is a finitely generated k-algebra ([BAC5,p.63]), and

(ii) A is a Dedekind domain (Proposition 2.2.5).

Thus the second equality case of (¶) in §2.1.3 implies (##).

(c) See for example [BAC6,p.149]. Note that their extensions v' of v to M include the ramification index, so that $v'(z) = v(z)$ for $z \in L$, and $v'(x)$ may have a fractional value. $\blacksquare$

With L, M as in the proposition, let $\theta : L \hookrightarrow M$ denote the embedding and define

$$\theta^* : \mathsf{Div}_{k_L}(L) \longrightarrow \mathsf{Div}_{k_M}(M)$$

as the $\mathbf{Z}$-linear extension of

$$P_v \longmapsto \sum_{w|v} e(w,v)P_w.$$

Since $\deg(P_w) = f(w,v)\deg(P_v)$, we have the

Corollary 2.2.8 *For $A \in \mathsf{Div}_{k_L}(L)$,*

$$\deg(\theta^* A) = [M : L]\deg(A). \quad \blacksquare$$

Continuing with the same extension M/L, let M_s be the maximal separable extension of L in M, let $w_1, \ldots, w_g$ be the extensions of v to M, and let $w'_1, \ldots, w'_g$ be their restrictions to M_s. Since M is purely inseparable over M_s, each w'_i has a unique extension to M and therefore the w'_i are distinct.

For use in Chapter 6 we state a special case of the preceding proposition:†

Proposition 2.2.9 *With the above notation, suppose now that $k = \overline{k}$ is algebraically closed, so that L, M_s and M are all function fields with constant field $\overline{k}$. Then $f(w_i, v) = f(w_i, w'_i) f(w'_i, v) = 1$, hence*

$$\sum e(w_i, v) = [M : L], \quad \sum e(w'_i, v) = [M_s : L], \quad e(w_i, w'_i) = [M : M_s] \ \forall i.$$

Therefore if $e(w'_i, v) = 1$ for $1 = 1, \ldots, g$ (and there are only finitely many v for which this is not true by Proposition 2.2.5)*, then*

$$g = [M_s : L]. \quad \blacksquare$$

We quote another useful fact without proof in a

Lemma 2.2.10 *Let L be a one variable function field with constant field k and let $M = L(y)$ be a finite extension, say the minimum polynomial of y over L is ϕ. Let $v \in \mathrm{gam}_k L$ with valuation ring V, suppose that y is integral over V, and let ϕ have the factorization $\phi_1 \cdots \phi_g$ over the v-adic completion of K. Then the ϕ_i are in one to one correspondence with the extensions $w_1, \ldots, w_g$ of v to M, and $\deg \phi_i = e(w_i, v) f(w_i, v)$.* $\blacksquare$

Here are two examples relevant to our work.

Proposition 2.2.11 (a) *Let E be the elliptic curve defined by the nonsingular Weierstrass equation $y^2 + a_1 xy + a_3 = x^3 + a_2 x^2 + a_4 x + a_6$ over the field K, and let $L = K(x, y)$ be the function field. Then L is a degree 2 Galois extension of $K(x)$, and for each $v \in \mathrm{gam}_K(K(x))$ we have*

$$efg = 2.$$

In the case $v = v_\infty$ we have $e = 2$, hence $f = g = 1$. Let w_∞ denote the unique extension on L. Then

$$w_\infty(x) = -2, \qquad w_\infty(y) = -3.$$

(b) *Let $\mathrm{char}\, K \neq 2$ and let $L = K(x, y)$ where*

$$y^2 = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0, \quad a_i \in K, \ a_n \neq 0,$$

and the polynomial on the right has no repeated factors in $K[x]$. Then $A := K[x, y]$ is the integral closure of $K[x]$ in L, and consequently A is Dedekind. For v_∞ we have $efg = 2$ and

†In §6.1 the theory of one variable function fields is resumed where a self-contained proof of the Riemann-Roch theorem is presented.

- *if n is odd then $e = 2$, hence $f = g = 1$;*
- *if n is even and $\sqrt{a_n} \in K$ then $e = f = 1$ and $g = 2$ (the case of two rational places at infinity);*
- *if n is even and $\sqrt{a_n} \notin K$ then $e = g = 1$ and $f = 2$.*

Remarks. See Propositions 2.2.14 and 2.2.15 for a continuation of these results.

In part (b), the assumption that the polynomial has no repeated factors is implied by the assumption that it has no repeated roots in an algebraic closure $\overline{K}$. But the converse is not true in general when $\operatorname{char} K > 0$. For example, if K is an imperfect field of characteristic 3 and $t \in K^* - K^{*3}$, then $x^3 - t$ has no repeated factors in $K[x]$, and so the proposition applies to $y^2 = x^3 - t$. But the polynomial has 3 equal roots, hence this Weierstrass equation has $\Delta = 0$.

Proof. Let $v = v_\infty$ and $w|v$.

(a) We saw in §1.6 that L is separable over $K(x)$ (also over $K(y)$), and (#) takes the form $efg = 2$. From

$$y(y + a_1x + a_3) = x^3 + a_2x^2 + a_4x + a_6$$

we deduce

$$w(y) + w(y + a_1x + a_3) = ev(x^3 + a_2x^2 + a_4x + a_6) = -3e.$$

If $w(y) \geq w(x)$ then since $w(x) = -e$ we would have $w(y + a_1x + a_3) \geq -e$ and the above equation would imply the contradiction $-3e \geq -2e$. Hence $w(y) < w(x)$, so $w(y + a_1x + a_3) = w(y)$, and $2w(y) = -3e$. This implies $e = 2$, therefore $f = g = 1$, and $w(y) = -3$.

(b) The assumptions ensure that L is a quadratic and separable over $K(x)$, so again we have $efg = 2$.

Let $\xi = \alpha + \beta y \in L$ with $\alpha, \beta \in K(x)$ so the Galois conjugate is $\bar{\xi} = \alpha - \beta y$. We must prove that if both $\xi + \bar{\xi} = 2\alpha$ and $\xi\bar{\xi} = \alpha^2 - \beta^2(a_nx^n + \cdots)$ are in $K[x]$ then $\alpha, \beta \in K[x]$. Since $\operatorname{char} K \neq 2$, 2 is invertible and $(\xi + \bar{\xi})/2 = \alpha \in K[x]$, hence $\alpha^2 - \xi\bar{\xi} = \beta^2(a_nx^n + \cdots) \in K[x]$. Since the polynomial has no repeated factor, β can have no factor in its denominator, *i.e.*, $\beta \in K[x]$.

Since $a_n \neq 0$, it follows that $2w(y) = w(a_nx^n + \cdots) = -ne$, hence $e = 2$ when n is odd. Let n be even. Then

$$P_w(y/x^{n/2})^2 = P_w(a_n + a_{n-1}x^{-1} + \cdots) = a.$$

Thus the residue field of w contains $\sqrt{a_n}$, which obliges $f = 2$ when $\sqrt{a_n} \notin K$.

Conversely suppose that $a_n = s^2$, $s \in K$. Referring to the lemma,

$$(y/x^{n/2})^2 = s^2 + a_{n-1}x^{-1} + \cdots$$

is integral over the valuation ring V of v_∞ and it remains to show that $\phi(Y) = Y^2 - (s^2 + a_{n-1}T + \cdots)$ has roots in the completion $\widehat{V} = K[[T]]$. In fact any

series in $\widehat{V}$ whose constant term is a nonzero square has square roots in $\widehat{V}$: for any given $b_1, b_2, \ldots \in K$, write

$$\begin{aligned} s^2 + b_1T + b_2T^2 + \cdots &= (s + c_1T + c_2T^2 + \cdots)^2 \\ &= s^2 + 2sc_1T + (2sc_2 + c_1^2)T^2 + \cdots \end{aligned}$$

and solve succesively for $c_1, c_2, \ldots$ ■

2.2.3 Elliptic function fields

Proposition 2.2.12 *Let E be an elliptic curve defined over the field K by the nonsingular Weierstrass equation $F(x, y) = 0$ with function field $L = K(x, y)$ and affine coordinate ring $A = K[x, y]$. Then A is the integral closure of $K[x]$ in L, and consequently A is Dedekind.*

Proof. It is convenient to deal separately with 3 cases: $\operatorname{char} K \neq 2$, $\operatorname{char} K = 2$ with $j = 0$ (the supersingular case), and $\operatorname{char} K = 2$ with $j \neq 0$ (the ordinary case). We may make substitutions of the form

$$x = u^2x' + r, \quad y = u^3y' + su^2x' + t, \quad \text{where } \ r, s, t \in K, \ u \in K^*,$$

then divide the resulting equation by u^6, in order to obtain simplified Weierstrass equations, since $K[x] = K[x']$ and $A = K[x', y']$. Calculation shows that the discriminant of the new equation is $\Delta' = u^{-12}\Delta$. (Such simplifications are studied sytematically in Proposition 4.2.2.)

After such a simplification we can revert to the original notation x, y, a_1 *etc.* Then let $\xi = \alpha + \beta y \in L$ with $\alpha, \beta \in K(x)$ so the Galois conjugate is $\bar{\xi} = \alpha + \beta(-y - a_1x - a_3)$. We must prove that if both

$$\begin{aligned} \xi + \bar{\xi} &= 2\alpha - \beta(a_1x + a_3) \qquad \text{and} \\ \xi\bar{\xi} &= \alpha^2 - \alpha\beta(a_1x + a_3) - \beta^2(x^3 + a_2x^2 + a_4x + a_6) \end{aligned}$$

are in $K[x]$ then $\alpha, \beta \in K[x]$.

First assume $\operatorname{char} K \neq 2$. Then 2 is invertible so we may replace y by $y + (a_1x + a_3)/2$ — in effect we can assume $a_1 = a_3 = 0$; and since the discriminant of the cubic is $\Delta/16 \neq 0$, the cubic has no repeated factor. Hence this case is included in part (b) of the previous proposition.

Now let $\operatorname{char} K = 2$. Since $\Delta = a_1^4b_8 + a_3^4 + a_1^3a_3^3 \neq 0$, therefore a_1 and a_3 are not both 0.

Subcase $a_1 = 0$, $a_3 \neq 0$: $\xi + \bar{\xi} = a_3\beta \in K[x]$ hence $\beta \in K[x]$. Now $\xi\bar{\xi} = \alpha^2 + \alpha\beta a_3 + \beta^2(x^3 + \cdots) \in K[x]$ gives an equation of integral dependence for α over $K[x]$, hence $\alpha \in K[x]$.

Final subcase $a_1 \neq 0$: Replacing x, y with

$$a_1^2x + a_3/a_1, \qquad a_1^3y + a_4/a_1 + a_3^2/a_1^3,$$

respectively, and dividing the resulting equation by a_1^6, we obtain

$$y^2 + xy = x^3 + a_2x^2 + a_6, \quad \text{so} \quad \Delta = a_6 \neq 0,$$

with new values for a_2, a_6. Thus our assumptions are that $\beta_1 := \xi + \bar{\xi} = \beta x \in K[x]$ and, discarding the terms $\beta^2(x^3 + a_2x^2)$ from $\xi\bar{\xi}$,

$$\delta := \alpha^2 + \alpha\beta x + \beta^2 a_6 \in K[x].$$

If α (resp. β) is in $K[x]$ this gives an equation of integral dependence of β (resp. α) over $K[x]$, hence both α, $\beta \in K[x]$ and we have the result.

There remains the possibility that $\alpha = \alpha_1/x$ and $\beta = \beta_1/x$ where $\alpha_1 = A_0 + A_1x + \cdots$ and $\beta_1 = B_0 + B_1x + \cdots$ are in $K[x]$ with $A_0B_0 \neq 0$. But the last inequality is contradicted by comparing the coefficient of x in $\alpha_1^2 + \alpha_1\beta_1 x + \beta_1^2 a_6 = \delta x^2$. ■

Referring to Proposition 2.2.11, let P_∞ denote the place corresponding to w_∞. Since the integral closure of $K[x]$ in L is the intersection of all valuation rings V_w for w in $\mathrm{gam}_K(L)$ except w_∞, the following is essentially equivalent to the proposition.

Corollary 2.2.13 *With the notation of the proposition, an element $\xi = \alpha + \beta y$ of L, where $\alpha, \beta \in K(x)$, has polar divisor $\mathsf{Pol}(\xi) = nP_\infty$, where n is a non-negative integer, iff α and β are polynomials in x satisfying*

$$\max\{2\deg(\alpha), 2\deg(\beta) + 3\} = n \quad (\text{where } \deg 0 = -\infty). \quad ■$$

This proposition and its converse that $K[x, y]$ is not integrally closed when $\Delta = 0$, at least when K is algebraically closed, exemplify well known general facts about "normality" in algebraic geometry; in fact a **normal domain** is by definition a noetherian Krull domain, equivalently, a noetherian integrally closed domain.

With E, K, L as in the proposition, let $w \in \mathrm{gam}_K L$.Then the residue field $\widetilde{w}$ is an extension of K, and the corresponding place $P_w : L \longrightarrow \widetilde{w}$ is **trivial** on K, *i.e.,* P_w is the identity map on K. Recall from §2.2.2 that

$$\deg w = \deg P_w = [\widetilde{w} : K] < \infty.$$

Consequently

$$P_w(x), P_w(y) \in \widetilde{w} \cup \{\infty\} \subset \overline{K} \cup \{\infty\},$$

where $\overline{K}$ denotes an algebraic closure of K.

Proposition 2.2.14 *Let E be an elliptic curve defined over the field K with function field $L = K(x, y)$, and $w \in \mathrm{gam}_K L$.*

(a) *w and P_w are uniquely determined by the values of $P_w(x)$ and $P_w(y)$ in $\overline{K} \cup \{\infty\}$.*

(b) *$(P_w(x), P_w(y)) \in E(\widetilde{w}) \supset E(K)$ (where (∞, ∞) is interpreted as O). Thus $P_w \leftrightarrow (P_w(x), P_w(y))$ establishes a canonical bijection between the rational places on L and the points on E defined over K.*

Proof. w_∞, which is characterized as the only member of $\mathrm{gam}_K L$ whose ring does not contain the coordinate ring $A = K[x, y]$, has degree 1 and corresponds to $O \in E(K)$.

For all other w we have $x, y \in V(w)$, and since P_w is a K-algebra homomorphism, therefore

(i) $(P_w(x), P_w(y))$ is in $E(\widetilde{w})$, and

(ii) $\forall \alpha \in A$, $P_w(\alpha)$ is uniquely determined by the values $P_w(x)$ and $P_w(y)$. The kernel M in A is a nonzero prime ideal. Since A is Dedekind, M is maximal and A_M is a valuation ring, which is therefore the ring V_w of w. Thus V_w, w and P_w are determined by $P_w(x)$ and $P_w(y)$.

Therefore, taking w of degree 1, $P_w \mapsto (P_w(x), P_w(y))$ is an injection from the set of rational places to $E(K)$.

Conversely let $(a, b) \in E(K)$. The ideal M generated by the two elements $x-a$, $y-b$ in A is maximal, hence the local ring A_M is the ring of a valuation w extending v_{x-a}. The place sends (x, y) to (a, b) and this implies that the degree is 1. (In general there are two places over v_{x-a}, the other one corresponding to $[-1](a, b) = (a, -b - a_1 a - a_3)$; but for points $(a, b) = [-1](a, b)$ of order 2, there is a unique ramified w extending v_{x-a}, as for the point O.) ■

If $P \in E(K)$, a **local parameter** or **uniformizer** at P is a uniformizer for the valuation, canonically associated with P. Thus $\pm x/y$ are local parameters at O since $w_\infty(\pm x/y) = -2 - (-3) = 1$, as we saw in Proposition 2.2.11.

Using Proposition 2.2.11(b) and reasoning as in the previous proof, we have

Proposition 2.2.15 *Let* $\mathrm{char}\, K \neq 2$ *and* $L = K(x, y)$ *where*

$$y^2 = f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0, \quad a_i \in K,$$

and the polynomial on the right has no repeated factors in $K[x]$. *Then the places* P *of* L *not at infinity*, i.e., *those not above* $P_{1/x}$ *on* $K(x)$, *are in bijection with the rational points* (a, b) *satisfying* $b^2 = f(a)$. ■

2.3 The group of reversible power series

Let A be a commutative ring and let $\mathcal{R}$ denote the ideal $TA[[T]]$ in the ring of formal power series $A[[T]]$. For $\alpha = a_0 + a_1 T + a_2 T^2 + \cdots \in A[[T]]$ and $\rho = r_1 T + \cdots \in \mathcal{R}$, we have functional composition

$$\alpha \circ \rho = a_0 + a_1(r_1 T + r_2 T^2 + \cdots) + a_2(r_1 T + \cdots)^2 + \cdots$$

$$= a_0 + a_1 r_1 T + (a_1 r_2 + a_2 r_1^2) T^2 + (a_1 r_3 + 2a_2 r_1 r_2 + a_3 r_1^3) T^3 + \cdots.$$

If both α and ρ are in $\mathcal{R}$ then so is $\alpha \circ \rho$. But note that $\circ$ is always a noncommutative operation on $\mathcal{R}$; for example,

$$(T + T^3) \circ (T + T^2) - (T + T^2) \circ (T + T^3) = T^4 + 3T^5.$$

The following rules are easily verified: for $\alpha, \beta \in A[[T]]$ and $\rho, \sigma \in \mathcal{R}$,

$$\begin{aligned}(\alpha \circ \rho) \circ \sigma &= \alpha \circ (\rho \circ \sigma),\\ (\alpha + \beta) \circ \rho &= \alpha \circ \rho + \beta \circ \rho,\\ (\alpha\beta) \circ \rho &= (\alpha \circ \rho)(\beta \circ \rho).\end{aligned}$$

Clearly these results are sufficient to prove the

Proposition 2.3.1 *With $+$ inherited from $A[[T]]$ and $\circ$ serving as multiplication, $\mathcal{R}$ is a noncommutative ring with T acting as "1".*

Next, d/dT denotes as usual the standard A-linear derivation on $A[[T]]$:

$$\frac{d}{dT}(a_0 + a_1T + a_2T^2 + \cdots) = a_1 + 2a_2T + \cdots.$$

We also use the prime notation $d\alpha/dT = \alpha'$, etc., as well as the symbol D for d/dT. The simple verifications of the following three familiar rules is again left to the reader: for $\alpha, \beta \in A[[T]]$ and $\rho \in \mathcal{R}$,

$$\begin{aligned}D(\alpha + \beta) &= \alpha' + \beta',\\ D(\alpha\beta) &= \alpha'\beta + \alpha\beta',\\ D(\alpha \circ \rho) &= (\alpha' \circ \rho)\rho' \quad \text{(the chain rule)}.\end{aligned}$$

As usual, $\mathcal{R}^*$ denotes the group of units of the ring $\mathcal{R}$.

Proposition 2.3.2 (a) *The group $\mathcal{R}^*$ consists of series of the form $r_1T + \cdots$ with $r_1 \in A^*$.*

(b) *$\mathcal{R}^*$ acts on $A[[T]]$ making $A[[T]]$ a right $\mathcal{R}^*$-module*, i.e., *a module over the integral group ring $\mathbf{Z}[\mathcal{R}^*]$ as follows. For $\alpha \in A[[T]]$ and $\rho \in \mathcal{R}^*$, using exponential notation,*

$$\alpha^\rho = (\alpha \circ \rho)\rho'.$$

Remarks. The elements in $\mathcal{R}^*$ are called **reversible power series**, and $\mathcal{R}^*$ is the **group of reversible power series**.

$A[[T]]$ has the simpler right $\mathcal{R}^*$-module structure given by $\alpha \circ \rho$. However the action of $\mathcal{R}^*$ given in the proposition will be the one of interest to us later.

Proof. (a) Given $\rho = r_1T + \cdots \in \mathcal{R}^*$ with inverse $\sigma = s_1T + \cdots$, we see from

$$\rho \circ \sigma = r_1s_1T + (r_1s_2 + r_2s_1^2)\,T^2 + \cdots = T$$

that $r_1 \in A^*$. Conversely if $r_1 \in A^*$ then by comparing coefficients in this equation we can solve recursively for $s_1, s_2, \ldots$, obtaining $\rho \circ \sigma = T$. By the same token we find ρ' such that $\sigma \circ \rho' = T$, and applying the associative law to $\rho \circ \sigma \circ \rho'$, we find $\rho = \rho'$, and therefore $\rho \in \mathcal{R}^*$.

(b) We verify the module axioms using the rules above:

$$(\alpha+\beta)^\rho = ((\alpha+\beta)\circ\rho)\rho' = (\alpha\circ\rho+\beta\circ\rho)\rho' = \alpha^\rho+\beta^\rho;$$

$$(\alpha^\rho)^\sigma = ((\alpha\circ\rho)\rho')^\sigma = ((\alpha\circ\rho)\circ\sigma)\,(\rho'\circ\sigma)\sigma'$$

coincides with

$$\alpha^{\rho\circ\sigma} = (\alpha(\rho\circ\sigma))\,(\rho\circ\sigma)'. \qquad \blacksquare$$

In order to avoid confusion with the ordinary ring inverse we denote the group inverse of ρ in $\mathcal{R}^*$ by $\rho^{(-1)}$, and in general any power by $\rho^{(n)}$. When it is necessary to make it clear what ring is being discussed we denote the group $\mathcal{R}^*$ by $\mathcal{R}^*(A)$ or $\mathcal{R}^*(A,T)$.

Elements of the form $T+r_2T^2+\cdots$ form a subgroup of $\mathcal{R}^*$. For such elements the begining of "generic reversion" is

$$\begin{aligned}(T+r_2T^2+\cdots)^{(-1)} \quad = \quad & T-r_2T^2+(2r_2^2-r_3)\,T^3+\\ & (-5r_2^3+5r_2r_3-r_4)\,T^4+\\ & (14r_2^4-21r_2^2r_3+6r_2r_4+3r_3^2-r_5)\,T^5+\cdots.\end{aligned}$$

We will also need formal integration:

$$\int(a_0+a_1T+\cdots)\,dT = a_0T+\frac{a_1}{2}T^2+\cdots.$$

This brings up the matter of having inverses of integers available. Consider the canonical ring homomorphisms

$$\mathbf{Z}\xrightarrow{i}A\xrightarrow{j}A\otimes_{\mathbf{Z}}\mathbf{Q}.$$

If $\ker i = n\mathbf{Z}$, $n\geq 0$, we call n the **characteristic** of A; thus characteristic 0 means that i is injective. A is **flat**, more precisely *flat as a* $\mathbf{Z}$*-module*, when j is injective; this is equivalent to (since $\mathbf{Z}$ is a PID) the additive group of A being torsion-free. Flat implies characteristic 0, but not conversely as is shown by the example $\prod\mathbf{F}_p$, and sometimes in discussions of formal groups these two conditions are confused. (In this connection we note that [Haz78] *defines* "characteristic 0" to mean flat.) To say that A is a $\mathbf{Q}$-algebra is equivalent to saying that j is an isomorphism. For convenience we denote the tensor product by $A_{\mathbf{Q}}$; and when A is flat we regard it as a subring of $A_{\mathbf{Q}}$.

Proposition 2.3.3 *Let A be flat. Then the set of $f\in A_{\mathbf{Q}}[[T]]$ of the form*

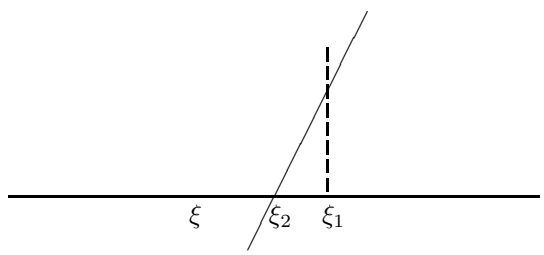
$$f=\sum_{n=1}^{\infty}\frac{a_n}{n!}T^n, \qquad a_n\in A,\ a_1\in A^*$$

is a subgroup of $\mathcal{R}^*(A_{\mathbf{Q}})$. In particular, the coefficients of $f^{(-1)}$ can be written in the form $a_n'/n!$.

Proof. Let $\mathcal{R}^*{}_!$ denote the set of such f. Clearly $\mathcal{R}^*{}_!$ contains the group identity element T. Secondly, let f and $g = b_1T + b_2T^2/2! + \cdots$ be members of $\mathcal{R}^*{}_!$ and write $h = f \circ g = c_1T + c_2T^2/2! + \cdots$. Then $c_1 = a_1b_1 \in A^*$ and differentiating $h(T) = f(g(T))$ n times gives the n-th derivative D^nh as a polynomial with integer coefficients in the derivatives D^if, D^jg, $1 \le i, j \le n$. Substituting $T = 0$ thus gives c_n as a polynomial with integer coefficients in a_i, b_j, which proves that $f \circ g \in \mathcal{R}^*{}_!$. Finally, suppose $g = f^{(-1)}$ so that $h = T$. Then these polynomial relations for $n > 0$ are of the form $0 = c_n = a_1b_n +$ (monomials involving the a_i and b_j with $j < n$). Since $a_1 \in A^*$, this proves by induction that $b_n \in A$. ■

2.4 Hensel's lemma

Newton's method for the approximate numerical calculation of a root ξ of a differentiable function f is best explained by a picture:



If ξ_1 is an approximation to the root ξ then intersecting the tangent to the curve $y = f(x)$ at $x = \xi_1$ with the ξ-axis gives the presumably better approximation

$$\xi_2 = \xi_1 - \frac{f(\xi_1)}{f'(\xi_1)}.$$

Here f' denotes the derivative of f and an obviously minimal requirement for this to be useful is that $f'(\xi_1) \neq 0$. In favorable circumstances the formula can be iterated to give a sequence $(\xi_1, \xi_2, \ldots)$ converging to ξ. In fact this algorithm, called **Newton's method** has wide application in both the archimedean cases $\mathbf{R}$ and $\mathbf{C}$ and in the nonarchimedean or ultrametric cases which include the case of a discrete valuation.

In the latter case problems of convergence tend to be quite trivial. For example a series $x_1 + x_2 + \cdots$ of elements in a complete discretely valued field K converges to an element in K iff $\lim |x_i| = 0$. Thus one easily proves the following fact. Let V denote the valuation ring in K and let $f(T) = a_0 + a_1T + \cdots \in V[[T]]$.

If $f(t)$ converges for a particular $t \in K$ (which is certainly true if $|t| < 1$), then the derivative $f' = Df = a_1 + 2a_2T + \cdots$ also converges for $T = t$.

In fact we can generalize this to the **Taylor series**

$$f(T+\Delta) = f(T) + f_1(T)\Delta + f_2(T)\Delta^2 + \cdots \in V[[T]][[\Delta]]$$

where $f_1 = Df$ and in general $f_i \in V[[T]]$ has coefficients which are linear combinations with integer binomial coefficients of the a_j. (Obviously this Taylor expansion can be defined for a series over any commutative ring V; when V is flat (in the present situation that means simply that char $K = 0$) one has the formulas $f_i = D^i f/i!$, but the series is properly defined in any case.) If $f(T)$ converges for $T = t$ then so do all the $f_i(t)$, and the Taylor series $f(t+\delta)$ converges for all δ satisfying $|\delta| < 1$. Of course if f is a polynomial in T then there is no restriction on t and δ.

The following result is known as **Hensel's lemma** (for discrete valuations); for later applications we have made a mild generalization from the usual case of finding a root of a polynomial to finding a root of a power series.

Proposition 2.4.1 *Let K be complete with respect to the discrete valuation v, let V be the valuation ring, let $f \in V[[T]]$ and let $\xi_1 \in V$ be such that $f(\xi_1)$ converges and*

$$|f(\xi_1)| < |f'(\xi_1)|^2.$$

Then Newton's method produces a convergent sequence $(\xi_1, \xi_2, \ldots)$ whose limit ξ is a root of f in V. The progress of the convergence is measured by

$$|\xi - \xi_i| = |f(\xi_i)/f'(\xi_1)|.$$

Moreover, ξ is the only root of f in the closed disc

$$|\xi - \xi_1| \leq |f(\xi_1)/f'(\xi_1)|.$$

Remarks.

(i) The convergence of the algorithm is quadratic rather than merely linear as will become clear in the proof. Roughly speaking this means that the accuracy ("number of digits") doubles at each iteration. A "save work simplification" such as

$$\xi_{i+1} = \xi_i - \frac{f(\xi_i)}{c}$$

where the denominator c has the constant value $f'(\xi_1)$ produces a sequence that converges only linearly to the root, and is much less efficient.

(ii) A corollary is that if $f \in V[T]$ is a monic polynomial whose reduction $\overline{f}$ to the residue field k has a simple root $\overline{\xi} \in k$, then the factorization $\overline{f} = (T - \overline{\xi})\overline{g}$ in $k[T]$ can be lifted to $V[T]$; for if $\xi_1 \in V$ is any lifting of $\overline{\xi}$, then $v(f(\xi_1)) > 0$ and $v(f'(\xi_1)) = 0$ since the root is simple.

In fact V is *henselian* in the sense of the following definition (*cf.* [Mil80], p.32): a local noetherian domain R with maximal ideal m and residue field $k = R/m$ is **henselian** if for every monic polynomial $f \in R[T]$ and factorization

$\overline{f} = g_0 h_0$ in $k[T]$ where g_0 and h_0 are monic and coprime, there exist monic $g, h \in R[T]$ such that $f = gh$ and $\overline{g} = g_0$, $\overline{h} = h_0$.

Proof. The assumption $|f(\xi_1)| < |f'(\xi_1)|^2$ guarantees that $f'(\xi_1) \neq 0$ and Newton's method can be written as $\xi_2 = \xi_1 + \delta_1$ where $f(\xi_1) + \delta_1 f'(\xi_1) = 0$. Let

$$f(T + \Delta) = f(T) + f_1(T)\Delta + \cdots$$

be the Taylor expansion of f, so $f_1 = f'$. Each $f_j(\xi_1) \in V$, and $|\delta_1| < |f'(\xi_1)| \leq 1$, so $\delta_1,\ \xi_2 \in V$. Substituting $T = \xi_1,\ \Delta = \delta_1$ in the Taylor expansion and writing $C = |f(\xi_1)/f'(\xi_1)^2|$ yields, using $f(\xi_1) + \delta_1 f'(\xi_1) = 0$,

$$|f(\xi_2)| \leq \max\{|f_2(\xi_1)\delta_1^2|, \ldots, |f_n(\xi_1)\delta_1^n|, \cdots\} \leq |\delta_1|^2 = C|f(\xi_1)|,$$

where $C < 1$. The Taylor expansion for $f' = g$, say, gives

$$|f'(\xi_2)| = |f'(\xi_1) + \delta_1 g_1(\xi_1) + \delta_1^2 g_2(\xi_1) \cdots| = |f'(\xi_1)|$$

since $|\delta_1^j g_j(\xi_1)| \leq |\delta_1| < |f'(\xi_1)|$ for $j > 0$.

Thus $|f(\xi_2)/f'(\xi_2)^2| < C^2$ and the process can be repeated. By induction,

$$\begin{aligned} |f'(\xi_i)| &= |f'(\xi_1)| \\ |f(\xi_i)/f'(\xi_i)^2| &\leq C^{2^{i-1}} \\ |f(\xi_{i+1})| &\leq C^{2^{i-1}}|f(\xi_i)| \end{aligned}$$

where $\xi_{i+1} = \xi_i + \delta_i$ and δ_i is defined by $f(\xi_i) + \delta_i f'(\xi_i) = 0$. (Of course the process stops if $f(\xi_i) = 0$.) A simple induction gives

$$|\delta_i| \leq C^{2^{i-1}-1}|\delta_1|.$$

If $i < j$ then

$$|\xi_i - \xi_j| = |\delta_i + \cdots \delta_{j-1}| \leq C^{2^{i-1}-1}|\delta_1|.$$

Since $C < 1$ this shows that $(\xi_1, \xi_2, \ldots)$ is cauchy, say $\lim \xi_i = \xi$, and since $|f(\xi_i)| = |\delta_i f'(\xi_1)| \to 0$ and a powerseries function is continuous on a closed disc in its domain of convergence, therefore $f(\xi) = 0$. Since $|\delta_i| > |\delta_{i+1}| > \cdots$,

$$|\xi - \xi_i| = \lim_{j\to\infty} |\delta_i + \cdots + \delta_j| = \lim_{j\to\infty} |\delta_i| = |\delta_i|.$$

Suppose $\xi' = \xi + \eta \in V$ is another root of f satisfying $|\xi' - \xi_1| \leq |f(\xi_1)/f'(\xi_1)|$. Then

$$\begin{aligned} |\eta| &= |\xi' - \xi_1 + \xi_1 - \xi| \leq \max\{|\xi' - \xi_1|, |\xi_1 - \xi|\} \\ &\leq |f(\xi_1)/f'(\xi_1)| < |f'(\xi_1)| = |f'(\xi)|, \end{aligned}$$

the last inequality being our basic assumption on ξ_1. Hence for $i > 1$, and assuming $\eta \neq 0$, we have $|f_i(\xi)\eta^i| \leq |\eta^2| < f'(\xi)\eta|$, which is incompatible with the Taylor expansion

$$0 = f(\xi') = f(\xi + \eta) = f'(\xi)\eta + f_2(\xi)\eta^2 + \cdots. \quad \blacksquare$$

Distinct $\xi_1 \in V$ can lead to the same root ξ. For instance

$$f \equiv X^2 - 1 \bmod 16$$

has the four roots $\pm 1, \pm 7 \in \mathbf{Z}/16\mathbf{Z}$, but each of these must lift to one of the two roots ± 1 of f in $\mathbf{Q}_2$. In fact 7 lifts to -1 since $|7+1| = |f(7)/f'(7)|$.

The 'factor theorem' for polynomials, that a root corresponds to a linear factor, extends to power series:

Proposition 2.4.2 *Let V be a complete discrete valuation ring, let $f \in V[[T]]$ and let $\xi \in V$ be a root of f : the series $f(\xi)$ converges to 0. Then $f(T) = (T - \xi)g(T)$ where $g(T) \in V[[T]]$.*

Proof. We have $g(T) = f(T)/(T - \xi) = b_0 + b_1 T + \cdots \in K[[T]]$, and we wish to show that $b_i \in V$. This is clear if $v(\xi) = 0$ since then $T - \xi$ is an invertible element of $V[[T]]$ (and we don't need $f(\xi) = 0$):

$$g(T) = \frac{f(T)}{-\xi(1 - T/\xi)} = \frac{f(T)}{-\xi}\left(1 + \frac{T}{\xi} + \frac{T^2}{\xi^2} + \cdots\right) \in V[[T]].$$

Thus suppose $v(\xi) > 0$ and let $f = \sum a_i T^i$. Then

$$\begin{aligned} g(T) = \frac{f(T) - f(\xi)}{T - \xi} &= \frac{1}{T - \xi}\sum_{i=0}^{\infty} a_i(T^i - \xi^i) \\ &= \sum_{i=1}^{\infty} a_i(T^{i-1} + T^{i-2}\xi + \cdots + \xi^{i-1}), \end{aligned}$$

so
$$b_i = a_{i+1} + a_{i+2}\xi + a_{i+3}\xi^2 \cdots ,$$

and for each i the series converges to an element of V. $\blacksquare$

The ring $V[[T]]$ is in fact a UFD; *cf.* [BAC7, p.42].

Corollary 2.4.3 (ascribed to Strassmann in [Cas78, p.52])

Let $f = \sum a_i T^i \in V[[T]]$ and suppose $|a_i| \longrightarrow 0$. If N denotes the largest subscript such that $|a_N| = \max\{|a_i|\}$, then f has at most N roots in V. It follows that as T ranges over V, f takes any value at most a finite number of times, and in particular f cannot be periodic.

Proof. If $\xi \in V$ is a root of f then $g = f/(T-\xi) = b_0 + b_1 T + \cdots \in V[[T]]$ and

$$b_i = a_{i+1} + a_{i+2}\xi + \cdots \longrightarrow 0.$$

Also $N-1$ is the largest subscript such that $b_{N-1} = \max\{|b_i|\}$. Thus the process can be repeated on g. ■

For example the polynomial $2a + T + 2bT^n$, where $a, b \in \mathbf{Z}$ and $n \geq 2$, has at most one integral root since it has at most one root in $\mathbf{Z}_2$ (and therefore exactly one root in $\mathbf{Z}_2$ since Hensel can be applied with $T_1 = -2a$ as the starting 2-adic approximation.)

2.4.1 An application to P-adically reversible series

Let A be a commutative ring and P an ideal in A that is **topologically nilpotent**, *i.e.,*

$$\bigcap_{n=1}^{\infty} P^n = 0. \tag{$*$}$$

With $\{a + P^n : n = 1, 2, \ldots\}$ taken as a neighborhood basis at $a \in A$, A is a (Hausdorff) topological ring. When $P = 0$ the topology is discrete.

We define $v : A \longrightarrow \{0, 1, 2, \ldots\} \cup \{\infty\}$ by declaring $v(a) = 0$ if $a \notin P$, $v(0) = \infty$ and otherwise $v(a)$ is the smallest n such that $a \notin P^{n+1}$. Clearly

$$v(a+b) \geq \min\{v(a), v(b)\}, \quad v(ab) \geq v(a) + v(b).$$

Lemma 2.4.4 (Krull, Chevalley) *Let A, P satisfy $(*)$. One has $v(ab) = v(a) + v(b)$ for all $a, b \in A$ iff*

(i) *A is an integral domain, and*

(ii) *P is a prime ideal.*

Then, assuming $P \neq 0$, v extends to a discrete valuation on the quotient field of A if we define $v(a/b) = v(a) - v(b)$ for $a, b \in A$ and $b \neq 0$.

Proof. If $a \neq 0, b \neq 0, ab = 0$ then $v(ab) = \infty > v(a) + v(b)$; if $a \notin P, b \notin P, ab \in P$ then $v(ab) > 0 = v(a) + v(b))$; and if $a/b = c/d$ then $ad = bc$, hence if v acts like log, then $v(a) + v(d) = v(b) + v(c)$, and $v(a/b)$ is well-defined.

Conversely assume that P is a prime ideal in the integral domain A. By Corollary 2.1.3, there exists a generalized valuation ring V lying between A and its quotient field Q whose maximal ideal intersected with A is P. The valuation ring is discrete because of the special assumption $(*)$. Hence v acts like log. ■

We now assume in addition to $(*)$ that A is ***P*-adically complete**, *i.e.,* the canonical ring homomorphism

$$A \xrightarrow{\sim} \varprojlim A/P^n \tag{$**$}$$

is an isomorhism. Then a series $a_1 + a_2 + \cdots$, where $a_i \in A$, converges to an element of A iff $v(a_i) \longrightarrow \infty$ (since then the partial sums form a consistent sequence mod P^n for ever increasing n). It follows that

$$a \in A^* \quad \text{and} \quad x \in P \Longrightarrow a + x \in A^*.$$

We also assume

$$A \text{ is an integral domain, and } P \text{ is a prime ideal.} \qquad (***)$$

For A, P satisfying $(*)$, $(**)$ and $(***)$, we define

$$\mathcal{R}_P = \mathcal{R}_P(A, T) = \{a_0 + a_1 T + \cdots \in A[[T]] : a_0 \in P\}.$$

Note that $\mathcal{R}$ defined in §2.3 is $\mathcal{R}_0$ — the **discrete case** — except that here we assume that A is an integral domain; we will have no need for greater generality. With that minor proviso, the following generalizes Propositions 2.3.1 and 2.3.2.

Proposition 2.4.5 (a) *Let A, P satisfy $(*)$, $(**)$ and $(***)$. Then $\mathcal{R}_P$, with the $+$ of $A[[T]]$ and $\circ$ as operations, is a noncommutative ring with T serving as the ring "1".*

(b) *The group of units $\mathcal{R}_P^*$ consists of the series of the form $r_0 + r_1 T + r_2 T^2 + \cdots$ where $r_0 \in P$ and $r_1 \in A^*$.*

(c) *$A[[T]]$ is a right $\mathcal{R}_P^*$-module, where, using exponential notation,*

$$\alpha^\rho = (\alpha \circ \rho)\rho', \quad \textit{for} \;\; \alpha \in A[[T]],\, \rho \in \mathcal{R}_P^*.$$

Remark. The series in $\mathcal{R}_P^*$ are said to be ***P*-adically reversible**, and $\mathcal{R}_P^*$ is the **group of *P*-adically reversible series**.

Proof. First one checks that the composition $\alpha \circ \rho$ of $\alpha = a_0 + a_1 T + \cdots \in A[[T]]$ and $\rho = r_0 + r_1 T + \cdots \in \mathcal{R}_P$ is defined. If we write $\alpha \circ \rho = \beta$, then

$$\begin{aligned} b_0 &= a_0 + a_1 r_0 + a_2 r_0^2 + \cdots, \\ b_1 &= a_1 r_1 + 2a_2 r_0 r_1 + 3a_3 r_0^2 r_1 + \cdots, \\ b_2 &= a_1 r_2 + a_2(2r_0 r_2 + r_1^2) + \cdots, \\ \text{etc.} & \end{aligned}$$

One easily sees that the series for all the b_i converge, hence $\alpha \circ \rho \in A[[T]]$, and that $\mathcal{R}_P$ and $\mathcal{R}_P^*$ are closed under the operation $\circ$.

The only detail that is not entirely elementary is the existence of inverses. As in the proof of Proposition 2.3.2, we need only consider one-sided inverses: given $\alpha = a_0 + a_1 T + \cdots$ with $a_0 \in P$ and $a_1 \in A^*$, we wish to calculate $\rho \in \mathcal{R}_P$ so that $\beta = \alpha \circ \rho = T$. Thinking of $b_0 = \alpha(r_0)$ as a function of r_0, we have $v(\alpha(-a_0 a_1^{-1})) = v(a_2 a_0^2 a_1^{-2} + \cdots) \geq 2$ and $v(\alpha'(-a_0 a_1^{-1})) = v(a_1 + \cdots) = 0$. Thus Hensel can be applied to the equation $\alpha(r_0) = 0$ with the initial

approximation $r_0 = -a_0 a_1^{-1}$ to obtain r_0. We note that inductively at each step of Newton's method, $\alpha'(\xi_i) = a_1 + 2a_2\xi_i + \cdots \in A^*$, hence $\xi_{i+1} \in P$ and therefore in the limit $r_0 \in P$. (It is necessary to make this remark since the valuation ring involved may be distinctly larger than A.) Then $r_1, r_2, \ldots$ are calculated directly one after the other:

$$\begin{aligned} b_1 &= r_1(a_1 + 2a_2r_0 + 3a_3r_0^2 + \cdots) = 1, \\ b_2 &= r_2(a_1 + 2a_2r_0 + 3a_3r_0^2 + \cdots) + (a_2r_1^2 + 3a_3r_0r_1^2 + \cdots) = 0, \end{aligned}$$

etc.

The successive coefficients of $r_1, r_2, \ldots$ are of the form $a_1 + x$ where $x \in P$, hence they are in A^*.

The $\mathcal{R}_P^*$-module structure of $A[[T]]$ can now be verified as in the proof of Proposition 2.3.2. ■

Corollary 2.4.6 *With A, P as in the proposition and $\alpha \in \mathcal{R}_P^*$, the equation $\alpha(T) = 0$ has a unique root $T = r_0$ in P. This root will be denoted N_α: $\alpha(N_\alpha) = 0$.*

The uniqueness of N_α is a consequence of the uniqueness of the group inverse and the fact that $b_1, b_2, \ldots$ are determined once b_0 is specified.

Here is an example: $R[[S,T]]$, the ring of power series in two variables over the integral domain R, regarded as the ring of series in T over $A = R[[S]]$, with the prime ideal $P = SA$. Then $\mathcal{R}_P^*$ consists of all series $\sum_{i,j=0}^{\infty} r_{ij}S^iT^j$ where $r_{ij} \in R$, $r_{00} = 0$ and $r_{01} \in R^*$. For example, let $\alpha = aS + bT$ and $\beta = cS + dT$, where $a, c \in R$ and $b, d \in R^*$. Then

$$\alpha \circ \beta = (a + bc)S + bdT, \quad \alpha^{(-1)} = -\frac{a}{b}S + \frac{1}{b}T, \quad N_\alpha = -ab^{-1}S.$$

2.5 Applications to elliptic curves

2.5.1 Infinitesimal shifts

We introduce a technical device that will be useful in the proof of the proposition in the next section and elsewhere.

Let E be an elliptic curve defined over the field K by the nonsingular Weierstrass equation

$$F(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 = 0,$$

and suppose $P = (a,b)$ is point in $E(K)$ that is not in $E(K)[2]$. Since $-(a,b) = (a, -b - a_1a - a_3) \neq (a,b)$, the partial derivative $F_y(a,b) \neq 0$. Let t be an indeterminate, so that the power series field $K((t))$ is complete with respect to

the t-adic valuation with ring $V = K[[t]]$. We assign x the value $a+t$ and apply Hensel to $F = 0$ regarded as a function of the variable y starting at $y = b$:

$$v_t(F(a+t,b)) \geq 1, \quad v_t(F_y(a+t,b) = 0.$$

This leads to a solution $y = b+\beta$, where $\beta = \beta_1 t + \beta_2 t^2 + \cdots \in tK[[t]]$, and thus we obtain a point $P' = (a+t, b+\beta) \in E(K((t)))$, which we call an **infinitesimal shift** of P — the shift by the **infinitesimal** t.

If two different shifts of P are needed, another infinitesimal can be introduced; alternatively, one might shift the x-coordinate by t and also by t^2, for instance.

Lemma 2.5.1 *With the above notation,*

(a) $\forall Q \in E(K)$ *except* $Q = -P$, *the point* $P' + Q = (x_3, y_3)$ *is* t*-integral*, i.e., $x_3, y_3 \in K[[t]]$;

(b) $[2]P' \notin E(K)$.

Proof. (a) First we observe that this is true when $Q = O$ since then $x_3 = a+t$ and $y_3 = b+\beta$ are in $K[[t]]$. Taking $Q = (x_1, y_1)$ and $P' = (a+t, b+\beta) = (x_2, y_2)$ in the addition formula in Proposition 1.7.1, we see that we wish to prove

$$v_t(\lambda) = v_t\left(\frac{b+\beta_1 t + \cdots - y_1}{a+t-x_1}\right) \geq 0.$$

The only time this is not true is when $x_1 = a$ and $y_1 \neq b$, *i.e.,* when $Q = -P$. Intuitively, this says that when we apply the ring homomorphism $K[[t]] \longrightarrow K$ where $t \mapsto 0$, the equation $P' + Q = R$ maps to $P + Q = R_0$ where $R_0 \neq (\infty, \infty)$ except when $Q = -P$.

(b) By definition, $x(P') = x(P) + t$ is not algebraic over K. It follows by Proposition 1.7.3 that $[2]P' \notin E(K)$. ■

2.5.2 Reduction mod π: a first look

Let v be a valuation on the field K with ring V, uniformizer π, and residue field k. Recall from §2.1.2 that each point P in projective space $\mathbf{P}^n(K)$ has a set of v-proper coordinates $(X_0, \ldots, X_n)$, *i.e.,* $\min\{v(X_i)\} = 0$, and they are unique up to multiplication by a unit $u \in V^*$. Let the canonical surjection $V \longrightarrow k$ be denoted $a \mapsto \widetilde{a}$. If $u \in V^*$ then $\widetilde{u} \in k^*$, hence $P = (X_0, \ldots) \mapsto \widetilde{P} = (\widetilde{X}_0, \ldots)$ is a well defined map

$$\mathbf{P}^n(K) \longrightarrow \mathbf{P}^n(k),$$

called **reduction mod $\boldsymbol{\pi}$**. (Since the element π of K could be simultaneously a uniformizer for other valuations on K, the terminology is not accurate; but in practice the context, or clarification if necessary, makes the meaning unambiguous.) Obviously this map is surjective.

Now suppose that E is an elliptic curve defined by a nonsingular Weierstrass equation

$$F = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0,$$

with all the coefficients $a_i \in V$. We denote by $\widetilde{F}$ the reduction of $F \bmod \pi$:

$$\widetilde{F} = Y^2Z + \widetilde{a_1}XYZ + \cdots - \widetilde{a_6}Z^3,$$

which is a Weierstrass equation over k. Affine equations can be reduced in the same way:

$$\text{if} \quad F = y^2 + a_1xy + \cdots - a_6, \quad \text{then} \quad \widetilde{F} = y^2 + \widetilde{a_1}xy + \cdots - \widetilde{a_6}.$$

Since $V \longrightarrow k$ is a ring homomorphism, therefore the quantities associated in §1.1 to $\widetilde{F}$ are $\widetilde{b_2} = \widetilde{a_1}^2 + 4\widetilde{a_2}$, *etc.*,, and in particular, the discriminant of $\widetilde{F}$ is $\widetilde{\Delta}$. Since all the $a_i \in V$, and Δ is a polynomial in the a_i with integer coefficients, therefore $v(\Delta) \geq 0$ with equality iff $\widetilde{\Delta} \neq 0$, *i.e.,* iff $\widetilde{F}$ represents an elliptic curve over k.

If $\widetilde{\Delta} \neq 0$ we define the **reduction mod $\boldsymbol{\pi}$ of $\boldsymbol{E}$** to be the elliptic curve over k defined by $\widetilde{F} = 0$, and denoted $\widetilde{E}$.

However $\widetilde{F}$ *must not be taken as the definition of the reduction of* E when $\widetilde{\Delta} = 0$. The problem is that F may not be "v-minimal"; the relevant definition, which is not as simple as that of v-proper, will be discussed in Chapter 5. Using that,

$$\boxed{\text{the general definition of } \widetilde{E} \text{ is given in §7.1}}$$

Let us consider an example. If v is the 2-adic valuation on $\mathbf{Q}$ then $y^2 = x^3 + 5x^2 + 8x + 16$ ($\Delta = -2^{12}15$) reduces to the singular equation $y^2 = x^3 + x^2$. However if we first substitute $x = 4x'$, $y = 8y' + 4x' + 4$, and divide by 64, we obtain the equation

$$y'^2 + x'y' + y' = x'^3 + x'^2, \quad \Delta = -15. \qquad \textbf{A15}$$

The present definition now applies: the reduction mod 2 of **A15** is

$$\widetilde{E} : y^2 + xy + y = x^3 + x^2 \ /\mathbf{F}_2.$$

Proposition 2.5.2 *Let v be a valuation on the field K with valuation ring V, residue field k and uniformizer π. Let E be an elliptic curve given by a nonsingular Weierstrass equation F defined over V, and suppose that $v(\Delta) = 0$. Then the reduction* $\bmod \pi$ *map $\mathbf{P}^2(K) \longrightarrow \mathbf{P}^2(k)$ induces a group homomorphism $E(K) \longrightarrow \widetilde{E}(k)$ with kernel $E_1(K)$ (the 1-st term in the v-adic filtration).*

Remarks. It follows that $E_1(K)$ is a subgroup of $E(K)$ when $v(\Delta) = 0$. This will be superseded by Proposition 2.6.7 where we prove that all terms in the v-adic filtration are subgroups regardless of the value of $v(\Delta)$; *cf.* the discussion in §2.1.2.

We will take F in projective form, but intuitively the affine form gives the right result: if $(x, y) \in E_1(K)$, then x and y have π in their denominators, and since $\pi \mapsto 0$ in the reduction, therefore $(x, y) \mapsto (\infty, \infty) =$ the point O at infinity.

Proof. A point in $P \in E(K)$ with v-proper coordinates (X, Y, Z) reduces to $O = (0, 1, 0)$ iff $v(X) > 0$ and $v(Z) > 0$. By definition, this is equivalent to $P \in E_1(K)$.

If $P_1 + P_2 + P_3 = O$, where $P_i \in E(K)$, we wish to prove that

$$\widetilde{P_1} + \widetilde{P_2} + \widetilde{P_3} = O. \tag{¶}$$

Now $\widetilde{O} = O$ and, since $-(X, Y, Z) = (X, -Y - a_1X - a_3Z, Z)$, therefore $\widetilde{-P} = -\widetilde{P}$. Thus (¶) is true if some $P_i = O$, hence we can assume that no $P_i = O$.

Consider any line $L : aX + bY + cZ = 0$ in $\mathbf{P}^2(K)$. The coefficients can be chosen so that $\min\{v(a), v(b), v(c)\} = 0$, and then $\widetilde{L} : \widetilde{a}X + \widetilde{b}Y + \widetilde{c}Z = 0$ is a line in $\mathbf{P}^2(k)$. If L meets $E(K)$ in the points P_i then, since $V \longrightarrow k$ is a ring homomorphism, $\widetilde{L}$ meets $\widetilde{E}(k)$ in the points $\widetilde{P_i}$. This proves (¶) when the $\widetilde{P_i}$ are distinct.

Suppose $\widetilde{P_1} = \widetilde{P_2}$, and suppose to begin with that P_1 and P_2 are not both points of order 2, say $P_1 \notin E(K)[2]$. We replace P_1 with an infinitesimal shift P_1' and define $P_3' = -P_1' - P_2$. Extend v to $K((t))$ by Proposition 2.1.7 so that π is still a uniformizer, the reduction homomorphism is now $V((t)) \longrightarrow k((t))$ and $\widetilde{P_1'}$ is an infinitesimal shift of $\widetilde{P_1}$. Using part (b) of the previous lemma, we see that no two of $\widetilde{P_1'}$, $\widetilde{P_2}$, $\widetilde{P_3'}$ are equal, hence

$$\widetilde{P_1'} + \widetilde{P_2} + \widetilde{P_3'} = O. \tag{¶$'$}$$

Since no $P_i = O$, by (a) of the lemma, all of P_1', P_2 and P_3' are t-integral, and therefore the same is true of the three points in (¶$'$). This allows us to apply the substitution homomorphism $t \mapsto 0$, which results in (¶).

Finally, consider the case where P_1 and P_2 are points of order 2. Then $P_1 \neq P_2$ since $P_3 \neq O$. Therefore $\operatorname{char} K \neq 2$ (*cf.* Corollary 1.7.7), and the P_i are the three distinct points of order 3. From $P_i = -P_i$ and $\widetilde{-P_i} = -\widetilde{P_i}$, it follows that $\widetilde{P_i} \in \widetilde{E}(k)[2]$, and by assumption, $\widetilde{P_1} = \widetilde{P_2}$. Thus we wish to show that $\widetilde{P_3} = O$. In x, y-coordinates, $P_i = (e_i, g_i)$ where $g_i = -(a_1e_i + a_3)/2$. The Weierstrass equation in b-form is

$$\eta^2 = (x - e_1)(x - e_2)(x - e_3) = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4}, \tag{♭}$$

from which it follows that $v(4e_i) \geq 0$.

Consider the case $\operatorname{char} k \neq 2$. Then $v(e_i) \geq 0$, hence $\widetilde{P_i} = (\widetilde{e_i}, \widetilde{g_i}, 1)$. Since $\widetilde{P_1} = \widetilde{P_2}$, we have $\widetilde{e_1} = \widetilde{e_2}$. The relation

$$\Delta = 16(e_1 - e_2)^2(e_1 - e_3)^2(e_2 - e_3)^2$$

then leads to the contradiction $\widetilde{\Delta} = 0$. In other words, when $\operatorname{char} k \neq 2$ and P_1, P_2 are distinct points of order 2, then $\widetilde{P_1}$, $\widetilde{P_2}$ are distinct points of order 2.

We are left with the case $\operatorname{char} K = 0$, $\operatorname{char} k = 2$. By Proposition 1.7.8, when $\widetilde{a_1} = 0$ then $\widetilde{E}$ is supersingular and has no points of order 2. Then all $\widetilde{P_i} = O$ and (¶) is true. The final detail that completes the proof when $\widetilde{E}$ is ordinary, *i.e.,* when $v(a_1) = 0$, is contained in part (b) of the following proposition; we include a number of other details that will be useful for working out examples.

Proposition 2.5.3 *Let K be a field of characteristic 0 and v a valuation on K with ring V and residue field k of characteristic 2. Let E be an elliptic curve defined over V.*

(a) *If $v(\Delta) = 0$ then at least one of $v(a_1)$ and $v(a_3)$ is 0. If $v(a_1) > 0$ and $v(a_3) = 0$, then $v(\Delta) = 0$, E has supersingular reduction at v, and if (x, y) is a point of order 2 defined over K then $v(x) < 0$; in order that such a point exist it is necessary that $v(2) \geq 3$.*

(b) *Suppose that $v(a_1) = 0$.*

(b1) *If the three points $P_i = (e_i, g_i)$ of order 2 are defined over K then exactly one $P_i \in E_1(K)$,* i.e., *for exactly one i, $v(e_i) < 0$; for that i, $v(4e_i) = 0$.*

(b2) *If K is v-complete then (at least) the point of order 2 with $v(e_i) < 0$ is defined over K.*

(c) *Suppose $E(K)[2]$ contains a fractional point $P_1 = (e_1, g_1)$, $v(e_1) < 0$, but not necessarily all three points of order 2.*

(c1) *If $v(2) = 1$ then $v(a_1) = 0$, $P_1 = (s/4, t/8)$ where $v(s) = v(t) = 0$, and any other points of order 2 present are v-integral.*

(c2) *If $v(2) = 2$ then either*

- *$v(a_1) = 0$ and the conclusions of* (c1) *are valid, or*
- *$v(a_1) = 1$, $v(a_3) > 0$, $v(\Delta) \geq 4$, $P_1 = (s/2, t/2\pi)$ where $v(\pi) = 1$ and $v(s) = v(t) = 0$, and any other points of order 2 present are v-integral.*

Proof. (a) Look at the reduction of the Weierstrass equation over k. We calculate $\widetilde{b_2} = \widetilde{a_1}^2, \ldots, \widetilde{\Delta} = \widetilde{a_1}^4\widetilde{b_8} + \widetilde{a_3}^4 + \widetilde{a_1}^3\widetilde{a_3}^3$. Now suppose $v(a_1) > 0$ and $v(a_3) = 0$. Then $v(\widetilde{\Delta}) = 0$ and $\widetilde{E}$ is supersingular by Proposition 1.7.8. Since $\widetilde{E}$ has no points of order 2, any point P of order 2 in $E(K)$ must reduce to $\widetilde{O}$, *i.e.,* $P \in E_1(K)$. The remark about $v(2) \geq 3$ will follow from part (c).

(b1) Suppose $v(P_i) < 0$. Then $v(e_i) = -2n$, $v(g_i) = v(-(a_1e_i + a_3)/2) = -3n$ for some positive integer n. Using $v(a_1) = 0$ and $v(a_3) \geq 0$ we deduce that $n = v(2)$, hence $v(e_i) = -v(4)$. Comparison of coefficients in ($\flat$) gives

$$e_1 + e_2 + e_3 = \frac{b_2}{4} = a_2 + \frac{a_1^2}{4}, \tag{1}$$

$$e_1e_2 + e_1e_3 + e_2e_3 = \frac{b_4}{2} = a_4 + \frac{a_1a_3}{2}, \tag{2}$$

$$e_1e_2e_3 = \frac{b_6}{4} = a_6 + \frac{a_3^2}{4}. \tag{3}$$

(1) implies that at least one $v(e_i) < 0$, and then $v(e_i) = -v(4)$. Conversely, (3) implies that at most one $v(e_i) < 0$.

(b2) The values of e_i are $u/4$ where u runs through the roots of $f(u) = u^3 + b_2u^2 + 8b_4u + 16b_6$. Since $v(a_1) = 0$, therefore $v(b_2) = 0$, hence $v(f(-b_2)) \geq 3v(2)$ and $v(f'(-b_2)) = 0$. Thus Hensel produces a point $(x, \eta) = (-b_2/4 + \cdots, 0)$ defined over K.

(c) Suppose $v(e_1) = -2n$, $v((a_1e_1 + a_3)/2) = -3n$, and $v(2) < 3$. Then $\min\{v(a_1e_1), v(a_3)\} \leq -3n + v(2) < 0$, hence $v(a_1e_1) = v(a_1) - 2n = -3n + v(2)$, and therefore $v(a_1) = v(2) - n$.

(c1) $v(a_1) = 1 - n$, hence $n = 1$ and $v(a_1) = 0$. By extending K to include all three points of order 2 (it is immaterial whether or not v ramifies), the other remarks follow from (b).

(c2) If $n = 2$ then $v(a_1) = 0$ and proceed as in (c1). So let $n = 1$ and therefore $v(a_1) = 1$. We can write $e_1 = s/2$ where $v(s) = 0$. Suppose $P_2, P_3 \in E(K)[2]$. If either is fractional, the corresponding $e_i = s_i/2$ where $v(s_i) = 0$. Thus (3) implies that not all P_i are fractional. Suppose P_2 but not P_3 is fractional. (We cannot appeal to the fact that $E_1(K)$ is a subgroup since that is not yet proved.) Then (2) is

$$\frac{s}{2}\left(\frac{s_2}{2} + e_3\right) + \frac{s_2}{2}e_3 = a_4 + \frac{a_1a_3}{2},$$

where $v(e_3) \geq 0$. Since $v(a_1) = 1$, the value of the right side is ≥ -1, whereas the value of the left side is -4 — a contradiction. Thus P_2 and P_3 are both v-integral. Looking at (3) again, $v(e_1e_2e_3) \geq -2$ which obliges $v(a_3) > 0$. ■

Here is an example of the second case of (c2):

$$y^2 + \sqrt{2}xy + \sqrt{2}y = x^3, \quad \Delta = -100, \quad \eta^2 = (x + 1/2)(x^2 + 1),$$

and in (x, y) coordinates,

$$P_1 = \left(-\frac{1}{2}, -\frac{1}{2\sqrt{2}}\right), \quad P_2, P_3 = \left(\pm i, -\frac{1 \pm i}{\sqrt{2}}\right).$$

So far we have discussed reduction in terms of a valuation ring. If A is a Krull domain and P is a minimal prime ideal of A, then by **reduction mod P** we understand reduction from the valuation ring A_P to the residue field $k_P = A_P/P_P$. We saw in Proposition 2.2.4 that when A is a UFD, the coordinates of a point can be chosen to be simultaneously proper for all P; but in general the coordinates must be adjusted to accomodate different P.

For any abelian group Γ, let $\mathcal{T}(\Gamma)$ denote the torsion subgroup. A standard application of reduction is to obtain information about $\mathcal{T}(E(K))$. For example, when E is defined over $\mathbf{Z}$, we can reduce E mod p for various primes p to obtain information about $\mathcal{T}(E(\mathbf{Q}))$.

Corollary 2.5.4 *Let the elliptic curve E be defined over the Krull domain A with quotient field K. Let $\mathcal{S}$ denote the set of minimal prime ideals P such that $v_P(\Delta) = 0$ and the torsion subgroups of the following two groups are finite: the reduction $\widetilde{E}(k_P)$ mod P, and the first subgroup $E_1(K)$ in the P-adic filtration. If $\mathcal{S}$ is not empty, then $\mathcal{T}(E(K))$ is finite, and*

$$|\mathcal{T}(E(K))| \quad \textit{divides} \quad \gcd_{P\in\mathcal{S}}\{|\mathcal{T}(E_1(K))||\mathcal{T}(\widetilde{E}(k_P))|\}.$$

Proof. For each P such that $v_P(\Delta) = 0$, the reduction homomorphism restricts to $\mathcal{T}(E(K)) \longrightarrow \mathcal{T}(\widetilde{E}(k_P))$ with kernel $\mathcal{T}(E_1(K))$. ■

See §2.10.1 for examples.

2.5.3 Local expansions

Let E be an elliptic curve given by a nonsingular Weierstrass equation $y^2 + a_1xy + \cdots - a_6 = 0$ over the field K and let $P \in E(K)$. The place corresponding to P has degree 1 (*cf.* §2.2.2), that is, if V is the valuation ring and z is a local parameter, then the residue field V/Vz is K. It follows that the completion of V is the power series ring

$$\widehat{V} = \varprojlim V/Vz^n = K[[z]],$$

and the completion of the function field $L = K(x, y)$ at this place is the quotient field $Q(\widehat{V}) = K((z))$, the field of formal Laurent series. A pertinent application of Hensel's lemma is the calculation of x and y as elements of this field.

The principal case is $P = O$, but let us first take a numerical example of a P in the affine part of the plane.

On

$$E: \quad y^2 + xy = x^3 + x^2 + 4x + 5, \tag{A89}$$

$P = (-5/4, 5/8)$ is a point of order 2 in $E(\mathbf{Q})$. The corresponding valuation v on $\mathbf{Q}(x, y)$ is ramified over $\mathbf{Q}(x)$, and a local parameter is

$$z = \eta = y + \frac{a_1x + a_3}{2} = y + \frac{x}{2}.$$

If we substitute $x = 2(z - y)$, the Weierstrass equation becomes $g(y) = 0$ where

$$g(y) = (-5 - 8z - 4z^2 - 8z^3) + (8 + 10z + 24z^2)y + (-5 - 24z)y^2 + 8y^3.$$

Since

$$v(g(5/8)) = v\left(-\frac{89}{8}z + 11z^2 - 8z^3\right) = 1,$$

$$\text{and} \quad v(g'(5/8)) = v\left(\frac{89}{8} - 20z + 24z^2\right) = 0,$$

we can start Hensel at $y_1 = 5/8$. Using the big-O notation ($A = B + O(z^n)$ means $v(A - B) \geq n$),

$$y_1 - \frac{\left(-89z/8 + O(z^2)\right)}{89/8 + O(z)} = 5/8 + z + O(z^2)$$

— since $y = y_1 + O(z)$, there is no point calculating to higher accuracy — and we take $y_2 = 5/8 + z$. We have $y = y_2 + O(z^2)$.

The next step happens to involve no truncation:

$$y_3 = \frac{5}{8} + z - \frac{z^2}{89/8} = \frac{5}{8} + z - \frac{8}{89}z^2,$$

which is $y + O(z^4)$. One more iteration gives

$$y = \frac{5}{8} + z - \frac{8}{89}z^2 - \frac{2^{10}5}{89^3}z^4 - \frac{2^{15}3 \cdot 37}{89^5}z^6 + O(z^8),$$

"and so on". Thus

$$x = 2z - 2y = -\frac{5}{4} + \frac{2^4}{89}z^2 + \frac{2^{11}5}{89^3}z^4 + \frac{2^{16}3 \cdot 37}{89^5}z^6 + O(z^8).$$

In the limit we get

$$(x, y) = (-5/4 + (16/89)z^2 + \cdots, 5/8 + z + \cdots)$$

which can now be regarded as a point in $E(\mathbf{Q}((z)))$ where z is a trancendental over $\mathbf{Q}$. With this altered perspective of z, the above point, whose coordinates now have particular values in the extended ground field $K' = \mathbf{Q}((z))$, is one of the three points of intersection of E with the line $x + 2y = 2z$ in the affine plane over K'. Actually the other two points of intersection are defined only over the quadratic extension $K'(\sqrt{-1})$; we leave the investigation to the reader.

To describe the situation precisely, let $L = K(x, y)$ denote the function field of E (here $K = \mathbf{Q}$), and let z be transcendental over K. The point $P = (-5/4, 5/8) \in E(K)$ has given rise first to the K-algebra embedding $\rho_P : K(z) \longrightarrow L$, defined by $z \mapsto y + x/2$, which makes L into a $K(z)$-algebra, and

second to the $K(z)$-algebra embedding $\sigma_P : L \longrightarrow K((z))$ defined by $(x, y) \mapsto (-5/4 + (16/89)z^2 + \cdots, 5/8 + z + \cdots)$. We have the following network of field extensions:

$$\begin{tikzcd}
 & & K((z)) \arrow[dd, dash] \\
 & L \arrow[ur, "\sigma_P"] \arrow[dl, dash, "2"'] \arrow[d, dash, "3"'] & \\
K(x) \arrow[dr, dash] & K(y) \arrow[d, dash] & K(z) \arrow[ul, dash, "3"] \arrow[uul, "\rho_P"'] \arrow[dl, dash] \\
 & K &
\end{tikzcd}$$

The numbers are degrees: $[L : K(x)] = 2$, etc.

We now treat the point O on a general E. We choose the local parameter $z = -x/y$ with the minus sign to make certain formulas come out a little neater. We apply a linear transformation to translate O to the origin. With the usual projective coordinates (X, Y, Z), we have $x = X/Z$, $y = Y/Z$ so that $z = -X/Y$. The transformation in matrix notation is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X \\ Z \\ -Y \end{pmatrix}.$$

If we define $w = -1/y$ then the new affine coordinates are

$$z = -X/Y = -x/y \quad \text{and} \quad w = -Z/Y = -1/y.$$

Thus the (z, w)-coordinates of O are $(0, 0)$, and the equation of E in the affine z, w plane is

$$z^3 + (-1 + a_1 z + a_2 z^2)w + (a_3 + a_4 z)w^2 + a_6 w^3 = 0. \qquad (\#)$$

If we regard the left side as a function of w, say $f(w)$, then

$$f'(w) = -1 + a_1 z + a_2 z^2 + 2(a_3 + a_4)w + 3a_6 w^2.$$

Since z is a local parameter,

$$v(f(z^3)) = v(a_1 z^4 + \cdots) \geq 4 \quad \text{and} \quad v(f'(z^3)) = v(-1 + \cdots) = 0.$$

Thus Hensel can be implemented starting with $w_1 = z^3$. The convergence is very rapid since $w = w_2 + O(z^6)$, etc., and the result is

$$w = z^3[1 + A_1 z + A_2 z^2 + \cdots]$$

where the first few coefficients are as follows.

i	A_i
1	a_1
2	$a_1^2 + a_2$
3	$a_1^3 + 2a_1a_2 + a_3$
4	$a_1^4 + 3a_1^2a_2 + 3a_1a_3 + a_2^2 + a_4$
5	$a_1^5 + 4a_1^3a_2 + 6a_1^2a_3 + 3a_1a_2^2 + 3a_1a_4 + 3a_2a_3$
6	$a_1^6 + 5a_1^4a_2 + 10a_1^3a_3 + 6a_1^2a_2^2 + 6a_1^2a_4 + 12a_1a_2a_3$
	$+a_2^3 + 3a_2a_4 + 2a_3^2 + a_6$

Since the series for f' begins with -1, therefore

$$A_i \in \overline{\mathbf{Z}}[a_1, a_2, a_3, a_4, a_6], \quad \forall i$$

where $\overline{\mathbf{Z}}$ denotes $\mathbf{Z}$ mod char K.

To obtain the z-expansions of x and y we need to invert the series for w.

(Alternatively, we could have calculated the expansions of x and y directly, without calculating w, in the following way. Extend the ground field to $K((z))$, where z is transcendental over K, and intersect E with the line $x = -yz$. The calculation is facilitated by introducing a parameter t and writing the line as $x = t/z^2$, $y = -t/z^3$. Substituting these values into the Weierstrass equation and multiplying by z^6, the three points of intersection correspond to the roots of the cubic $g(t) = 0$ where

$$g(t) = t^3 + (-1 + a_1 z + a_2 z^2)t^2 + (a_3 z^3 + a_4 z^4)t + a_6 z^6.$$

Denoting the z-adic valuation by v, we have

$$v(f(1)) = v(a_1 z + a_2 z^2 + \cdots) \geq 1,$$
$$v(f'(1)) = v(1 + 2a_1 z + \cdots) = 0,$$

hence we can start Hensel at $t_1 = 1$ with the result

$$t = 1 - a_1 z - a_2 z^2 - a_3 z^3 - (a_4 + a_1 a_3)z^4 + \cdots + B_i z^i + \cdots.$$

However we will need the series for w when we construct the formal group $\widehat{E}$ in §2.5.2.)

The result is

$$w^{-1} = z^{-3}[1 + B_1 z + B_2 z^2 + \cdots] \quad \text{where}$$

i	B_i
1	$-a_1$
2	$-a_2$
3	$-a_3$
4	$-a_4 - a_1a_3$
5	$-a_1a_4 - a_2a_3 - a_1^2a_3$
6	$-a_6 - a_1^2a_4 - a_2a_4 - a_3^2 - a_1^3a_3 - 2a_1a_2a_3$
7	$-2a_1a_6 - 2a_3a_4 - 2a_1a_2a_4 - a_2^2a_3 - a_1^3a_4 - 3a_1^2a_2a_3 - 3a_1a_3^2 - a_1^4a_3$

Since the series for $z^{-3}w$ begins with 1, therefore all $B_i \in \overline{\mathbf{Z}}[a_1, \ldots, a_6]$. For reference we put the final results in a proposition.

Proposition 2.5.5 *Let $y^2 + a_1xy + \cdots - a_6 = 0$ define an elliptic curve. The expansion at O of (x, y) in terms of the local parameter $z = -x/y$ is*

$$(x, y) = (t/z^2, -t/z^3), \quad \textit{where}$$

$$t = 1 - a_1z - a_2z^2 - a_3z^3 - (a_4 + a_1a_3)z^4 + \cdots + B_iz^i + \cdots.$$

The coefficients B_i are in $\overline{\mathbf{Z}}[a_1, \ldots, a_6]$ and can be calculated by the method described above using Hensel's Lemma.

Note that the network of fields pictured above is valid if we define

$$\rho_P(z) = -x/y, \quad \text{and} \quad \sigma_P(x, y) = (t/z^2, -t/z^3).$$

2.6 Formal groups

Let A be a commutative ring and S and T independent transcendentals. A power series $F \in A[[S, T]]$ is a commutative **formal group** if

FG1 $F(S, T) \equiv S + T \bmod \deg 2$,

FG2 $F(F(R, S), T) = F(R, F(S, T))$ and

FG3 $F(T, S) = F(S, T)$.

Notice that the associative law implies $F(S, 0) = S$, *i.e.,* there are no pure powers in S beyond the first, and similarly $F(0, T) = T$. For if k were minimal such that S^k occurs with nonzero coefficient a, then $F(S, 0) = F(S, F(0, 0)) = F(F(S, 0), 0)) = F(S, 0) + aF(S, 0)^k + \cdots$ which is a contradiction. Thus a

formal group, *i.e.,* a power series satisfying FG1 and FG2, but not necessarily FG3, has a power series expansion beginning

$$F = S + T + a_{11}ST + a_{12}ST^2 + a_{21}S^2T + \cdots.$$

It is a fact that the commutative law FG3, equivalently $a_{ij} = a_{ji}\ \forall i, j$ is "almost always" a consequence of the first two axioms: all formal groups over A are commutative iff the ideal of nilpotent elements of A is torsion-free as an additive group. In one direction the proof is easy: suppose A contains a torsion nilpotent element a. By replacing a by an appropriate integer multiple of a power of itself we can assume $a \neq 0,\ a^2 = 0,\ pa = 0$ for a prime p. Then $F(S,T) = S + T + aS^pT$ is a noncommutative formal group. For the converse see [Laz54], [Con66], [Hon68] and [Hon70]. Since the noncommutive case will not play a role in our work we make the terminological convention

$$\boxed{\textit{formal group} \ \text{ means } \ \textit{commutative formal group.}}$$

There is a unique series $N_F = N_F(S) \in SA[[S]]$ satisfying

$$F(S, N_F(S)) = 0.$$

We already observed this in §2.4.1 in the case when A is an integral domain and actually the general case follows by writing A as the quotient of an integral domain B. However it is simpler to substitute $N_F = b_1S + b_2S^2 + \cdots$ and solve recursively for the b_i. This negative series begins

$$N_F = -S + a_{11}S^2 - a_{11}^2S^3 + (a_{11}^3 + a_{11}a_{12} + 2a_{13} - a_{22})S^4 + \cdots.$$

Let F and G be formal groups over A. A **morphism** from F to G is a power series f in one variable over A with 0 constant term such that

$$f(F(S,T)) = G(f(S), f(T)).$$

The set of morphisms from F to G is denoted $\hom(F,G)$, or when necessary, $\hom_A(F,G)$. Occasionally we use the categorical notation $f : F \longrightarrow G$.

We now list a few "formalities", leaving the easy details for the most part to the reader.

- $\hom(F,G)$ is an abelian group: if $e, f \in \hom(F,G)$ their sum is defined by $(e+f)(T) = G(e(T), f(T))$. The 0 series is the group 0 and $f(N_F(T)) = N_G(f(T))$ serves as $-f$. When checking the group axioms it is essential that G be commutative, as it is in the case of ordinary groups. Note that although we can identify $\hom(F,G)$ as a subset of the ring $\mathcal{R} = \mathcal{R}(A)$, the operation $+$ is (usually) different.

- If $f \in \hom(F,G)$ and $g \in \hom(G,H)$ then $g \circ f \in \hom(F,H)$ and $\circ$ defines a bilinear map

$$\hom(F,G) \times \hom(G,H) \longrightarrow \hom(F,H).$$

When we identify both $\hom(F,G)$ and $\hom(G,H)$ as subsets $\mathcal{R}$, then the above operation $\circ$ coincides with multiplication in the ring $\mathcal{R}$. Hence $f \in \hom(F,G)$ is an isomorphism iff $f \in \mathcal{R}^*$, *i.e.,* the coefficient of T is in A^*, the inverse isomorphism then being the group inverse $f^{(-1)} \in \mathcal{R}^*$:

$$f^{(-1)} \circ f = [1]_F, \quad f \circ f^{(-1)} = [1]_G.$$

- $\mathrm{end}(F) = \mathrm{end}_A(F) = \hom(F,F)$ is a ring (usually noncommutative) with unit element represented by the power series T which we denote $[1]$; in general for $n \in \mathbf{Z}$ we write $[n]$ for $n \in \mathrm{end}(F)$. Note that $[-1]$ is the power series N_F. Formulas for the first half dozen terms in the series $[n]$ are given in §2.7. If $f \in \hom(F,G)$ then $\forall n \in \mathbf{Z}$, $f \circ [n]_F = [n]_G \circ f$.

- If $f = a_1T + a_2T^2 + \cdots \in \hom(F,G)$ define $c(f) = a_1$. Then $c : \hom(F,G) \longrightarrow A$ is an abelian group homomorphism and when $F = G$ it is a ring homomorhism. In particular, $c([n]) = n$ for $n \in \overline{\mathbf{Z}}$, where $\overline{\mathbf{Z}}$ is the image of $\mathbf{Z}$ in A. Note that $c(f) \in A^* \Leftrightarrow f \in \mathcal{R}^* \Leftrightarrow f$ is an isomorphism. A **strict isomorphism** is one with $c(f) = 1$.

For future reference we repeat:

Proposition 2.6.1 *Let F be a formal group over the commutative ring A and let $n \in \overline{\mathbf{Z}}$. Then $c([n]) = n$. Hence $[n]$ is an automorphism of F iff $n \in A^*$.*

- If $f \in \hom(F,G)$ is an isomorphism then $F = f^{(-1)}(G(f(S), f(T))$, so F is uniquely determined by G, and of course conversely G is uniquely determined by F since $f^{(-1)} \in \hom(G,F)$ is an isomorphism. We use the notation $F = G^f$. Then if $e \in \hom(E,F)$ is also an isomorphism we have

$$(G^f)^e = G^{f \circ e}.$$

Again for reference we put the essential facts in a

Proposition 2.6.2 *Let $\mathcal{F} = \mathcal{F}(A)$ denote the set of formal groups over A, let $F \in \mathcal{F}$ and $f \in \mathcal{R}^*$, and define*

$$F^f = f^{(-1)}\left(F(f(S), f(T))\right).$$

Then $F^f \in \mathcal{F}$ and $f \in \hom_A(F^f, F)$ is an isomorphism. Moreover, all isomorphisms to F are obtained this way.

If also $g \in \mathcal{R}^*$ then

$$\left(F^f\right)^g = F^{f \circ g}.$$

Hence $\mathcal{R}^*$ acts on $\mathcal{F}$ on the right, and the orbits are the A-isomorphism classes in $\mathcal{F}$.

- If $\phi : A \longrightarrow B$ is a homomorphism of commutative rings, in other words B is an A-algebra, and $F = S + T + a_{11}ST + \cdots$ is a formal group over A, then $\phi_* F = S + T + \phi(a_{11})ST + \cdots$ is a formal group over B. If $f = a_1 T + a_2 T^2 + \cdots \in \hom_A(F, G)$, and we define $\phi_* f = \phi(a_1)T + \phi_*(a_2)T^2 + \cdots$, then $\phi_* f \in \hom_B(\phi_* F, \phi_* G)$.

  A case of importance for us will be the reduction map $V \longrightarrow k$ of a valuation ring to its residue field.

- If B is a commutative A-algebra and I is a topologically nilpotent ideal in B (see §2.4.1), for example when B is a discrete valuation ring with maximal ideal I, and assuming that B is I-adically complete (again see §2.4.1 for the definition), then F induces an abelian group structure on I by the convergent series

$$x +_F y = F(x, y), \qquad -_F x = N_F(x).$$

  We denote this group by $F(I)$ to distinguish it from the usual abelian group I given by the ring operations. Since I^n are (topologically) closed subgroups, we have the subgroup filtration

$$F(I) \supset F(I^2) \supset F(I^3) \supset \cdots \quad .$$

  Later (in Proposition 2.9.2) we will see circumstances in which the groups $F(I^n)$ and I^n are isomorphic. In any case, for $n \geq 1$ the identity map

$$F(I^n)/F(I^{n+1}) \longrightarrow I^n/I^{n+1}$$

  is a group isomorphism since for $x, y \in I^n$,

$$x +_F y = F(x, y) = x + y + a_{11}xy + \cdots \equiv x + y \bmod I^{2n}.$$

  In the case where B is a complete discrete valuation ring, by one of the remarks in §2.1, $I^n/I^{n+1} \approx k^+$,and so in this case we have

$$F(I^n)/F(I^{n+1}) \xrightarrow{\sim} k^+.$$

  We now restrict our attention to the case where B is a complete local ring with maximal ideal I. (In the present context a local ring is understood to be commutative.) Localness ensures that an element of B that is not in I is invertible in B.

Proposition 2.6.3 *Let B be a complete local ring with maximal ideal I, let F be a formal group over B and let p denote the characteristic of the field B/I. Then*

- *if $p = 0$ then the abelian group $F(I)$ is torsion free;*
- *if $p > 0$, in particular when* $\operatorname{char} B = p > 0$, *then $F(I)$ has no torsion prime to p: if $x \in I$ and $[n]_F(x) = 0$ for some positive integer n, then $[p^m]_F(x) = 0$ for sufficiently large m.*

Proof. Suppose $[n](x) = 0$ where n is positive. In the case $p > 0$ we can replace x by $[p^m](x)$ for sufficiently large m so that we can assume $\gcd(n, p) = 1$. We must prove that $x = 0$.

In both cases $n \notin I$, hence $n \in B^*$. By the previous proposition, $[n]$ is an automorphism of F, hence induces an automorphism on $F(I)$, and in particular, the kernel is 0. ■

2.6.1 The additive and multiplicative formal groups

Our first two examples of formal groups are the polynomials

$$\mathbf{A} = S + T, \qquad \mathbf{M} = S + T + ST = (1 + S)(1 + T) - 1$$

which are called, respectively, the **additive** and the **multiplicative** formal group. Alternative notation is $\mathbf{G}_a$ and $\mathbf{G}_m$. These are defined over any commutative ring A, and we write $\mathbf{A}_A$, $\mathbf{M}_A$ when necessary. An easy induction gives $\forall n \in \mathbf{Z}$

$$[n]_{\mathbf{A}}(T) = nT, \qquad [n]_{\mathbf{M}}(T) = (1 + T)^n - 1$$

(the latter being expanded as a series as usual when $n < 0$).

Taking $A = \mathbf{Z}$ and $B = \lim_{\leftarrow} B/I^n$ a complete ring as described in the previous section, for $n \geq 1$ we can identify the group $\mathbf{A}(I^n)$ with the additive group I^n, and the group $\mathbf{M}(I^n)$ with the multiplicative group $1 + I = \{1 + i : i \in I\}$. With these identifications we have exact sequences of abelian groups:

$$0 \longrightarrow \mathbf{A}(I) \longrightarrow B \longrightarrow B/I \longrightarrow 0,$$

$$1 \longrightarrow \mathbf{M}(I) \longrightarrow B^* \longrightarrow (B/I)^* \longrightarrow 1.$$

Also from the discussion in the previous section,

$$1 + i + I^{n+1} \longmapsto i + I^{n+1}, \quad i \in I^n$$

induces an isomorphism

$$\mathbf{M}(I^n)/\mathbf{M}(I^{n+1}) = 1 + I^n/1 + I^{n+1} \longrightarrow I^n/I^{n+1}.$$

By Proposition 2.6.2, the construction

$$\mathbf{A}^f = f^{(-1)}(f(S) + f(T)), \quad f \in \mathcal{R}^*(A)$$

yields all formal groups isomorphic to $\mathbf{A}_A$; *e.g.* when f is the unit element T of $\mathcal{R}^*$ we get $\mathbf{A}$ itself. We will see in §2.7 that when A is a $\mathbf{Q}$-algebra then *every* formal groups is of the form $\mathbf{A}^f$, *i.e.,* every formal group is isomorphic over A with $\mathbf{A}$. For example $\mathbf{M} = \mathbf{A}^{\log_{\mathrm{M}}}$ where

$$\log_{\mathbf{M}} = (\text{formal}) \int \frac{1}{1+T} dT = \sum_{n=1}^{\infty} (-1)^{n-1} T^n / n.$$

(The subscript $\mathbf{M}$ notation will be explained in §2.9.1.) To verify this statement we need to calculate $\log_{\mathbf{M}}^{(-1)}$, which we denote $\exp_{\mathbf{M}}$. We claim that (*cf.* Proposition 2.3.3)

$$\exp_{\mathbf{M}} = \sum_{n=1}^{\infty} T^n / n!.$$

Now when $A = \mathbf{C}$, the complex field, we have

$$\log_{\mathbf{M}} = \log(1+T), \quad \exp_{\mathbf{M}} = e^T - 1, \quad \text{hence}$$

$$\log_{\mathbf{M}} \circ \exp_{\mathbf{M}} = T = \exp_{\mathbf{M}} \circ \log_{\mathbf{M}}.$$

Either of the latter equations implies the other since $\mathcal{R}(\mathbf{C})$ is a group. They are true in $\mathbf{Q}[[T]]$ and therefore also in $A \otimes_{\mathbf{Z}} \mathbf{Q}[[T]] = A[[T]]$. In the same way,

$$\exp_{\mathbf{M}}(\log_{\mathbf{M}}(S) + \log_{\mathbf{M}}(T)) = e^{\log(1+S)+\log(1+T)} - 1 = S + T + ST$$

is true in $\mathbf{C}[[S,T]]$, hence in $\mathbf{Q}[[S,T]]$, hence in $A[[S,T]]$. Thus when A is a $\mathbf{Q}$-algebra and $f = \log_{\mathbf{M}}$, then $\mathbf{A}^f = \mathbf{M}$.

In contrast, we note

Proposition 2.6.4 *When A is a commutative ring of prime characteristic p, then*

$$\hom_A(\mathbf{M}, \mathbf{A}) = 0.$$

In particular, $\mathbf{A}$ *and* $\mathbf{M}$ *are not isomorphic.*

Proof. Since $[p]_{\mathbf{A}} = pT = 0$ and $[p]_{\mathbf{M}} = (1+T)^p - 1 = T^p$, if $f \in \hom(\mathbf{M}, \mathbf{A})$ then $[p]_{\mathbf{A}} \circ f = f \circ [p]_{\mathbf{M}}$ is $0 = f(T^p)$, hence $f = 0$. ■

Corollary 2.6.5 *If V is a discrete valuation ring of residue characteristic $p > 0$, then*

$$\hom_V(\mathbf{M}, \mathbf{A}) = 0.$$

Proof. Suppose $f \in \hom_V(\mathbf{M}, \mathbf{A})$ and $f \neq 0$, hence $\operatorname{char} V = 0$ by the proposition. Let π denote a uniformizer and $\phi : V \longrightarrow k$ the reduction map to the residue field. Then for appropriate $n \geq 0$, we have $\pi^{-n} f \in V[[T]]$ and $\phi_*(\pi^{-n} f) \neq 0$. From $f(S + T + ST) = f(S) + f(T)$ we deduce

$$\pi^{-n} f(S + T + ST) = \pi^{-n} f(S) + \pi^{-n} f(T),$$

hence $\pi^{-n} f \in \hom_V(\mathbf{M}, \mathbf{A})$, and therefore $\phi_*(\pi^{-n} f)$ is a nonzero element in $\hom_k(\mathbf{M}, \mathbf{A})$. This contradicts the proposition. ■

The story for $\hom(\mathbf{A}, \mathbf{M})$ is more complicated, as the following three examples indicate.

(i) If $\operatorname{char} A = p$ is prime and the ring endomorphism $a \mapsto a^p$ of A has 0 kernel, then

$$\hom_A(\mathbf{A}, \mathbf{M}) = 0.$$

For in these circumstances, $[p]_{\mathbf{M}} \circ f = f \circ [p]_{\mathbf{A}} \Longrightarrow (f(T))^p = 0 \Longrightarrow f(T) = 0$.

(ii) If $\operatorname{char} A = 2$ and A contains an element ϵ such that $\epsilon^2 = 0$, then $\epsilon T \in \hom(\mathbf{A}, \mathbf{M})$.

(iii) When $A = \mathbf{Z}_p$, the above corollary implies that $\hom_{\mathbf{Z}_p}(\mathbf{A}, \mathbf{M})$ does not contain an isomorphism. (Note $\log_{\mathbf{M}}$ is only in $\hom_{\mathbf{Q}_p}$ because of denominators.) But $\hom_{\mathbf{Z}_p}(\mathbf{A}, \mathbf{M}) \neq 0$, containing for example

$$\exp_{\mathbf{M}} \circ [p]_{\mathbf{A}} = pT + p^2 T^2/2! + \cdots + p^n T^n/n! + \cdots .$$

In fact $v_p(p^n/n!) > 0$ for $n > 0$, as follows from the

Lemma 2.6.6 *Let n be a positive integer, let its standard p-adic expansion be $a_0 + a_1 p + \cdots + a_N p^N$ where $a_i \in \{0, \ldots, p-1\}$, and put $a = \sum_{i=0}^N a_i$. Then*

$$v_p(n!) = \frac{n-a}{p-1}.$$

Proof. The **integral part** or **floor** of a real number x is the largest integer $\leq x$, and is denoted $\lfloor x \rfloor$. (I believe this notation of D. Knuth is becoming standard, replacing the older $[x]$. While we're at it: the **ceiling** of x is $\lceil x \rceil = -\lfloor -x \rfloor =$ the smallest integer $\geq x$.)

$$\begin{aligned}
v_p(n!) &= \sum_{i \geq 1} \lfloor n/p^i \rfloor = \sum_{i \geq 1} (a_i + a_{i+1} p + \cdots) \\
&= \sum_{i \geq 1} a_i (1 + p + \cdots + p^{i-1}) = \sum_{i \geq 1} a_i \frac{p^i - 1}{p - 1} \\
&= \sum_{i \geq 0} a_i \frac{p^i - 1}{p - 1} \quad \text{since } a_0 0 = 0 \\
&= (n-a)/(p-1) \quad ■
\end{aligned}$$

2.6.2 The formal group of an elliptic curve

Let E be an elliptic curve given by a nonsingular Weierstrass equation with coefficients $a_1, \ldots, a_6$ in the integral domain A. We associate to E a formal group $\widehat{E}$ over A, defined as follows. Let K be the quotient field of A and for $i = 1, 2$ let

$$(x_i, y_i) = (z_i/w_i, -1/w_i) = (t_i/z_i^2, -t_i/z_i^3)$$

be independent generic points as constructed in §2.5.3, independent meaning that the transcendentals z_i are independent. The definition is simply

$$\widehat{E} = z_3 = -x_3/y_3 \quad \text{where} \quad (x_1, y_1) + (x_2, y_2) = (x_3, y_3),$$

and $+$ stands for addition of points on E. However there are a few details to be clarified.

For $i = 1, 2$ let $K_i = K((z_i))$. First, the addition is taking place in $E(M)$ where M is the smallest field containing both K_1 and K_2, namely the quotient field of the integral domain $K_1 \otimes_K K_2$. This domain can be identified with $S^{-1}B$ where $B = K[[z_1, z_2]]$ and $S = \{z_1^i z_2^j : i \geq 0, j \geq 0\}$; thus M is more easily described as the quotient field of B. (There is no need to go to either of the larger fields $K_{12} = K((z_1))((z_2))$, $K_{21} = K((z_2))((z_1))$. Incidentally, do the images of these two fields in $K_{12} \otimes_M K_{21}$ coincide?)

Secondly, we wish to express $\widehat{E}$ as a power series in B. If one takes the formulas for x_3 and y_3 from Proposition 1.4.1 and substitutes $x_i = t_i/z_i^2$, $y_i = -x_i/z_i$ and $t_i = 1 - a_1 z_i - a_2 z_i^2 - \cdots$ for $i = 1, 2$, (with the t_i truncated for some desired degree of accuracy), one obtains a complicated and quite unmanageable expression.

A better way is to work in the z, w-plane. The line joining (z_1, w_1) and (z_2, w_2) is $w = \lambda z + \nu$ where, in the notation of §2.5.3,

$$\lambda = \frac{w_2 - w_1}{z_2 - z_1} = \sum_{n=3}^{\infty} A_{n-3} \frac{z_2^n - z_1^n}{z_2 - z_1}$$

and $\nu = w_1 - \lambda z_1$. We noted in §2.5.3 that all the A_n are in the ring $\overline{\mathbf{Z}}[a_1, \ldots, a_6] = R$, say, and it follows that λ and ν are in the ring $R[[z_1, z_2]]$. Since the transformation from x, y to z, w coordinates is linear, therefore the third point of intersection of this line with E is $-(z_3, w_3)$; let us denote the coordinates of this point (z_4, w_4). Using the equation labelled (#) in §2.5.3, we see that z_1, z_2 and z_4 are the three roots of the cubic

$$z^3 + (-1 + a_1 z + a_2 z^2)(\lambda z + \nu) + (a_3 + a_4 z)(\lambda z + \nu)^2 + a_6(\lambda z + \nu)^3.$$

Collecting terms and using the fact that the sum of the roots of a monic polynomial is minus the second coefficient, we obtain

$$z_4 = -z_1 - z_2 - \frac{a_1\lambda + a_2\nu + a_3\lambda^2 + 2a_4\lambda\nu + 3a_6\lambda^2\nu}{1 + a_2\lambda + a_4\lambda^2 + a_6\lambda^3}.$$

Since the constant terms of λ and ν are 0, therefore $z_4 \in R[[z_1, z_2]]$. Finally, since $(z_3, w_3) = -(z_4, w_4)$ we have

$$z_3 = -\frac{x_3}{y_3} = \frac{x_4}{y_4 + a_1 x_4 + a_3} = \frac{-z_4}{1 - a_1 z_4 - a_3 w_4}$$

$$= -z_4 \left(1 - a_1 z_4 - a_3 z_4^3 [1 + A_1 z_4 + \cdots]\right)^{-1},$$

using the z-expansion for w obtained in §2.5.3. Clearly this is also in $R[[z_1, z_2]]$. The final result is

$$\widehat{E} = z_1 + z_2 - a_1 z_1 z_2 - a_2(z_1 z_2^2 + z_1^2 z_2) + \cdots$$

where the general terms in the series are $a_{ij}(z_1^i z_2^j + z_1^j z_2^i)$ for $i < j$, and $a_{ii} z_1^i z_2^i$.

a_{11}	$-a_1$
a_{12}	$-a_2$
a_{13}	$-2a_3$
a_{22}	$-3a_3 + a_1 a_2$
a_{14}	$-2a_4 - 2a_1 a_3$
a_{23}	$-4a_4 - a_1 a_3 + a_2^2$
a_{15}	$-2a_1^2 a_3 - 2a_1 a_4 - 2a_2 a_3$
a_{24}	$-a_1^2 a_3 - a_1 a_4$
a_{33}	$-a_1^2 a_3 - a_1 a_2^2 + 2a_2 a_3$
a_{16}	$-3a_6 - 2a_3^2 - 2a_1^3 a_3 - 2a_1^2 a_4 - 2a_2 a_4 - 4a_1 a_2 a_3$
a_{25}	$-9a_6 - a_1^3 a_3 - a_1^2 a_4 - a_1 a_2 a_3$
a_{34}	$-15a_6 - a_2^3 + 4a_3^2 - a_1^2 a_4 - a_1^3 a_3 + 4a_2 a_4 - 2a_1 a_2 a_3$

A better way, and probably the best way, to calculate these coefficients is to use the formal logarithm; this will be explained in §2.7.

We have taken the table far enough to see that distinct E give distinct $\widehat{E}$, that is, we can recover the Weierstrass coefficients $a_1, \ldots, a_6$ from the coefficients a_{ij} — with the following exceptions:

- a_6 remains indeterminate in characteristic 3;
- a_4 remains indeterminate in characteristic 2 when $j = 0$.

The proof of Proposition 2.1.6 can now be completed.

Proposition 2.6.7 *Let E be an elliptic curve defined over the discrete valuation ring V with quotient field K, let M denote the maximal ideal of the completion $\widehat{V}$, let k denote the residue field and k^+ the additive group of k. Then for $m \geq 1$*

(i) *$E_m(K)$ is a subgroup of $E(K)$ and $O \mapsto 0$ and $P = (x, y) \mapsto -x/y$ for $P \neq O$ define an injective group homomorphism*

$$E_m(K) \longrightarrow \widehat{E}(M^m)\,;$$

this is an isomorphism when K is complete;

(ii) *there is an exact sequence of group homomorphisms*

$$0 \longrightarrow E_{m+1}(K) \longrightarrow E_m(K) \longrightarrow k^+\,;$$

the map on the right is surjective when K is complete.

Proof. (i) If $\widehat{K}$ denotes the v-adic completion of K, then

$$E_m(K) = E_m(\widehat{K}) \cap E(K), \quad m \geq 1, \qquad \P$$

and therefore to prove that $E_m(K)$ is a subgroup of $E(K)$, it is sufficient to prove that $E_m(\widehat{K})$ is a subgroup of $E(\widehat{K})$.

Now the map $E_m(\widehat{K}) \longrightarrow M^m$ defined by $O \mapsto 0$ and $P = (x, y) \mapsto -x/y$ for $P \neq O$ has the inverse $0 \mapsto O$ and for $z \neq 0$, $z \mapsto (t/z^2, -t/z^3)$ in the notation of Proposition 2.5.5; for $v(z) = m > 0$ and therefore the series converges to an element $t \in \widehat{K}$ with $v(t) = 0$. Thus these maps are bijections. Moreover, when we give the set M^n the group structure $\widehat{E}(M^n)$, the inverse map is a group homomorphism by construction of $\widehat{E}$, and (i) is proved.

(ii) Combining remarks from §2.1 and §2.5, we have group isomorphisms

$$E_m(\widehat{K})/E_{m+1}(\widehat{K}) \approx \widehat{E}(M^m)/\widehat{E}(M^{m+1}) \approx M^m/M^{m+1} \approx k^+, \quad m \geq 1.$$

By ¶ we can regard $E_m(K)/E_{m+1}(K)$ as a subgroup of the group on the left, and (ii) follows. ■

2.7 The invariant differential of a formal group

Let A be a commutative ring. Recall from [BA3] that the A-linear derivations are of the form

$$\alpha \frac{d}{dT}, \quad \alpha \in A[[T]],$$

and so comprise a free $A[[T]]$-module of rank 1; and the dual module of A-linear differentials are of the form

$$\alpha\, dT, \quad \alpha \in A[[T]],$$

and so also comprise a free $A[[T]]$-module of rank 1, which we denote $\mathcal{D} = \mathcal{D}(A,T)$. A member $(a_0 + a_1T + \cdots)\,dT \in \mathcal{D}$ is a **normalized differential** if $a_0 = 1$.

Let us now specialize to the case $A = R[[S]]$ where R is an integral domain and P is the prime ideal SA in A. As we saw in Proposition 2.4.5, $\mathcal{D}$ is an $\mathcal{R}_P^*$-module with the definition

$$(\alpha\,dT)^\rho = (\alpha \circ \rho)\rho'\,dT.$$

Since $(a\alpha\,dT)^\rho = a(\alpha\,dT)^\rho$ for all $a \in A$, we can regard $\mathcal{D}$ as an A–$\mathbf{Z}[\mathcal{R}_P^*]$ bimodule.

Let $F = F(S,T) = S + T + a_{11}ST + \cdots$ be a formal group over R and let $F_1(S,T)$ (resp. $F_2(S,T)$) denote the partial derivative of $F(S,T)$ with respect to S (resp. T):

$$F_1(S,T) = 1 + a_{11}T + a_{12}(2ST + T^2) + \cdots,$$

$$F_2(S,T) = 1 + a_{11}S + a_{12}(S^2 + 2ST) + \cdots.$$

Note that

$$F \in \mathcal{R}_P^*.$$

Proposition 2.7.1 *There is a unique differential $\alpha\,dT \in \mathcal{D}$ satisfying the three conditions*

(i) $\partial\alpha/\partial S = 0$, *that is,* $\alpha \in R[[T]]$; *we write* $\alpha = \alpha(T)$;

(ii) $\alpha\,dT$ *is normalized*, i.e., $\alpha(0) = 1$; *and*

(iii) $\alpha^F = \alpha$.

It is given by the formula

$$\begin{aligned}\alpha &= \frac{1}{F_1(0,T)}\\ &= \frac{1}{1 + a_{11}T + a_{12}T^2 + \cdots}\\ &= 1 - a_{11}T + (a_{11}^2 - a_{12})T^2 - (a_{11}^3 - 2a_{11}a_{12} + a_{13})T^3 + \cdots.\end{aligned}$$

Remarks. See the table below for more terms in the series. The differential described in the proposition is the **normalized invariant differential** of the formal group F and will be denoted ω_F. The simplest examples are

$$\begin{aligned}\text{for } \mathbf{A} &= S + T, \qquad F_1(0,T) = 1 \text{ and } \omega_F = dT;\\ \text{for } \mathbf{M} &= S + T + ST, \quad F_1(0,T) = 1 + T \text{ and } \omega_F = (1+T)^{-1}dT.\end{aligned}$$

Condition (i) cannot be dropped, at least when char $R > 0$. Here is an example in characteristic 2 of a normalized $\mathbf{A}$-invariant differential that is not free of S:

$$(1 + ST + T^2)\,dT.$$

Proof. A differential $\alpha(T)\,dT$ satisfying condition (i) also satisfies (iii) iff

$$\alpha(F(S,T))F_2(S,T) = \alpha(T). \tag{$\dagger$}$$

Let us first prove that this is true for $\alpha = 1/F_1(0,T)$, *i.e.*,

$$F_1(0,T)F_2(S,T) = F_1(0,F(S,T)). \tag{$\dagger\dagger$}$$

By FG2 and FG3 we have the identity

$$F(S,F(U,T)) = F(U,F(S,T)).$$

Differentiating this with respect to U and then setting $U = 0$ results in the required identity ($\dagger\dagger$).

Secondly, suppose $\alpha\,dT$ satisfies (iii) and ($\dagger$). Substituting $T = 0$ in ($\dagger$) gives the identity

$$\alpha(S)F_2(S,0) = \alpha(0).$$

Replacing S by T, we see that $\alpha(T)$ is uniquely determined by $\alpha(0)$, and in particular, there is just one such α that is normalized. ■

Here is a table of more terms in the series $\omega_F = 1 + \alpha_1 T + \alpha_2 T^2 + \cdots$; since $F_1(0,T)$ has the simple form $1 + \sum a_{1i}T^i$, this is just the "generic inverse".

i	α_i where $\omega_F = (1 + \alpha_1 T + \alpha_2 T^2 + \cdots)dT$
1	$-a_{11}$
2	$a_{11}^2 - a_{12}$
3	$2a_{11}a_{12} - a_{13} - a_{11}^3$
4	$a_{12}^2 - 3a_{12}a_{11}^2 + 2a_{11}a_{13} - a_{14} + a_{11}^4$
5	$-3a_{11}a_{12}^2 + 2a_{12}a_{13} + 4a_{12}a_{11}^3 - 3a_{13}a_{11}^2 + 2a_{11}a_{14} - a_{15} - a_{11}^5$
6	$-6a_{13}a_{11}a_{12} + a_{13}^2 + 4a_{13}a_{11}^3 + 2a_{14}a_{12} - 3a_{14}a_{11}^2 - a_{16} + 2a_{11}a_{15}$
	$-a_{12}^3 + 6a_{12}^2a_{11}^2 - 5a_{12}a_{11}^4 + a_{11}^6$
7	$2a_{11}a_{16} - a_{11}^7 - a_{17} - 6a_{14}a_{11}a_{12} + 12a_{13}a_{12}a_{11}^2 + 2a_{14}a_{13} + 4a_{14}a_{11}^3$
	$+2a_{15}a_{12} - 3a_{15}a_{11}^2 - 3a_{13}a_{12}^2 - 3a_{11}a_{13}^2 - 5a_{13}a_{11}^4$
	$+4a_{11}a_{12}^3 - 10a_{12}^2a_{11}^3 + 6a_{12}a_{11}^5$

Corollary 2.7.2 *Let F and G be formal groups over R, with invariant differentials $\omega_F = \alpha_F(T)\,dT$ and $\omega_G = \alpha_G(T)\,dT$, and let $f(T) \in \hom(F,G)$. Then*

$$\omega_G^f = f'(0)\omega_F, \quad \textit{i.e.} \quad (\alpha_G \circ f)\cdot f' = f'(0)\alpha_F.$$

Proof. Differentiate

$$f(F(S,T)) = G(f(S), f(T))$$

with respect to S, and then set $S = 0$:

$$f'(T)F_1(0,T) = G_1(0, f(T))f'(0).$$

Since $\alpha_F = 1/F_1(0,T)$ and $\alpha_G = 1/G_1(0,T)$, the result follows. ■

Corollary 2.7.3 *Let F be a formal group over R and n an integer. Then the derivative $[n]' \in nR[[T]]$. Thus when p is a prime number,*

$$[p](T) = p\lambda_p(T) + \mu_p(T^p) \quad \textit{and} \quad [-p](T) = p\lambda_{-p}(T) + \mu_{-p}(T^p)$$

for certain $\lambda_p, \mu_p, \lambda_{-p}, \mu_{-p} \in TR[[T]]$.

Proof. By straightforward induction starting with $[1] = T$ we have

$$[n] \equiv nT \bmod T^2 \qquad \forall n \in \mathbf{Z}. \tag{¶}$$

Hence $[n]'(0) = n$, and with $\omega_F = \alpha\, dT$, $F = G$ and $f = [n]$, the previous corollary implies

$$\alpha([n](T)) \cdot [n]' = n\alpha(T).$$

Since $\alpha(0) = 1$, therefore $\alpha \in R[[T]]^*$ and $[n]' \in nR[[T]]$. ■

For a general formal group $F = S + T + a_{11}ST + \cdots$ defined over R, using (¶), let

$$[n](T) = nT + M_2(n)T^2 + M_3(n)T^3 + \cdots.$$

The quantities $M_2, M_3, \ldots$ can be determined recursively as functions of n and the coefficients of F by repeated application of

$$[n+1](T) = F([n](T), T).$$

For example, the first step gives

$$[n+1] = F(nT + M_2(n)T^2 + \cdots, T) = nT + M_2(n)T^2 + \cdots + T + a_{11}(nT^2 + \cdots) + \cdots$$

hence

$$M_2(n) + a_{11}n = M_2(n+1).$$

Since $M_2(1) = 0$, it follows that

$$M_2(n) = (n-1)na_{11}/2 = \binom{n}{2}a_{11}.$$

The computer assures us that the results of the next few steps are as follows.

n	M_n where $[n]_F = nT + M_2T^2 + M_3T^3 + \cdots$
2	$\binom{n}{2}a_{11}$
3	$\binom{n}{3}a_{11}^2 + 2\binom{n+1}{3}a_{12}$
4	$\binom{n}{4}a_{11}^3 + \binom{n}{3}(2n+1)a_{11}a_{12}$
	$+\frac{1}{6}\binom{n}{2}\left\{3(n^2-n+2)a_{13} + 2(2n-1)a_{22}\right\}$
5	$\binom{n}{5}a_{11}^4 + \frac{2}{5}\binom{n+1}{4}(8n+1)a_{12}^2$
	$+\frac{1}{10}\binom{n+1}{3}\left\{2(6n^2-15n+16)a_{14} + 5(3n-2)a_{23}\right\}$
	$+\frac{1}{20}\binom{n}{3}\left\{20(2n-1)a_{11}a_{22} + 3(14n^2-13n+13)a_{11}a_{13}\right.$
	$\left.+2(11n^2-27n-8)a_{11}^2a_{12}\right\}$
6	$\binom{n}{6}a_{11}^5 + \frac{1}{15}\binom{n}{4}(26n^2-75n-11)a_{11}^3a_{12}$
	$+\frac{1}{60}\binom{n}{3}\left\{6(22n^3-30n^2+n+23)a_{11}a_{14} + 3(32n^3-105n^2+81n-42)a_{11}^2a_{13}\right.$
	$+3(2n-1)(11n-29)a_{11}^2a_{22} + 6(15n^3-9n^2+8n+12)a_{12}a_{13}$
	$+(68n^3-138n^2-125n+21)a_{11}a_{12}^2 + 12(8n^2-n-4)a_{12}a_{22}$
	$\left.+3(42n^2+n-21)a_{11}a_{23}\right\} + \binom{n}{2}^2 a_{33}$
	$+\frac{1}{30}\binom{n}{2}\left\{5(2n^4-4n^3+n^2+n+6)a_{15} + 2(2n-1)(3n^2-3n+4)a_{24}\right\}$

Substituting $n = -1$ in the first few M_i confirms our earlier formula

$$[-1] = -T + a_{11}T^2 - a_{11}^2T^3 + \left(a_{11}^3 + a_{11}a_{12} + 2a_{13} - a_{22}\right)T^4 + \cdots .$$

Similarly,

$$[2] = 2T + a_{11}T^2 + 2a_{12}T^3 + (2a_{13} + a_{22})T^4 + (2a_{14} + 2a_{23})T^5$$
$$+(2a_{15} + 2a_{24} + a_{33})T^6 + \cdots ,$$

and so on.

The above corollary implies that the derivative

$$[n]' = n + 2M_2T + 3M_3T^2 + \cdots \in nR[[T]], \quad \forall n \in \mathbf{Z}.$$

Hence for $i = 2, 3, \ldots$ and all $n \in \mathbf{Z}$,

$$iM_i(n) \equiv 0 \bmod n \quad (\textit{i.e.,} \bmod nR).$$

These requirements, which are ultimately a consequence of the associative law FG2, produce a sequence of nontrivial conditions:

$$\begin{aligned}
i = 4, n = 3: &\quad a_{11}a_{12} - a_{22} \equiv 0 \bmod 3,\\
i = 5, n = 3: &\quad a_{14} + a_{23} - a_{12}^2 - a_{11}(a_{11}a_{12} - a_{22}) \equiv 0 \bmod 3,\\
i = 5, n = 4: &\quad 2a_{23} + a_{11}a_{13} + 2a_{12}^2 \equiv 0 \bmod 4,\\
\text{etc.}
\end{aligned}$$

2.7.1 The elliptic curve case

Let E be an elliptic curve defined over the field K (of any characteristic), and assume all the relevant notation involved in $\widehat{E}(z_1, z_2) = z_3$ as defined in §2.6.2. To streamline the notation for the next proposition, let us replace x_2, y_2, z_2 by x, y, z respectively, where $x = -yz$. Thus z_1 and z play the roles of S and T of the previous section.

Proposition 2.7.4 *The normalized invariant differential of $\widehat{E}$ is*

$$\frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}.$$

Remark. To have $\omega_{\widehat{E}}$ in the format $\alpha\, dz$ it is necessary to replace the numerators in these fractions by $(dx/dz)dz$ and $(dy/dz)dz$ respectively.
Proof. Differentiate the two equations

$$x + yz = 0, \tag{1}$$

$$y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x_a - 6 = 0 \tag{2}$$

with respect to z (prime denotes derivative with respect to z), using $z = -x/y$ to obtain

$$yx' - xy' = -y^2, \tag{1'}$$

$$(a_1y - 3x^2 - 2a_2x - a_4)x' + (2y + a_1x + a_3)y' = 0. \tag{2'}$$

$(2')$ establishes the equality of the two forms in the proposition. Solving the system $(1'), (2')$ we find

$$\frac{dx/dz}{2y + a_1x + a_3} = \frac{y^2}{a_3 + x^3 - a_4x - 2a_6}.$$

By Proposition 2.7.1 we wish to show that

$$\widehat{E}_1(0, z) = \frac{a_3y + x^3 - a_4x - 2a_6}{y^2}. \tag{3}$$

The left side of this equation, where $z_3 = -x_3/y_3$ and x_3, y_3 are given by the formulas of Proposition 1.71 (with $x_2 = x$ and $y_2 = y$), is the value when $z_1 = 0$ of

$$\frac{\partial z_3}{\partial z_1} = \frac{\partial}{\partial z_1}(-x_3/y_3) = -y_3^{-1}\frac{\partial x_3}{\partial z_1} + x_3 y_3^{-2}\frac{\partial y_3}{\partial z_1}.$$

Noting that $x_3 = x_2$ and $y_3 = y_2$ when $z_1 = 0$, and substituting $x_1 = t_1/z_1^2$ and $y_1 = -t_1/z_1^3$ where $t_1 = 1 - a_1 z_1 - \cdots$ as in §2.5.3, the computer finds that

$$\left.\frac{\partial x_3}{\partial z_1}\right)_{z_1=0} = 2y_2 + a_1 x_2 + a_3.$$

It follows by (2′) that

$$\left.\frac{\partial y_3}{\partial z_1}\right)_{z_1=0} = 3x_2^2 + 2a_2 x_2 + a_4 - a_1 y_2,$$

and equation (3) now follows. ■

We specialize the tables of the previous section to obtain $\omega_{\widehat{E}}$ and $[n]_{\widehat{E}}$ by substituting $a_{11} = -a_1$, etc. as in the table in §2.6.2, and writing z for T; alternatively, the α_i can be calculated directly using the above proposition with $x = t/z^2$, $y = -t/z^3$ and the series for t determined in §2.5.3

i	α_i where $\omega_{\widehat{E}} = (1 + \sum_{i=1}^{\infty} \alpha_i z^i)\, dz$
1	a_1
2	$a_2 + a_1^2$
3	$2a_1a_2 + 2a_3 + a_1^3$
4	$a_2^2 + 3a_1^2a_2 + 6a_1a_3 + 2a_4 + a_1^4$
5	$3a_1a_2^2 + 6a_2a_3 + 4a_1^3a_2 + 12a_1^2a_3 + 6a_1a_4 + a_1^5$
6	$24a_1a_2a_3 + 6a_3^2 + 20a_1^3a_3 + 6a_2a_4 + 12a_1^2a_4 + 3a_6 + a_2^3$
	$+6a_1^2a_2^2 + 5a_1^4a_2 + a_1^6$
7	$60a_1^2a_2a_3 + 24a_1a_2a_4 + 12a_1a_6 + 30a_1a_3^2 + 4a_1a_2^3 + 30a_1^4a_3$
	$+20a_1^3a_4 + 10a_1^3a_2^2 + 6a_1^5a_2 + 12a_2^2a_3 + 12a_3a_4 + a_1^7$

n	M_n where $[n]_{\widehat{E}} = nz + \sum_{n\geq 2} M_n z^n$
2	$-\binom{n}{2}a_1$
3	$\binom{n}{3}(a_1^2 - 2a_2)$
4	$-\binom{n}{4}a_1^3 + \binom{n+1}{3}(2n-3)a_1a_2 - \binom{n}{2}(n^2+n+1)a_3$
5	$\binom{n}{5}a_1^4 - \frac{2}{5}\binom{n+1}{4}(11n-18)a_1^2a_2 + \frac{1}{5} + \frac{2}{5}\binom{n+1}{3}\left\{(2n^2-3)a_2^2 - 6(n^2+1)a_4\right\}$
	$+\frac{1}{5}\binom{n}{2}(3n^3 - 7n^2 - 7n - 12)a_1a_3$
6	$-\binom{n}{6}a_1^5 + \frac{1}{15}\binom{n+1}{4}(2n-5)(13n-24)a_1^3a_2$
	$+\frac{1}{15}\binom{n+1}{3}\left\{18(2n-5)(n^2+1)a_1a_4 - (17n^3 - 30n^2 - 23n + 45)a_1a_2^2\right\}$
	$+\frac{1}{30}\binom{n}{2}\left\{10(n^4 + n^3 - 3n^2 - 6n - 6)a_2a_3\right.$
	$\left.-(8n^4 - 37n^3 + 48n^2 + 33n + 120)a_1^3a_3\right\}$

Of course now the congruences $iM_i(n) \equiv 0 \bmod n$ are automatically satisfied since the Weierstrass coefficients can be chosen arbitrarily.

Again one can assign any integer value to n, *e.g.,*

$$[-1] = -z - a_1z^2 - a_1^2z^3 - (a_1^3 + a_3)z^4 - (a_1^4 + 3a_1a_3)z^5$$

$$-(a_1^5 + 6a_1^2a_3 + a_2a_3)z^6 + \cdots .$$

We tabulate a few more cases where $\pm n$ is a prime number (*cf.* the λ, μ notation of Corollary 2.7.3):

$$\begin{aligned}
[2] &= 2\left\{z - a_2z^3 - (6a_4 + 3a_1a_3 - a_2^2)z^5 + \cdots\right\} \\
&\quad + \left\{-a_1z^2 + (a_1a_2 - 7a_3)z^4 + (-6a_1a_4 - 2a_2a_3 - 7a_1^2a_3 - a_1a_2^2)z^6 + \cdots\right\} \\
[3] &= 3\left\{z - a_1z^2 + (4a_1a_2 - 13a_3)z^4 - (32a_4 + 3a_1a_3 - 8a_2^2 + 2a_2a_1^2)z^5 + \cdots\right\} \\
&\quad + \left\{(a_1^2 - 8a_2)z^3 + (48a_1a_4 + 57a_2a_3 - 30a_1^2a_3 - 44a_1a_2^2 + a_2a_1^3)z^6 + \cdots\right\} \\
[5] &= 5\left\{z - 2a_1z^2 + 2(a_1^2 - 4a_2)z^3 - (62a_3 + a_1^3 - 28a_1a_2)z^4\right. \\
&\quad \left. + (41a_2a_1^3 + 624a_1a_4 + 426a_2a_3 - 124a_1^2a_3 - 348a_1a_2^2)z^6 + \cdots\right\} \\
&\quad + \left\{(-1248a_4 + 306a_1a_3 + 376a_2^2 + a_1^4 - 222a_2a_1^2)z^5 + \cdots\right\}
\end{aligned}$$

$$\begin{aligned}
[-2] &= (-2)\left\{z + (2a_1^2 - a_2)z^3 + (3a_1^4 - 8a_2a_1^2 + 15a_1a_3 + a_2^2 - 6a_4)z^5 \cdots\right\} \\
&\quad + \left\{-3a_1z^2 + (-5a_1^3 + 7a_1a_2 - 9a_3)z^4\right. \\
&\qquad \left. + (-7a_1^5 + 30a_2a_1^3 - 67a_1^2a_3 - 11a_1a_2^2 + 54a_1a_4 + 2a_2a_3)z^6 + \cdots\right\} \\
[-3] &= (-3)\left\{z + 2a_1z^2 + (5a_1^3 - 12a_1a_2 + 14a_3)z^4\right. \\
&\qquad \left. + (7a_1^4 - 34a_2a_1^2 + 54a_1a_3 + 8a_2^2 - 32a_4)z^5 + \cdots\right\} \\
&\quad + \left\{(8a_2 - 10a_1^2)z^3\right. \\
&\quad \left. + (-28a_1^5 + 231a_2a_1^3 - 420a_1^2a_3 - 164a_1a_2^2 + 528a_1a_4 + 78a_2a_3)z^6 + \cdots\right\}
\end{aligned}$$

2.8 Formal groups in characteristic p

Let A be a commutative ring of characteristic p where p is a prime number, and let $f : F \longrightarrow G$ be a formal group morphism defined over A. If $f \neq 0$ the **height** of f is the largest integer $h \geq 0$ such that

$$f(T) = f_1(T^{p^h})$$

for some $f_1 \in A[[T]]$. The notation used is $\mathrm{ht}(f) = h$. It is convenient to assign $\mathrm{ht}(0) = \infty$. The **height** of a formal group F is defined to be $\mathrm{ht}(F) = \mathrm{ht}([p]_F)$. Since $p = 0$ in A, by Corollary 2.7.3 we have $ht(F) \geq 1$. Here are two examples:

$$[p]_{\mathbf{A}} = pT = 0, \quad \text{hence} \quad \mathrm{ht}(\mathbf{A}) = \infty;$$

$$[p]_{\mathbf{M}} = (T+1)^p - 1 = T^p, \quad \text{hence} \quad \mathrm{ht}(\mathbf{M}) = 1.$$

Let $h = \mathrm{ht}(f)$ and $f(T) = f_1(T^{p^h})$. Since the derivative

$$f'(T) = p^h T^{p^h - 1} f_1'(T^{p^h}),$$

therefore $h > 0$ iff $f'(T) = 0$. Of course the latter implies that $f'(0) = 0$. What is not immediately obvious is that the converse is true:

Proposition 2.8.1 *Let $f : F \longrightarrow G$ be a nonzero morphism of formal groups defined over A where* $\mathrm{char}\, A = p$ *is a prime number. Let $h = \mathrm{ht}(f)$ and $f(T) = f_1(T^{p^h})$.*

(a) $\mathrm{ht}(f) > 0$ *iff* $f'(0) = 0$.

(b) $f_1'(0) \neq 0$.

(c) *Assume for this part that A is an integral domain. If $g : G \longrightarrow H$ is also a morphism defined over A then*

$$\mathrm{ht}(g \circ f) = \mathrm{ht}(f) + \mathrm{ht}(g).$$

Proof. (a) If $f'(0) = 0$ then by Corollary 2.7.2

$$(\alpha_G \circ f)f' = (1 + \cdots)f'(T) = 0,$$

hence $f'(T) = 0$.

(b) We deduce (b) from (a) by showing that f_1 is a morphism of height 0. Let $q = p^h$ and define

$$F_1 = F^q = \left(\sum a_{ij} S^i T^j\right)^q = \sum a_{ij}^q S^{qi} T^{qj}.$$

F_1 can be regarded as a series in the two variables $S_1 = S^q$, $T_1 = T^q$, and clearly $F_1(S_1, T_1)$ is a formal group — in the axioms FG1–FG3 for F, raise the equations to the power q. Moreover, $f_1 \in \hom_A(F_1, G)$ as we now verify:

$$\begin{aligned} f_1(F_1(S_1, T_1)) &= f_1(F(S,T)^q) \\ &= f(F(S,T)) \\ &= G(f(S), f(T)) \\ &= G(f_1(S^q), f_1(T^q)) \\ &= G(f_1(S_1), f_1(T_1)). \end{aligned}$$

If $k = \mathrm{ht}(f_1)$ and $f_1(T) = f_2(T^{p^k})$ then $f(T) = f_2(T^{p^{h+k}})$. Thus $k = 0$ since h is maximal.

(c) By (b) we have $f_1(T) = aT + \cdots$ where $a \neq 0$, and similarly $g_1(T) = bT + \cdots$ where $b \neq 0$ and $g(T) = g_1(T^{p^m})$ where $m = \mathrm{ht}(g)$. Now

$$(g \circ f)(T) = g_1\left(\left(f_1\left(T^{p^h}\right)\right)^{p^m}\right) = abT^{p^{h+m}} + \cdots$$

is of the form $s\left(T^{p^{h+m}}\right)$ where $s(T) \in A[[T]]$ and $s'(0) = ab \neq 0$. Hence $\mathrm{ht}(g \circ f) = h + m$. ■

We will take up this topic again in Chapter 6 where we will prove the following. Let E be an elliptic curve defined over an integral domain of characteristic $p > 0$. Then $\mathrm{ht}(\widehat{E}) = 1$ or 2:

- if E is ordinary then $\mathrm{ht}(\widehat{E}) = 1$;
- if E is supersingular then $\mathrm{ht}(\widehat{E}) = 2$.

For example in characteristic 2, from the previous section we have

$$[2] = a_1 T^2 + (a_1 a_2 + a_3)T^4 + a_1(a_1 a_3 + a_2^2)T^6 + \cdots .$$

Recall that E is ordinary (resp. supersingular) when $a_1 \neq 0$ (resp. $a_1 = 0$ and then $a_3 \neq 0$ in order that $\Delta \neq 0$).

2.9 Formal groups in characteristic 0

2.9.1 The formal logarithm

Recall from §2.3 that when the commutative ring A is flat we can regard it as a subring of $A_{\mathbf{Q}} = A \otimes_{\mathbf{Z}} \mathbf{Q}$.

Let A be flat and let F be a formal group over A with differential

$$\omega_F = \frac{1}{F_1(0,T)} dT = (1 - a_{11}T + \cdots)dT.$$

The **formal logarithm** of F is the formal antiderivative

$$\log_F(T) = \int \omega_F = T - \frac{a_{11}}{2}T^2 + \frac{a_{11}^2 - a_{12}}{3}T^3 \cdots \in A_{\mathbf{Q}}[[T]] \ .$$

Note that $\log_F$ is usually not in $A[[T]]$. The **formal exponential**, denoted $\exp_F$, is defined to be the reverse series in $\mathcal{R}^*{}_1(A_{\mathbf{Q}}, T)$; by Proposition 2.3.3 it has the form

$$\exp_F = \log_F^{(-1)} = \sum_{n \geq 1} \frac{b_n}{n!} T^n, \quad b_n \in A.$$

Using the "generic reversion" formula of §2.3 (or more easily by using Maple or Mathematica on the computer) we find that the first few values of b_n are as follows. (The coefficients α_{n-1}/n in $\log_F$ are immediately obtained from the tables of α's given in §§2.6–2.6.1.)

n	b_n where $\exp_F = \sum_{n=1}^{\infty}(b_n/n!)T^n$
1	1
2	a_{11}
3	$a_{11}^2 + 2a_{12}$
4	$a_{11}^3 + 8a_{11}a_{12} + 6a_{13}$
5	$a_{11}^4 + 22a_{11}^2a_{12} + 42a_{11}a_{13} + 16a_{12}^2 + 24a_{14}$
6	$a_{11}^5 + 52a_{11}^3a_{12} + 192a_{11}^2a_{13} + 136a_{11}a_{12}^2$
	$+264a_{11}a_{14} + 180a_{12}a_{13} + 120a_{15}$

We already mentioned the explicit examples $\log_{\mathbf{M}}$ and $\exp_{\mathbf{M}}$ in §2.6.1.

Proposition 2.9.1 *Let A be flat and F a formal group over A. Let $\mathbf{A} = S + T$ denote the additive formal group as usual. Then*

$$\log_F : F \longrightarrow \mathbf{A}, \quad \exp_F : \mathbf{A} \longrightarrow F$$

are strict isomorphisms, inverse to each other, defined over $A_{\mathbf{Q}}$.

Hence over a $\mathbf{Q}$*-algebra, all formal groups are strictly isomorphic, all being strictly isomorphic to the additive formal group;*[†] *in other words, the set of all formal groups is*

$$\left\{ f^{(-1)}\left(f(S)+f(T)\right) : f \in \mathcal{R}^* \right\}.$$

Proof. Let $\omega_F = \alpha\, dT$. By Proposition 2.7.1,

$$(\alpha(F(S,T))\frac{\partial F(S,T)}{\partial T}\, dT = \alpha(T)\, dT.$$

Formal integration gives

$$\log_F F(S,T) = \log_F(T) + C(S)$$

where $C(S) \in A_{\mathbf{Q}}[[S]]$ is a "constant of integration". Substituting $T = 0$ shows that $C(S) = \log_F(S)$, hence $\log_F \in \hom_{A_{\mathbf{Q}}}(F, \mathbf{A})$. Since $c(\log_F) = 1$, in the notation of §2.6, log is a strict isomorphism, and its inverse is $\exp_F$ by definition. ■

For an elliptic curve E, the beginning of the series for $\exp_{\widehat{E}}$ can be obtained by substituting $a_{11} = -a_1$, etc in $\exp_F$, or by calculating $\log_{\widehat{E}}^{(-1)}$ directly. Then

$$\widehat{E} = \exp_{\widehat{E}}\left(\log_{\widehat{E}}(S) + \log_{\widehat{E}}(T)\right),$$

which is probably the most efficient way of calculating the series $\widehat{E}$.

2.9.2 Formal groups over discrete valuation rings

Recall that an element z in an abelian group is a **torsion element** if $[n]z = 0$ for some positive integer n; then the smallest such n is the **order** of z. Thus 0 is a torsion element of order 1. The torsion elements form a subgroup.

Proposition 2.9.2 *Let* V *be a complete discrete valuation ring of characteristic* 0*, with residue field* k *of characteristic* $p \geq 0$*, and let* M *denote its maximal ideal. Let* F *be a formal group defined over* V*, and let* $\mathcal{T}$ *denote the torsion subgroup of* $F(M)$*.*

(a) *Suppose* $p > 0$*. Then*

- $\mathcal{T}$ *is a finite* p*-group of order*

$$|\mathcal{T}| \leq \frac{p}{p-1} v(p);$$

[†]This does not mean that all formal groups over $\mathbf{Q}$, for example, are "the same". This point will be made forcibly when we discuss the Atkin, Swinnerton-Dyer congruences in a later chapter.

- *if $z \in \mathcal{T}$ then*
$$v([p]z) \geq pv(z);$$
for nontorsion elements we have only
$$v([p]z) \geq \min\{v(p) + v(z),\, pv(z)\};$$

- *if z is a nonzero element of $\mathcal{T}$, say the order of z is p^n, then*
$$v(z) \leq \frac{v(p)}{p^n - p^{n-1}}.$$

(b) $\log_F$ *induces a group homomorphism*
$$F(M) \longrightarrow K^+$$
into the additive group of K; the kernel contains $\mathcal{T}$.

(c) *Define*
$$n_0 = \begin{cases} 0 & \text{if } \ p = 0, \\ v(p)/(p-1) & \text{if } \ p > 0. \end{cases}$$
Then for $n > n_0$, $\log_F$ induces a group isomorphism
$$F(M^n) \longrightarrow M^n,$$
addition in the latter group being ordinary ring addition. The inverse isomorphism is induced by $\exp_F$. Hence

- *for $n > n_0$, $F(M^n)$ is torsion-free; in particular, $\mathcal{T} = 0$ when $p = 0$;*

- *when $p > 0$ and $v(z) > n_0$,*
$$v([p]z) = v(p) + v(z).$$

Proof. (a) The finiteness of $\mathcal{T}$ was first proved (as far as I know) in [Fle-Oes90], and we follow their method.

By Proposition 2.6.3, $\mathcal{T}$ is a p-group. Let $\mathcal{H}$ be any finite subgroup of $\mathcal{T}$, and let its exponent be p^n so that $\mathcal{H}' = \mathcal{H}[p^{n-1}]$ (the elements in $\mathcal{H}$ annihilated by p^{n-1}) is a subgroup of index $\geq p$. Let $\mathcal{H} - \mathcal{H}' = \{z_1, \ldots, z_h\}$, where $h = |\mathcal{H}| - |\mathcal{H}'| \geq (1 - 1/p)|\mathcal{H}|$.

Write $[p](T) = Tu(T)$ where $u(0) = p$, and define $w(T) = u([p^{n-1}](T))$ so that
$$w(0) = p \quad \text{and} \quad [p^n](T) = [p]([p^{n-1}](T)) = [p^{n-1}](T)w(T).$$
Since $[p^n](z_i) = 0 = [p^{n-1}](z_i)w(z_i)$ and $[p^{n-1}](z_i) \neq 0$, we have $w(z_i) = 0$, $\forall i$. By Proposition 2.4.2,
$$w(T) = (z_1 - T)(z_2 - T)\cdots(z_h - T)g(T) \quad \text{where} \ \ g(T) \in V[[T]].$$

Putting $T = 0$ we deduce that

$$v(p) \geq v(z_1) + \cdots + v(z_h) \geq h \geq (1 - 1/p)|\mathcal{H}|.$$

Since $\mathcal{H}$ is an arbitrary finite subgroup of $\mathcal{T}$, it follows that $\mathcal{T}$ itself is finite and of order $\leq v(p)/(1 - 1/p)$.

Next, take $\mathcal{H}$ to be the cylic subgroup generated by an element z of order p^n. Then each of the $h = \phi(p^n) = p^n - p^{n-1}$ elements $z_i = [m_i]z = m_i z + \cdots$, where $1 \leq m_i < p^n$ and $p \nmid m_i$, has value $v(z_i) = v(z)$. Thus $v(p) \geq v(z_1) + \cdots + v(z_h)$ gives the estimate $v(p) \geq (p^n - p^{n-1})v(z)$.

Earlier proofs of this estimate (*e.g.* in [Sil86, p.123]) are based on Corollary 2.7.3:

$$[p](T) = p\lambda(T) + \mu(T^p)$$

for certain $\lambda, \mu \in V[[T]]$, and the expansion of λ begins $T + \cdots$. Thus for $z \in M$,

$$v(p\lambda(z)) = v(p) + v(z + \cdots) = v(p) + v(z) \quad \text{and} \quad v(\mu(z^p)) \geq pv(z),$$

and therefore

$$v([p]z) \geq \min\{v(p) + v(z), pv(z)\}. \qquad \P$$

The earlier proofs now proceed by induction on the order p^n. However now we already know that

$$v(z) \leq v(p) \,/ p^{n-1}(p-1) \leq v(p)/(p-1),$$

hence in ¶ the minimum is $pv(z)$.

For the proof of (b) and (c) we require the

Lemma 2.9.3 *As in the proposition, let K be complete with respect to the valuation v, with ring V and maximal ideal M, and of characteristic 0 with residue characteristic $p \geq 0$.*

(i) *A series of the form*

$$a(x) = \sum_{n=1}^{\infty} \frac{a_n}{n} z^n, \quad a_n \in V$$

converges to an element in K for $z \in M$;

(ii) *a series of the form*

$$b(x) = \sum_{n=1}^{\infty} \frac{b_n}{n!} z^n, \quad b_n \in V$$

converges in K for z satisfying $v(z) > 0$ if $p = 0$ or $v(z) > v(p)/(p-1)$ if $p > 0$. If $b_1 \in V^*$ then

$$v(b(z)) = v(z).$$

Proof of the lemma. Recall that to prove convergence one proves that the value of the n-th term $\to \infty$ as $n \to \infty$. Thus the convergence of both series is obvious when $p = 0$: $v(n) = 0$ for all $n \geq 1$, so $v(a_n z^n/n) \geq nv(z)$ and $v(b_n z^n/n!) \geq nv(z)$. The last statement in (ii) is also clear when $p = 0$: for $n \geq 2$, $v(b_n z^n/n!) \geq 2v(z) > v(b_1 z)$ when $v(b_1) = 0$. Thus suppose $p > 0$.

(i) $v(a_n z^n/n) \geq nv(z) - v(n) \to \infty$ since $v(z) > 0$ and $v(n)$ is at most the real logarithm of n to the base p.

(ii)

$$\begin{aligned} v(b_n z^n/n!) &\geq nv(z) - (n-1)v(p)/(p-1) \quad \text{using Corollary 2.6.6} \\ &= v(z) + (n-1)\left(v(z) - \frac{v(p)}{p-1}\right) \\ &\longrightarrow \infty \quad \text{by the assumption on } v(z). \end{aligned}$$

Also when $v(b_1) = 0$,

$$v(b_n z^n/n!) > v(z) = v(b_1 z) \quad \text{for } n \geq 2,$$

so the value of the series is $v(z)$.

We now complete the proof of the proposition. As we saw in §2.9.1, $\log_F$ and $\exp_F$ are examples of the types of series in (i) and (ii) respectively.

(b) By (i), $\log_F(z)$ converges for $z \in M$. Since

$$\log_F F(z_1, z_2) = \log_F(z_1) + \log_F(z_2)$$

is true as a power series identity, it is true for the covergent series when $z_1, z_2 \in M$. Thus $z \mapsto \log_F(z)$ is a homomorphism. Its kernel must contain all torsion elements of $F(M)$ since $\mathrm{char} K = 0$ and therefore K^+ is torsion free.

(c) By parts (i) and (ii) of the lemma, $\log_F(z)$ and $\exp_F(z)$ converge for $z \in M^n$. We have seen that $\log_F$ is homomorphic, and the identities

$$\log_F(\exp_F(z)) = z = \exp_F(\log_F(z))$$

complete the proof. ■

Here is an example of a nontorsion point with $v([p]z) = v(p) + v(z) < pv(z)$: over the 3-adic field,

$$y^2 + y = x^3 + x^2 \qquad \textbf{A43}$$

has the point $P = (-2/9, 1/27)$. In the notation of §2.1.2,

$$v(P) = v(z) = -v(x)/2 = 1,$$

and $$v([3]P) = -v(13^3 37 \cdot 8191 \cdot 3^{-4} 11^{-2} 59^{-2})/2 = 2.$$

Similarly, for

$$Q = (369/64, 4941/512) \quad \text{on} \quad y^2 + xy + y = x^3 - x^2, \qquad \textbf{A53}$$

for the 2-adic valuation one finds

$$v(Q) = 3, \qquad v([2]Q) = 4.$$

Here is part (c) of the proposition stated for the elliptic curve case.

Corollary 2.9.4 *Let E be defined over the valuation ring V with quotient field K of characteristic 0, and residue field k of characteristic $p \geq 0$; let v denote the valuation map v, and $E(K) \supset E_1(K) \supset \cdots$ denote the v-adic filtration. Then $E_m(K)$ is torsion-free for*

- *$m > 0$ if $p = 0$,*
- *$m > v(p)/(p-1)$ if $p > 0$.*

It is not necessary to assume that K is complete — we can regard $E_m(K)$ as a subgroup of $E_m(\widehat{K})$.

2.10 The Nagell-Lutz theorem for Krull domains

The original theorem refers to elliptic curves defined over $\mathbf{Z}$ and is usually stated as follows: if $P = (s, t)$ is a non-zero rational point of finite order on

$$E : y^2 = x^3 + bx + c \quad \text{where} \quad b, c \in \mathbf{Z}, \tag{1}$$

then $s, t \in \mathbf{Z}$ and either $t = 0$ (then P has order 2) or $t^2 | 4b^3 + 27c^2$. This form of the theorem, as given by Nagell [Nag35] and Lutz [Lut37], is awkward to apply, and can be inefficient, when one or more of the coefficients a_1, a_2, a_3 is nonzero. For in general one must first complete the square in y, then the cube in x and finally clear denominators. The problems are illustrated by the example

$$Y^2 + XY + Y = X^3. \qquad \textbf{A26}$$

Substituting $X = (x-3)/36$, $Y = (y - 3x - 99)/216$ gives an equation of the form (1) with

$$b = 27 \cdot 23, \quad c = 54 \cdot 181, \quad 4b^3 + 27c^2 = 2^9 3^{12} 13.$$

This yields an uncomfortably large class of candidate divisors t^2. The 'right' way to state Nagell-Lutz for $\mathbf{Z}$ is given in Proposition 2.10.4 below; for this example it reduces the number of candidates to *two*. But first we prove some general results.

Nagell-Lutz type theorems involve two estimates for the coordinates x, y of torsion points: one which bounds the denominators, the other the numerators. In the original theorem quoted above, the denominators are 1; in general the denominator bound is obtained by applying Proposition 2.9.2 to $\widehat{E}$. (Nagell and Lutz had less precise information since formal group theory was not yet well developed.) The bound for the numerators comes from the identity proved in Proposition 1.7.12(b); in the theorem quoted above it takes the form $y^2|(4b^3+27c^2)$.

The first step is to bound denominators.

A point P on an elliptic curve defined over an integral domain A is said to be **integral**, or if necessary, A**-integral**, when

(i) $P \neq O$, say $P = (x, y)$, and

(ii) $x \in A$, $y \in A$.

When A is integrally closed, *e.g.* when A is Krull, it is sufficient to check that one of x, y is in A because of the Weierstrass equation.

Proposition 2.10.1 *Let A be a Krull domain, let E be an elliptic curve defined over A, let K denote the quotient field of A, and let $P = (x, y)$ be a nonzero point in $E(K)$.*

(a) *If $[n]P$ is integral for some $n \in \mathbf{Z}$ then P is integral.*

(b) *Now let P be a torsion element of order $m \geq 2$. Then either P is integral, or m is a prime power, say $m = p^n$.*

Suppose $x \notin A$. Then

- *if* $\operatorname{char} K > 0$, *then* $p = \operatorname{char} K$;
- *if* $\operatorname{char} K = 0$, *then for all essential valuations v of A,*

$$v(x) < 0 \Longrightarrow p = \operatorname{char} k \quad \textit{where } k \textit{ is the residue field of } v, \textit{ hence} \quad v(p) > 0,$$

$$\textit{and} \quad v(x) \geq -2\left\lfloor v(p)/(p^n - p^{n-1})\right\rfloor \quad \textit{(integral part).}$$

Hence $px \in A$ except possibly when $p = 2$, $n = 1$ and then $4x \in A$.

Proof. (a) If P is not integral then $x \notin A$, which is to say $v(x) < 0$ for some essential valuation v, and then in fact $v(P) = \nu$ where $v(x) = -2\nu$, $v(y) = -3\nu$, and $P \in E_\nu$ for the filtration associated to v, as explained in §2.2.2. We can regard E as being defined over the v-adic completion of K. Since E_ν is a subgroup, $[n]P \in E_\nu$ for all $n \in \mathbf{Z}$ and therefore $[n]P$ is not integral.

(b) The case where char $K > 0$ is covered by Proposition 2.6.3. Thus assume char K=0.

Suppose $[m]P = O$ and $v(x) < 0$. Then applying Proposition 2.9.2 to $F = \widehat{E}$, we see that $m = p^n$ where $p > 0$ is the residue characteristic, so $v(p) > 0$, and $n \geq 1$. The lower bound for $v(x) = -2v(z)$ follows from that proposition. Then $v(px) \geq 0$, or $v(4x) \geq 0$ when $p^n = 2$, are simple deductions. ■

Examples

1. The two variable polynomial ring $A = \mathbf{Z}[a, b]$ is a UFD, hence is Krull. The curve

$$E : y^2 + abxy + by = x^3 + ax^2 + x$$

is elliptic since Δ is a nonzero polynomial, and $(0, 0)$ is a point of infinite order because

$$[-2](0, 0) = (1/b^2, 1/b^3)$$

has infinite order, by the proposition with the b-adic valuation.

Under most specializations the order of $(0, 0)$ remains infinite. *E.g.* if we set $b = 3$, then Δ is a nonzero polynomial in a, and $(1/9, 1/27)$ has infinite order by applying the proposition with the 3-adic valuation on $A = \mathbf{Z}[a]$. (By Gauss's Lemma, Proposition 2.1.7, the 3-adic valuation on $\mathbf{Z}$ extends to $\mathbf{Z}[a]$.)

2. Let us take example 1 with $a = b^2$ over the polynomial ring $\mathbf{F}_2[b]$ in characteristic 2:

$$y^2 + b^3xy + by = x^3 + b^2x^2 + x, \quad \Delta = b^4 \pmod 2.$$

Of course we still have

$$[2](0, 0) = (1/b^2, 1/b^3),$$

but now this is the unique point of order 2, by Proposition 1.7.3; this is consistent with the proposition since $\operatorname{char} K = 2$. Thus $(0, 0)$ and $-(0, 0) = (0, b)$ are points of order 4.

3. Over the polynomial ring $\mathbf{Z}[a]$, $y^2 - y = x^3 + (4a - 1)x - a$ has $\Delta = -(256a - 37)(4a - 1)^2$. By the proposition, if $(1/4, 3/8)$ has finite order then the order must be 2; but that is not the case since

$$-(1/4, 3/8) = (1/4, 5/8).$$

By the same token, $(1/4, 3/8)$ has infinite order when we specialize a to any value in $\mathbf{Z}$.

Now let us describe the second ingredient in Nagell-Lutz type theorems. From Proposition 1.7.12(b), if E is defined over any field K and $P = (x, y) \in E(K)$ is such that $[2]P \neq O$, not necessarily of finite order, then

$$(\sigma - x([2]P)\tau)\kappa^2 = \Delta \qquad \P$$

where

$$\begin{aligned}
\kappa^2 &= (2y + a_1x + a_3)^2 \\
&= 4x^3 + b_2x^2 + 2b_4x + b_6, \\
\sigma &= 12x^3 - b_2x^2 - 10b_4x + b_2b_4 - 27b_6, \\
\tau &= 48x^2 + 8b_2x + 32b_4 - b_2^2.
\end{aligned}$$

The subscript 2 has been dropped from σ and τ to simplify the notation. Recall that when $\operatorname{char} K \neq 2$, $\kappa = 2\eta$ and then a nonzero point has order 2 iff $\eta = 0$ (*cf.* Proposition 1.7.3).

This implies for example that if E is defined over the integral domain A and the points P and $[2]P$ are both integral (hence $\sigma, \tau \in A$ and $[2]P \neq O$), then $\kappa^2 | \Delta$ in A, and, when also $b_6 = 0$ (hence $x(P) \neq 0$ since $[2]P \neq O$) then $x(P) | \kappa^2$ and therefore $x(P) | \Delta$.

We introduce some convenient notation that will be used not only in the rest of this chapter, but will be standard for the rest of the notes:

$\mathcal{T}$ denotes the torsion subgroup of $E(K)$, and $\mathcal{T}^*$ the set $\mathcal{T}$ with O removed.

When necessary we write $\mathcal{T}_{E/K}$.

When A is Krull of characteristic 0 and $P \in \mathcal{T}^*$ has order > 2, then also $[2]P \in \mathcal{T}^*$ and we can use the lower bound on $v(x([2]P))$ from the previous proposition:

Proposition 2.10.2 *Let E be an elliptic curve defined over the integral domain A with quotient field K, and let $P = (x, y) \in E(K)$ where $[2]P \neq O$.*

(a) *If P and $[2]P$ are A-integral then $\kappa^2 | \Delta$, and if also $b_6 = 0$ then $x \neq 0$ and $x | \Delta$.*

(b) *Now let A be a Krull domain of characteristic 0 and P an integral point in $\mathcal{T}^*$. Then for all essential valuations v,*

$$2v(\kappa) \leq \begin{cases} v(\Delta) & \text{if } v(x([2]P)) \geq 0, \\ v(\Delta) + 2\left\lfloor v(p)/(p^n - p^{n-1}) \right\rfloor & \text{if } v(x([2]P)) < 0, \end{cases}$$

where p^n denotes the order of $[2]P$ in the latter case. ■

The simplest Nagell-Lutz type theorem that avoids all denominator problems is the following.

Proposition 2.10.3 *Let A be a Krull $\mathbf{Q}$-algebra, let E be defined over A and let $P \in \mathcal{T}^*$. Then P is A-integral, hence $\eta \in A$, and either $\eta = 0$ (when P is of order 2) or $\eta^2 | \Delta$.*

Remark. Since 2 and 3 are invertible in A we can take the equation of E in the form $y^2 = x^3 + bx + c$ and then the statement becomes

$$y = 0 \quad \text{or} \quad y^2 |\, 4b^3 + 27c^2;$$

when there is a point of order 2 we can take the equation in the form $y^2 = x(x^2 + ax + b)$ and then the statement is

$$y = 0 \quad \text{or} \quad y^2 |\, b^2(a^2 - 4b).$$

Proof. Since A is a $\mathbf{Q}$-algebra, $\operatorname{char} K = 0$ and for the residue field k of every essential valuation, $\operatorname{char} k = 0$. By Proposition 2.10.1, every nonzero torsion point is integral. The rest follows from ¶ since 2 is now a unit. ■

Example 1 Let t be an indeterminate and

$$E : y^2 = x^3 - t(t-1)^2 x, \quad \Delta = 64t^3(t-1)^6, \quad j = 1728.$$

$T = (0,0)$ is a point of order 2. We will show that $\mathcal{T}_{E/K} = \{O, T\}$ where K is the field $\mathbf{C}(t)$ by applying the proposition to the Krull $\mathbf{Q}$-algebra $\mathbf{C}[t]$; it follows that $\mathcal{T}_{E/K} = \{O, T\}$ when K is $\mathbf{Q}(t)$ or $\mathbf{R}(t)$, for example, or even $K = \mathbf{Q}_p(t)$ (p any prime) since there are field embeddings $\mathbf{Q}_p(t) \hookrightarrow \mathbf{C}(t)$.

Since 0 is the only root of $x^3 - t(t-1)^2 x$ in $\mathbf{C}(t)$, if (x, y) is a torsion point distinct from O and T, then $x, y \in \mathbf{C}[t]$, $y \neq 0$ (hence $x \neq 0$) and $y^2 \mid t^3(t-1)^6$, and in particular, $\deg y \leq 4$. From the equation for E we deduce that

$$2 \deg y = \begin{cases} \deg x + 3 & \text{if} \quad \deg x < 2, \\ 3 \deg x & \text{if} \quad \deg x \geq 2. \end{cases}$$

Hence the possibilities for $(\deg x, \deg y)$ are $(1, 2)$ and $(2, 3)$. From $y^2 = x(x^2 - t(t-1)^2)$ we see that if y is divisible by t or by $t-1$, then so is x, and any irreducible divisor of x also divides y. Combining these observations with $y^2 \mid t^3(t-1)^6$ and the degree constraints leaves the possibilities

$$\begin{array}{ll} x = c(t-1), & y = d(t-1)^2; \\ x = ct(t-1), & y = dt(t-1)^2; \\ x = c(t-1)^2, & y = d(t-1)^3, \end{array}$$

where $c, d \in \mathbf{C}^*$. (*Prima facie,* $\mathcal{T}$ might be infinite.) Substituting into the equation, the last one is quickly eliminated, and the first two uncover the points (and only these points)

$$\pm P_1 = (1-t, \pm(t-1)^2) \quad \text{and} \quad \pm P_1 + T = (t(t-1), \pm t(t-1)^2).$$

Now
$$[2]P_1 = ((t+1)^2/4, (t+1)(t^2 - 6t + 1)/8)$$

is of infinite order since $y^2 \nmid \Delta$, hence so is P_1. Thus $\mathcal{T} = \{O, T\}$.

If we extend the field to $\mathbf{C}(t, t')$ where $t'^4 = t(t-1)^2$, and substitute $x = t'^2 x'$, $y = t'^3 y'$, the equation becomes $y'^2 = x'^3 - x'$ whose division polynomials are all defined over $\mathbf{Z}$. Now $\mathcal{T}$ is infinite (in a later chapter we will see that it is the divisible group $(\mathbf{Q}/\mathbf{Z}) \oplus (\mathbf{Q}/\mathbf{Z})$) which is typical of **constant** E, *i.e.,* j is in the field of constants $\mathbf{C}$ of the function field; see the example in §4.3.

Example 2 We leave the necessary calculations for this example to the reader.

$$y^2 + 6xy + (t+3)(t^2+3)y = x^3, \quad \Delta = \left[-3(t+1)(t+3)(t^2+3)\right]^3,$$

or $$\eta^2 = x^3 + 9x^2 + 3(t+3)(t^2+3)x + \frac{1}{4}(t+3)^2(t^2+3)^2.$$

Again $\mathcal{T}$ denotes the torsion subgroup of $E(\mathbf{C}(t))$.

There is no solution x with $\eta = 0$: by degree considerations we would have $x = c_0 + c_1 t + c_2 t^2$, and then substituting and comparing coefficients shows there is no solution.

Thus for $(x, \eta) \in \mathcal{T}^*$ we have $\eta \in \mathbf{C}[t]$, $\eta \neq 0$, and $\eta^2 | \Delta$. One finds that

$$\mathcal{T} = C_3 \oplus C_3,$$

where C_3 is the cyclic group of order 3, and the 8 points of order 3 in (x, η)-coordinates are as follows. Let $\zeta = (-1 + \sqrt{-3})/2$.

$$\begin{aligned}
P_1 &= \left(0, (t+3)(t^2+3)/2\right), \\
P_2 &= \left(-(t^2+3), \sqrt{-3}(t+1)(t^2+3)/2\right), \\
P_3 = P_1 + P_2 &= \left(-\zeta(t+3)(t-\sqrt{-3}), \sqrt{-3}(t+1)(t+3)(t-\sqrt{-3})/2\right), \\
P_4 = P_1 - P_2 &= \overline{P_3} \ \text{(complex conjugate)},
\end{aligned}$$

and the negatives $-P_i = (x_i, -\eta_i)$ of these four points.

This is a non-constant curve with

$$j = 2^{12} 3 \frac{t^3(t^2+3t+3)^3}{(t+1)^3(t+3)^3(t^2+3)^3},$$

and it can be shown that in every finite extension of $\mathbf{C}(t)$, $\mathcal{T}$ remains finite; *cf.* the discussion in the opening section of the next chapter.

2.10.1 Nagell-Lutz for Z

In the following proposition we combine Nagell-Lutz with reduction mod p. For $E_{/\mathbf{Z}}$, and for a prime p not dividing Δ, we let $\widetilde{E}_p$ denote the reduction of E mod p as defined in §2.5.2.

Proposition 2.10.4 *Let E be an elliptic curve defined over $\mathbf{Z}$, let $P = (x, y) \in \mathcal{T}^*$, and $\eta = y + (a_1 x + a_3)/2$. Then*

(a) *P is integral except possibly for one point of order 2 of the form $(s/4, t/8)$, where s and t are odd integers. In order that this unique nonintegral torsion point exist, it is necessary that a_1 be odd;*

(b) *$u = 0 \Leftrightarrow [2]P = O$, and otherwise*

$$\begin{array}{ccl}
\text{any } a_1, \text{ any } a_3 & \Longrightarrow & 2\eta \in \mathbf{Z} \text{ and } (2\eta)^2 | 4\Delta \\
a_1 \text{ even, any } a_3 & \Longrightarrow & 2\eta \in \mathbf{Z} \text{ and } (2\eta)^2 | \Delta \\
a_1 \text{ and } a_3 \text{ even}^{\dagger} & \Longrightarrow & \eta \in \mathbf{Z},\ 16 | \Delta, \text{ and } \eta^2 | \frac{1}{16}\Delta
\end{array}$$

(c) *Let $\mathcal{S}$ denote the set of primes p not dividing Δ. Then*

$$|\mathcal{T}| \quad \textit{divides} \quad \gcd_{p\in\mathcal{S}}\{t_p|\widetilde{E}_p(\mathbf{F}_p)|\},$$

where $\quad t_p = \begin{cases} 2 & \textit{if } p=2 \textit{ and } \mathcal{T} \textit{ contains a fractional point of order } 2, \\ 1 & \textit{otherwise.} \end{cases}$

In particular,

- *if a_1 is even and a_3 is odd then E has supersingular reduction at 2 and $\mathcal{T}$ has order 1, 3 or 5;*
- *if a_1 and Δ are odd then E has ordinary reduction at 2 and $|\mathcal{T}|$ divides 8.*

Remark. The finiteness of $\mathcal{T}$, which of course is a corollary of the proposition, also follows from

Mordell's theorem: $E(\mathbf{Q})$ is a finitely generated abelian group.

This is a special case of the Mordell-Weil theorem which will be proved in the next chapter.

Proof. (a) Suppose P is a nonintegral point of order m. Since $p^n = 2$ is the only case where $v(p)/(p^n - p^{n-1}) \geq 1$, by Proposition 2.10.1 we have $m = 2$ and $4x \in \mathbf{Z}$. Thus P is a point of order 2, and the remaining details follow from Proposition 2.5.3(c1).

(b) The statement concerning $\eta = 0$ is just a reminder; *cf.* Proposition 1.7.1. Thus suppose $\eta \neq 0$.

By (a), $x, y \in \mathbf{Z}$ hence $2\eta = 2y + a_1x + a_3 \in \mathbf{Z}$; also $[2]P$ is integral except possibly when a_1 is odd and $[2]P$ has order 2, and then $4x([2]P) \in \mathbf{Z}$. The first two implications now follow from relation ¶ of the previous section. Finally, let a_1 and a_3 be even. Then $4|b_2$, $2|b_4$, $4|b_6$, hence $4|\sigma$ and $16|\tau$. Thus ¶ implies that $4(2\eta)^2|\Delta$.

(c) The first statement follows from Corollary 2.5.4. When a_1 is even and a_3 is odd, then Δ is odd by Proposition 2.5.3(a). Thus reduction mod 2 of the Weierstrass equation produces a supersingular elliptic curve $\widetilde{E}$ over $\mathbf{F}_2$. Also the kernel $\mathcal{T}(E_1(\mathbf{Q}))$ of the reduction of the torsion subgroup is trivial, *i.e.*, $t_2 = 1$, hence the reduction homomorphism embeds $\mathcal{T}$ as a subgroup in $\widetilde{E}(\mathbf{F}_2)$. From the table at the end of Chapter 1, the possible orders of the latter group are 1, 3 and 5. Similarly when a_1 and Δ are odd, $\mathcal{T}$, possibly divided by a subgroup of order 2, embeds in $\widetilde{E}(\mathbf{F}_2)$, and the latter group has order 2 or 4. ■

†This case includes the original Nagell-Lutz theorem.

Examples over Q.

1. The curves **E15** and **F15** (§1.7) and **C17** (§1.7.1) have a point of order 2 of the form $(s/4, t/8)$, and $a_1 = 1$. The curve **A26**: $y^2 + xy + y = x^3$ has $a_1 = 1$ but no points of order 2.

2.

$$y^2 + xy = x^3 - 39x + 90 \qquad \textbf{C21}$$

or $$(2\eta)^2 = 4x^3 + x^2 - 156x + 360 = (4x - 15)(x^2 + 4x - 24).$$

There is one point of order 2

$$(x, \eta) = (15/4, 0) \quad \text{or} \quad (x, y) = (15/4, -15/8),$$

and any other point of finite order satisfies $2\eta \in \mathbf{Z}$ and

$$(2\eta)^2 \,|4\Delta = 2^2 3^8 7.$$

$2\eta = 3$ is the smallest positive value that yields a rational solution, namely $x = 3$. This gives the point $P = (x, y) = (3, 0) \in E(\mathbf{Q})$, which could still be a point of infinite order. As it turns out, P has order 8:

n	$[n](3,0)$	order	$2\eta = 2y + x$
1	$(3, 0)$	8	3
2	$(6, 6)$	4	18
3	$(-3, 15)$	8	27
4	$(15/4, -15/8)$	2	0
5	$(-3, -12)$	8	-27
6	$(6, -12)$	4	-18
7	$(3, -3)$	8	-3

By the last point in the proposition, $|\mathcal{T}| = 8$ and therefore $\mathcal{T} = C_8 = \langle P \rangle$. (Reduction mod 2 is less work than testing the remaining potential 2η.)

Note that the occurrence of $2\eta = \pm 18$ shows that the factor 4 cannot be omitted from 4Δ in the first case in (b) of the proposition.

3.

$$y^2 + y = x^3 + x^2 - 3x + 1 \quad \text{or} \quad (2\eta)^2 = 4x^3 + 4x^2 - 12x + 5. \qquad \textbf{B37}$$

There is no rational root x when $\eta = 0$, *i.e.,* no point of order 2. Thus $\eta \neq 0$, $2\eta \in \mathbf{Z}$ and

$$(2\eta)^2 = (2y + 1)^2 \,|\Delta = 37,$$

hence $2\eta = \pm 1$. Both values yield rational x and we find that the torsion subgroup of $\mathbf{B37}(\mathbf{Q})$ (in x, y coordinates, as usual) is $C_3 = \{O, (1, 0), (1, -1)\}$.

4.

$$y^2 + y = x^3 - x \quad \text{or} \quad (2\eta)^2 = 4x^3 - 4x + 1. \qquad \textbf{A37}$$

As in the previous example, there are no points of order 2, $\Delta = 37$ and the only candidates are $2\eta = \pm 1$. Taking $2\eta = 1$ we obtain the point $P = (0,0)$ and we calculate

n	$[n](0,0)$	$[-n](0,0)$
1	$(0,0)$	$(0,-1)$
2	$(1,0)$	$(1,-1)$
3	$(-1,-1)$	$(-1,0)$
4	$(2,-3)$	$(2,2)$
5	$(1/4,-5/8)$	$(1/4,-3/8)$
6	$(6,14)$	$(6,-15)$

By Proposition 2.10.4, the fifth line of this table shows that $(0,0)$ is a point of infinite order, and therefore the torsion subgroup of $\mathbf{A37}(\mathbf{Q})$ is trivial. In fact $\mathbf{A37}(\mathbf{Q}) = \langle (0,0) \rangle \approx \mathbf{Z}$, as we will see in Corollary 3.7.3, and then it will be an easy matter to see that all the $\mathbf{Z}$-integral points are contained in the above table.

5. Although the gcd in part (c) of the proposition stabilizes after a finite number of p (and in practice, a relatively small number), this final value may still be a proper multiple of $|\mathcal{T}|$. This phenomenon will be illustrated in the present example. The theoretical explanation will be given in Chapter 6; for now we can describe the situation as follows. As a temporary definition, we say that two elliptic curves E and E' defined over $\mathbf{Z}$ are *isogenous* if $|\widetilde{E}(\mathbf{F}_p)| = |\widetilde{E'}(\mathbf{F}_p)|$ for all primes p not dividing the discriminant of either curve. Note that the groups $\widetilde{E}(\mathbf{F}_p)$ and $\widetilde{E'}(\mathbf{F}_p)$ need not be isomorphic. Then the gcd is always a multiple of the *largest* among $|\mathcal{T}(E'(\mathbf{Q}))|$ where E' runs through the (at most 8) curves isogenous to E.

The following three E are isogenous, as will be explained in Chapter 6. We (that is, **apecs**) tabulate $\widetilde{E}(\mathbf{F}_p)$ for a few good p, *i.e.*, $p \neq 11$.

$$y^2 + y = x^3 - x^2, \quad \Delta = -11 \qquad \mathbf{A11},$$

$$y^2 + y = x^3 - x^2 - 10x - 20, \quad \Delta = -11^5 \qquad \mathbf{B11},$$

$$y^2 + y = x^3 - x^2 - 7820 - 263580, \quad \Delta = -11 \qquad \mathbf{C11}.$$

good p	$\mathbf{A11}(\mathbf{F}_p)$	$\mathbf{B11}(\mathbf{F}_p)$	$\mathbf{C11}(\mathbf{F}_p)$
2	C_5	C_5	C_5
3	C_5	C_5	C_5
5	C_5	C_5	C_5
7	C_{10}	C_{10}	C_{10}
31	C_{25}	$C_5 \oplus C_5$	C_{25}
47	$C_2 \oplus C_{20}$	$C_2 \oplus C_{20}$	$C_2 \oplus C_{20}$
53	C_{60}	$C_2 \oplus C_{30}$	$C_2 \oplus C_{30}$
101	$C_5 \oplus C_{20}$	$C_5 \oplus C_{20}$	C_{100}

Nagell-Lutz says to try

$$(2\eta)^2 = (2y+1)^2 \,\big|\, -11 \quad \text{resp.} \quad -11^5 \text{ , } i.e.,$$

$$y = 0, -1, \quad \text{resp.} \quad 0, -1,\ 5, -6,\ 60, -61.$$

There results (abbreviating the curve names)

$$\begin{aligned}
\mathcal{T}(\mathbf{A}(\mathbf{Q})) &= \{O, P = (0,0), [2]P = (1,-1), [3]P = (1,0), [4]P = (0,-1)\} \\
\mathcal{T}(\mathbf{B}(\mathbf{Q})) &= \{O, P = (5,5), [2]P = (16,-61), [3]P = (16,60), [4]P = (5,-6)\} \\
\mathcal{T}(\mathbf{C}(\mathbf{Q})) &= \{O\}
\end{aligned}$$

Thus reduction mod p tells us that $|\mathcal{T}(\mathbf{C}(\mathbf{Q}))|$ is a divisor of 5, but never reveals that the actual value is 1.

In fact in the three cases this is the whole group $E(\mathbf{Q})$. This is proved for **A11** in Corollary 3.7.2, and in Chapter 6 we will see that if E and E' are isogenous over $\mathbf{Q}$ then $E(\mathbf{Q})$ and $E'(\mathbf{Q})$ have the same rank.

When Nagell-Lutz gives a long list of candidates, this problem of "isogeny bloat" can make the determination of $\mathcal{T}$ a lot of work. An alternative work-around is to jettison Nagell-Lutz and carry out these two steps:

1. by reduction mod approriate p, find a multiple N of $|\mathcal{T}|$;
2. find the rational roots of the division polynomials ψ_n for divisors n of N.

(As soon as a new torsion point is found, the group generated by all the points found so far is calculated. Thus only $n = p^m$ that are prime powers need be taken, and of course if ψ_{p^m} has no rational roots then there is no need to look at higher powers of p.)

When applied to the example **C11**, step 1 gives $N = 5$, so $|\mathcal{T}| = 1$ or 5. For step 2 we use the apecs command `div(5)` whose value is ψ_5 (al5 in apecs notation), and the Maple command `roots` which finds the rational roots of a polynomial:

`roots(div(5));` ===> [ ],

the empty list, hence there are no points of order 5, and $\mathcal{T} = O$. *

The 'roots of ψ_n' method of determining $\mathcal{T}$ can be applied to E over other fields. Some examples over quadratic fields are given in the following section.

2.10.2 Nagell-Lutz for quadratic fields

An **algebraic number field**, or simply **number field**, is an extension field of $\mathbf{Q}$ of finite degree. The **degree** of a number field is the degree $[K : \mathbf{Q}]$. Number

*`div(5)` yields $\psi_5 = 5X^{12} + \cdots + 26251755532203500725560$, which would be somewhat daunting to treat by hand. But Maple's `roots` is very fast. As of version 4.37, the apecs procedure `tor` uses the `roots(div(n))` approach. This was brought on by an innocuous looking example whose Δ was highly composite — Nagell-Lutz had over 36,000 candidate divisors, and the old `tor` would have taken an absurdly long time. The present `tor` did the job in a matter of seconds.

fields of degree 2, 3, ... are referred to as **quadratic fields**, **cubic fields**, ... The **ring of integers** in a number field K is the integral closure of $\mathbf{Z}$ in K; this ring is a Dedekind domain.

Let K be a number field of degree d with ring of integers $\mathcal{O}$. We recall some basic number theory: let p be a prime number and let the principal ideal $p\mathcal{O}$ have the prime ideal factorization

$$p\mathcal{O} = \mathcal{P}_1^{e_1} \cdots \mathcal{P}_g^{e_g}.$$

e_i is the **ramification index** of $\mathcal{P}_i$ over p. The field extension degree

$$f_i = [\mathcal{O}/\mathcal{P}_i : \mathbf{Z}/p\mathbf{Z}]$$

is the **inertial degree** of $\mathcal{P}_i$ over p, and we have

$$\sum_{i=1}^{g} e_i f_i = d.$$

The possibilities range from $g = 1$, $f_1 = d$: p is **inert**, to $g = d$ when necessarily all $e_i = 1$ and all $f_i = 1$: p **splits**. If any $e_i > 1$ one says that p is **ramified**, or, more precisely, that $\mathcal{P}_i$ is ramified.

If v denotes the p-adic valuation on $\mathbf{Q}$, then the extensions of v to K are the $\mathcal{P}_i$-adic valuations (see §2.2.1); let us denote them $w_1, \ldots, w_g$. Thus $w_i(p) = e_i$. One refers to the $\mathcal{P}_i$ that occur in the factorization of $p\mathcal{O}$, and the corresponding w_i, as the primes and the valuations that **lie over** p, or are **above** p, and one writes $\mathcal{P}_i|p$, $w_i|v$.

Now let E be an elliptic curve defined over $\mathcal{O}$ and consider the torsion subgroup $\mathcal{T}$ of $E(K)$. Proposition 2.10.1 says that in order that a point $P = (x, y)$ of $\mathcal{T}$ be *fractional, i.e.,* $x \notin \mathcal{O}$, say $w(x) < 0$, the order of P must be a prime power p^n, w must lie over p and p must be sufficiently ramified at w. In the other direction, Proposition 2.10.2 gives "upper bounds" on P.

Let us write the details for quadratic fields in a corollary.

Corollary 2.10.5 *Let K be a quadratic field with ring of integers $\mathcal{O}$, and let E be an elliptic curve defined over $\mathcal{O}$. The possible orders for a fractional torsion point $P = (x, y)$ are*

(a) 2*;*

(b) 3 *if* 3 *is ramified;*

(c) 4 *if* 2 *is ramified.*

Hence if $E \supset E_1 \supset \cdots$ is the filtration for a valuation of residue characteristic ≥ 5, then the torsion subgroup of $E_1(K)$ is trivial.

In case (a) *there is at most one such P with $w(x) < 0$ for each valuation w over* 2*, and in cases* (b) *and* (c) *there is at most one such pair P, $-P$. Further*

details in these cases are as follows, where w denotes a valuation on K and, $\kappa = 2\eta = 2y + a_1x + a_3$.

(a) *If $w|2$ then $w(x) = -2$ if 2 is unramified, and $w(x) = -2$ or -4 if 2 is ramified. For all w not above 2, $w(x) \geq 0$.*

(b) *If w is the unique valuation above 3, then $w(x) = -2$ and $w(\kappa) \leq w(\Delta)/2 + 1$. If $w \nmid 3$, then $w(x) \geq 0$ and $w(\kappa) \leq w(\Delta)/2$.*

(c) *If w is the unique valuation above 2, then $w(x) = -2$, $w(x([2]P)) = -4$, and $w(\kappa) \leq w(\Delta)/2 + 1$. If $w \nmid 2$, then $w(x) \geq 0$ and $w(\kappa) \leq w(\Delta)/2$.*

If P is an integral torsion point of order > 2 then for all w

$$w(\kappa) \leq w(\Delta)/2 - \min\{0, w(x([2]P))\}.$$

Proof. The possibilities for the factorization of $p\mathcal{O}$ are $\mathcal{P}_1$, $\mathcal{P}_1\mathcal{P}_2$ and $\mathcal{P}_1^2$, and the three cases of fractional $P = (x, y) \in \mathcal{T}$ follow from $w(x) < 0 \Longrightarrow w(x) \geq -2\lfloor w(p)/(p^n - p^{n-1})\rfloor$.

There can be at most one such P in case (a) with $w(x) < 0$ by Proposition 2.5.3.

In case (b), two such pairs $\pm P$ and $\pm P'$ would give rise to two more $\pm(P+P')$ and $\pm(P - P')$. Each pair is determined by an x-coordinate with $w(x) = -2$ which is a root of $\psi_3 = 3X^4 + \cdots + b_8$. The product of these x's is $b_8/3$, and we have a contradiction from $w(b_8/3) \geq -2$.

Similarly in case (c), the x-coordinate is a root of $\psi_4/\kappa = 2X^6 + \cdots$.

The remaining statements are straightforward interpretations of Proposition 2.10.2. ■

We now consider a few examples.

Example 1 Let us determine the group of torsion points of

$$E : y^2 + xy + y = x^3 + 4x - 6 \qquad \textbf{C14}$$

over the field $K = \mathbf{Q}(\sqrt{-3})$. The roots of

$$\psi_2^2 = \kappa^2 = 4x^3 + x^2 + 18x - 23$$

are
$$1, \quad -\frac{5}{8} \pm \frac{7}{8}\sqrt{-7} = \frac{1+7\omega}{4}, \ -\frac{6+7\omega}{4},$$

where ω is the algebraic integer $(-1 + \sqrt{-7})/2$. Calculation shows that the points of order 2 are

$$(1, -1), \ \left(\frac{1+7\omega}{4}, -\frac{5+7\omega}{8}\right), \ \left(-\frac{6+7\omega}{4}, \frac{2+7\omega}{8}\right),$$

of which only the first is in $E(K)$. With $\eta = y + (x+1)/2$ and $x_1 = x - 1$, the Weierstrass equation becomes

$$\eta^2 = x_1(x_1^2 + (13/4)x_1 + 8).$$

Since $\sqrt{8} \notin K^*$, by Proposition 1.7.3(c) there are no points of order 4 in $E(K)$.

The roots of

$$\psi_3 = 3x^4 + x^3 + 27x^2 - 69x - 26$$

are

$$2, \quad -1/3, \quad -1 \pm 2\sqrt{-3},$$

and one finds that all 8 points of order 3 are rational over K:

$$(2,\, 2), \quad (2,\, -5), \quad \left(-\frac{1}{3},\, -\frac{1}{3} \pm \frac{14}{9}\sqrt{-3}\right),$$

$$(-1 \pm 2\sqrt{-3},\, -5), \quad (-1 \pm 2\sqrt{-3},\, 5 \mp 2\sqrt{-3}).$$

Finally, there are no other points of finite order in $E(K)$, and so the torsion group has order 18; for otherwise there would be an integral point P of order p, for some prime $p \geq 5$, and $[2]P$ would also be integral. The verification can then be completed by checking all possible $\kappa \in \mathcal{O}$ satisfying

$$\kappa^2 \,\big|\, \Delta = -2^6 7^3 = -2^6(3+\zeta)^3(2-\zeta)^3,$$

where $\zeta = (-1+\sqrt{-3})/2$, the last product being the prime factorization in the PID $\mathcal{O}$. Thus

$$\kappa = (1+\zeta)^\alpha 2^\beta (3+\zeta)^\gamma (2-\zeta)^\delta \quad \text{with} \;\; \alpha \bmod 6,\, \beta \leq 3,\, \gamma \leq 1,\, \delta \leq 1.$$

The number of candidate values of κ^2 is 48, already a substantial number for one of the simplest examples. However, reduction mod appropriate π relieves us of this burden. Noting that $13 = (4+3\zeta)(1-3\zeta)$ is split and is prime to Δ, reduction of $E \bmod \pi = 4 + 3\zeta$ gives the upper bound 18.

Example 2 We consider the points of order 2 of example 1 over the field $\mathbf{Q}(\sqrt{-7})$. Again $\mathcal{O} = \mathbf{Z}[\omega]$, where $\omega = (-1+\sqrt{-7})/2$, is a PID, and this time $\mathcal{O}^*$ is simply ± 1. We find $2 = -\omega(1+\omega)$, and using the norm $\mathrm{N}(a+b\omega) = a^2 - ab + 2b^2$ as an aid, we find that $23 = (5+\omega)(3-2\omega)$. Thus the factorizations of x for the fractional points (x, y) of order 2 are

$$\frac{1+7\omega}{4} = \frac{3-2\omega}{\omega^2}, \quad -\frac{6+7\omega}{4} = \frac{5+2\omega}{(1+\omega)^2},$$

which are in agreement with the above corollary So we can indeed have two fractional points of order 2 when 2 splits.

Example 3 The points of order 2 on

$$y^2 + xy + y = x^3 + x^2 - 10x - 10 \qquad \textbf{C15}$$

are

$$(-1, 0), \quad (3, -2), \quad (-13/4, 9/8),$$

and from $\psi_4/\kappa = 2x^6 + 5x^5 - 95x^4 - 390x^3 - 1390x^2 - 1436x + 1120$

we find that the points of order 4 are

$$(-2,3),\ (-2,-2),\ (8,18),\ (8,-27),\ (-7,3\pm 15i),$$

$$\left(\frac{1}{2}, \frac{-3\pm 15i}{4}\right),\ (-1+3i, 6\pm 6i),\ (-1-3i, 6\pm 6i),$$

where $i^2 = -1$. Note that $2 = i(1+i)^2$ is ramified in $\mathbf{Q}(i)$.

So far, the torsion points listed give a group of type $C_4 \oplus C_4$ in $E(\mathbf{Q}(i))$; in fact this is clearly the whole $\mathcal{T}(E(\mathbf{Q}(i)))$ by reduction mod $\pi = 1+i$ to $\widetilde{E}(\mathbf{F}_2)$. We have

$$[2]\left(\frac{1}{2}, \frac{-3\pm 15i}{4}\right) = \left(\frac{-13}{4}, \frac{9}{8}\right).$$

This gives an example of type (c) of the corollary.

Applying Proposition 1.7.3 to $(8,-27)$, we find that

$$(3\epsilon + 2, 3), \quad \text{where} \quad \epsilon = (-1+\sqrt{5})/2,$$

is a point of order 8, and thus we have a subgroup of type $C_8 \oplus C_2$ in $E(\mathbf{Q}(\sqrt{5}))$. Since there are infinitely many units $\pm\epsilon^n$ in $\mathcal{O} = \mathbf{Z}[\epsilon]$, Nagell-Lutz gives infinitely many candidates, and now reduction mod π is a practical necessity. In fact $11 = (-3+\epsilon)(4+\epsilon)$ splits in $\mathbf{Q}(\sqrt{5})$, and $|E(\mathbf{F}_{11})| = 16$, so we have all of $\mathcal{T}$.

Example 4 Here is an example of isogeny bloat, as discussed in the previous section. Over $\mathbf{Q}(\sqrt{2})$, **A14**: $y^2 + xy + y = x^3 - x$ has $|\mathcal{T}| = 6$ (with all points actually defined $/\mathbf{Q}$), but reduction modπ always gives a multiple of 12. This is because the isogenous curve **B14**: $y^2 + xy + y = x^3 - 11x + 12$ over $\mathbf{Q}(\sqrt{2})$ has $|\mathcal{T}| = 12$. We leave the detailed verification to the reader, who may prefer simply to call on apecs (version ≥ 5.1) and type

```
Qfin(2);ell(1,0,1,-1,0);Torq();
```

There is another method to determine $\mathcal{T}$:

$$\mathcal{T}_{E/K} = \{P \in E(K) : \hat{h}(P) = 0\}$$

where $\hat{h}$ is the *canonical height*, to be defined in the next chapter. This applies to a class of K that includes all number fields. But for number fields this method is not practicable, and is only of theoretical interest.

Chapter 3

The Mordell-Weil theorem

The Mordell-Weil theorem for elliptic curves is this:

> *If K is a finitely generated field and E is an elliptic curve defined over K, then the group $E(K)$ is finitely generated.*

By *finitely generated field* we mean finitely generated over the prime subfield. Then, as an abstract abelian group, $E(K)$ has the form

$$\mathcal{T} \oplus \mathbf{Z}^r$$

where $\mathcal{T}$ is finite and $\mathbf{Z}$ denotes the infinite cyclic group. The non-negative integer r is called the **rank** of E over K.

Mordell proved the theorem for $K = \mathbf{Q}$ in [Mor22].

In his thesis [Wei29], Weil generalized Mordell's theorem in two ways: he proved that if K is a number field and A is an abelian variety defined over K, then the group $A(K)$ is finitely generated. (In this introduction to Chapter 3, we use some technical terms that will be defined properly only later. Abelian varieties are projective group varieties; elliptic curves are abelian varieties of dimension 1. Actually, Weil proved his result only for A that are Jacobians of curves — it was subsequently discovered that not every abelian variety is the Jacobian of a curve. But this would appear to be a minor point since every abelian variety is isogenous to a factor of a Jacobian.)

Extending Weil's result, Néron in his thesis [Nér52] proved that $A(K)$ is finitely generated if K is finitely generated. However, it is customary to call any of these theorems the Mordell-Weil theorem. A proof of Néron's generalization is given in [Lan83, ch.6].

The proof is split into two parts: the *weak* Mordell-Weil theorem, which is the statement that the quotient group $E(K)/[m]E(K)$ is finite for some integer $m \geq 2$, and the construction of a height function $h : E(K) \longrightarrow \mathbf{R}$ with appropriate properties.

In this chapter we will prove the Mordell-Weil theorem for elliptic curves defined over the following types of fields:

- number fields, and
- finitely generated fields of characteristic > 2.

Thus the fields that are missing from our treatment of the theorem are finitely generated fields of characteristic 0 or 2 that are transcendental over the prime subfield. The reasons are:

- our proof of weak Mordell-Weil (with $m = 2$) requires $\operatorname{char} K \neq 2$;
- our height function (in the form implemented) lacks a key finiteness property when $\operatorname{char} K = 0$ and the transcendence degree ≥ 1.

The proof will be essentially elementary, quoting results as needed from standard commutative algebra and number theory. The first section in this chapter contains a definition and some preliminary observations that will be used in the weak Mordell-Weil theorem.

After the proof we construct the canonical height and show some basic properties. Then we discuss ways of estimating r, and consider various examples.

There is also a *relative* theory due to Lang and Néron, which in the case of elliptic curves amounts to the following two statements; *cf.* [Lan83, p.139]. Let K be a finitely generated separable extension of the field K_1 such that K_1 is algebraically closed in K, and let E be an elliptic curve defined over K with invariant j.

(i) If E is defined over K_1, then the quotient group $E(K)/E(K_1)$ is finitely generated. Hence if $E(K_1)$ is finitely generated, then so is $E(K)$.

(ii) If $j \notin K_1$ then the group $E(K)$ is finitely generated. For example, if t is an indeterminate and E is defined over $\mathbf{C}(t)$ with $j \notin \mathbf{C}$, then $E(\mathbf{C}(t))$ is finitely generated. This is true in particular if E is defined over $\mathbf{Q}(t)$ with $j \notin \mathbf{Q}$, and then there exists a number field F such that

$$E(\mathbf{C}(t)) = E(F(t)).$$

(Assuming $G := E(\mathbf{C}(t))$ is finitely generated, the existence of F is easily explained: If $P = (x, y) \in G$ then $[\mathbf{Q}(x, y) : \mathbf{Q}] < \infty$; otherwise x, y would contain a transcendental, say α and since E is defined over $\mathbf{Q}$, α could be replaced by any other transcendental, giving uncountably many points. Letting P run through a finite set of generators, together their coordinates generate a finite extension F of $\mathbf{Q}$, and every point in G is defined over F. We note that F is normal over $\mathbf{Q}$, since $(x, y) \in G$ implies that every conjugate $(x^\sigma, y^\sigma) \in G$.)

3.1 F2-Krull domains

Let R be a Krull domain. The free abelian group on the essential valuations of R modulo principal divisors is the **class group**, and is denoted $\mathrm{Cl}(R)$.

We call R an **F2-Krull domain** if it satisfies the following two finiteness conditions, where C_2 denotes the cyclic group of order 2, and $\otimes$ and hom refer to $\mathbf{Z}$-modules:

$$R^*/R^{*2} \approx C_2 \otimes R^* \quad \text{is finite,}$$

and

$$\mathrm{Cl}(R)[2] \approx \hom(C_2, \mathrm{Cl}(R)) \quad \text{is finite.}$$

It is precisely these two finiteness properties that support the proof of weak Mordell-Weil, as we will see. We make two elementary observations:

Lemma 3.1.1 *Suppose R is an* F2*-Krull domain. Then so are*
- *the polynomial ring $R[\{t_i\}]$ for an arbitrary set of indeterminates, and*
- *$R[1/d_1, \ldots, 1/d_m]$ for nonzero elements $d_1, \ldots, d_m \in R$.*

Proof. Let $R_t = R[\{t_i\}]$. Then R_t is Krull by [BAC7, Prop.13], extended to the case of infinitely many t_i by Exercise 8 for §1. Trivially, $R_t^* = R^*$. Also, $\mathrm{Cl}(R_t) = \mathrm{Cl}(R)$ by [BAC7, Prop.18] and the fact that any divisor involves only finitely many t_i. Thus the F2-properties are preserved in polynomial extensions.

Secondly, the set of essential valuations v of R for which $v(d_i) > 0$, for some i, is finite by the definition of Krull domain; let $\mathcal{D}$ denote the free abelian subgroup of the divisor group generated by these v. By [BAC7, Prop.6], $R' := R[1/d_1, \ldots, 1/d_m]$ is Krull, and from [Lan83, p.41] we have the (elementary) exact sequences

$$0 \longrightarrow R^* \longrightarrow R'^* \longrightarrow \mathcal{D},$$

$$\mathcal{D} \longrightarrow \mathrm{Cl}(R) \longrightarrow \mathrm{Cl}(R') \longrightarrow 0.$$

Letting $\mathcal{D}_1$ and $\mathcal{D}_2$ denote an appropriate subgroup and factor group respectively of $\mathcal{D}$, these yield the short exact sequences

$$0 \longrightarrow R^* \longrightarrow R'^* \longrightarrow \mathcal{D}_1 \longrightarrow 0, \tag{1}$$

$$0 \longrightarrow \mathcal{D}_2 \longrightarrow \mathrm{Cl}(R) \longrightarrow \mathrm{Cl}(R') \longrightarrow 0. \tag{2}$$

Tensoring (1) with C_2 gives the exact sequence

$$C_2 \otimes R^* \longrightarrow C_2 \otimes R'^* \longrightarrow C_2 \otimes \mathcal{D}_1.$$

The group on the left is finite by assumption, and the one on the right is finite since $\mathcal{D}_1$ is finitely generated. Hence the group in the middle is finite.

Applying $\hom(C_2, —)$ to (2) gives

$$\hom(C_2, \mathrm{Cl}(R)) \longrightarrow \hom(C_2, \mathrm{Cl}(R')) \longrightarrow \mathrm{ext}^1(C_2, \mathcal{D}_2).$$

Since $\mathrm{ext}^1(C_2, \mathbf{Z}) = C_2$, and $\mathcal{D}_2$ is finitely generated, therefore $\mathrm{ext}^1(C_2, \mathcal{D}_2)$ is finite. Again the middle term is sandwiched between two finite groups and is therefore finite. ■

Example 1 The ring of integers R in a number field is F2-Dedekind. In fact the group R^* is finitely generated and $\mathrm{Cl}(R)$ is finite. These basic facts can be found in any text on algebraic number theory.

Example 2 Only certain texts on algebraic number theory, such as [Wei67], contain the function field analog of example 1: let R be the integral closure of the one variable polynomial ring $\mathbf{F}_q[t]$ in a finite separable extension of $\mathbf{F}_q(t)$. Then again R^* is finitely generated and $\mathrm{Cl}(R)$ is finite.

Example 3 A field K is F2-Krull when it is algebraically closed or real closed since then $\mathrm{Cl}(K) = 0$ and $K^*/K^{*2} = 1$ or C_2 respectively. Thus the polynomial rings $K[t_1, \ldots, t_n]$, where $K = \mathbf{C}$ or $\mathbf{R}$ are F2-Krull.

Example 4 Examples 1 and 2 generalize as follows.

We define a ring to be **FT** (for finite type) if it is finitely generated as a ring over the prime subring $\overline{\mathbf{Z}}$. Note the minor conflict of terminology: the finitely generated fields $\mathbf{Q}$ and $\mathbf{F}_p(t)$ are not FT.

Let the field K be finitely generated over the prime subfield K_0; let $\mathrm{char}\, K = p \geq 0$; let R_1 denote the integral closure of $\overline{\mathbf{Z}}$ in K; and let K_1 denote the algebraic closure of K_0 in K. Thus $K_1 = Q(R_1)$, where Q denotes quotient field.

When $p = 0$, K_1 is a number field with R_1 as ring of integers; when $p > 0$, then $R_1 = K_1 = \mathbf{F}_q$ for some power q of p. In any case, R_1 is FT.

Choose a separating transcendence basis $t_1, \ldots, t_n$ of K over K_1 (*cf.* [Zar-Sa58, p.105, theorem 31]); let $R_2 = R_1[t_1, \ldots, t_n]$ and $K_2 = Q(R_2) = K_1(\{t_i\})$; thus K is a finite separable extension of K_2, and R_2 is FT. R_3 denotes the integral closure of R_2 in K; thus $Q(R_3) = K$. By [Zar-Sa58, p.264], R_3 is a finite R_2-module, hence is FT. By Propositions 12 and 13 of [BAC7], all the R_i are Krull; thus we say that R_1, R_2 and R_3 are FT-Krull.

We have the following results of Roquette [Roq58]; see also [Lan83, p.37].

R1 Let R be an FT domain. Then the group of units R^* is finitely generated.

R2 If R is FT-Krull, then the divisor class group $\mathrm{Cl}(R)$ is finitely generated.

Consequently, an FT-Krull domain is F2.

Although these theorems can hardly be described as standard number theory, we will assume them so that weak Mordell-Weil applies to all finitely generated fields (of characteristic $\neq 2$), as will be explained in the next section. Of course in the case of number fields we only need to quote the standard theorems.

If the conclusion of the lemma is true for the integral closure of R in a finite separable extension of the quotient field, and if we we could obtain a reasonably simple proof of this, then we would have a self contained proof of weak Mordell-Weil, *sans* Roquette. Lang remarks [Lan83, p.43] that **R2** is a comparatively deep theorem.

3.2 The weak Mordell-Weil theorem

Let K be a field of characteristic $\neq 2$, but, for the moment, otherwise arbitrary, and let E be an elliptic curve defined over K. We take the equation in b-form

$$\eta^2 = x^3 + (b_2/4)x^2 + (b_4/2)x + b_6/4 = f(x),$$

and coordinates of points are (x, η)-coordinates.

Let the cubic factor over $\overline{K}$ as

$$\eta^2 = (x - e_1)(x - e_2)(x - e_3). \tag{1}$$

The number of e_i in K can be 0, 1 or 3. Thus the 2-division points are $(e_i, 0)$, $i = 1, 2, 3$, and the polynomial discriminant D of f is related to the elliptic curve discriminant Δ by

$$D = (e_1 - e_2)^2(e_1 - e_3)^2(e_2 - e_3)^2 = \Delta/16.$$

We define the field $L_i = K(e_i)$, the factor group $\Gamma_i = L_i^*/L_i^{*2}$ and a map $\phi_i : E(K) \longrightarrow \Gamma_i$ as follows. First, $\phi_i(O) = 1L_i^{*2}$; second, if $P \in E(K)$, $P \neq O$ and $x(P) \neq e_i$, then

$$\phi_i(P) = (x(P) - e_i)L_i^{*2}.$$

We notice that when $x(P) \neq e_{i+1}$ or e_{i+2} (taking subscripts mod 3), equation (1) implies, since Γ_i is "mod squares",

$$\phi_i(P) = (x(P) - e_{i+1})(x(P) - e_{i+2})L_i^{*2}.$$

Third, we take this as our guide to complete the definition:

$$\phi_i(e_i, 0) = (e_i - e_{i+1})(e_i - e_{i+2})L_i^{*2}.$$

Define $\Phi : E(\mathbf{Q}) \longrightarrow \Gamma_1 \times \Gamma_2 \times \Gamma_3$ by

$$\Phi(P) = (\phi_1(P), \phi_2(P), \phi_3(P)).$$

Proposition 3.2.1 (The weak Mordell-Weil theorem)
Let K be a field of characteristic $\neq 2$. Then, with the above notation,

(a) *ϕ_i is a group homomorphism, hence so is Φ;*

(b) $\ker \Phi = [2]E(K)$*; if $E(K)[2] = O$ then $\forall i$,* $\ker \phi_i = [2]E(K)$.

Now suppose K is the quotient field of an F2*-Krull domain R,* e.g. *when K is finitely generated, and* $\operatorname{char} K \neq 2$. *Then*

(c) $\operatorname{im} \Phi$ *is finite.*

Hence the group $E(K)/[2]E(K)$ is finite.

Proof. (a) We must show that for $P_1, P_2 \in E(K)$, $\phi_i(P_1 + P_2) = \phi_i(P_1)\phi_i(P_2)$.

- $P_1 = O$ or $P_2 = O$: clear;
- $P_1 = -P_2 \neq O$: clear from $x(P_1) = x(P_2)$;
- for $P \notin E(K)[2]$, $\phi_i([2]P) \equiv 1$: this follows from Corollary 1.7.4 with x replaced by $x - e_i$;
- the general case (none of the above): let $P_1 + P_2 = -P_3$, let $(x, \eta)(P_j) = (x_j, \eta_j)$, and let $\eta = mx + c$ be the line through (x_1, η_1) and (x_2, η_2). Then the equation
$$(mx + c)^2 = (x - e_1)(x - e_2)(x - e_3)$$
has the roots $x = x_1, x_2, x_3$, hence
$$(x - e_1)(x - e_2)(x - e_3) - (mx + c)^2 = (x - x_1)(x - x_2)(x - x_3).$$
Substituting $x \mapsto e_i$ shows that $\phi_i(P_1)\phi_i(P_2)\phi_i(P_3)$ is a square.

(b) is an immediate from Proposition 1.7.3 and the remark that when $E(K)[2] = O$, L_i are conjugate cubic extensions of K.

(c) For notational convenience, let us prove that $\operatorname{im}\phi_1$ is finite. Over L_1 write the Weierstrass equation as
$$\eta^2 = (x - e_1)(x^2 + Ax + B) = (x - e_1)F.$$
Choose $d_1, \ldots, d_m \in R$ so that $e_1, A, B, 1/D \in R' := R[1/d_1, \ldots, 1/d_m]$. Since D is a polynomial in e_1, A, and B, therefore D is in the unit group R'^*. By Lemma 3.1.1, R' is F2-Krull.

Let $\epsilon_1, \ldots, \epsilon_t$ be elements of R'^* that represent the classes of R'^*/R'^{*2}; let $I_1, \ldots, I_h$ be divisors representing $\mathrm{Cl}(R')[2]$; and fix generators of the principal divisors $2I_i = (m_i)$.

Let $P = (x, \eta)$ be a nonzero element in $E(K)$ with $x \neq e_1$. We claim that for every essential valuation w of R', $w(x - e_1)$ is even. This is true if $w(x - e_1) < 0$, equivalently $w(x) < 0$ since $e_1 \in R'$, as we saw in §2.1.2. Thus suppose $w(x - e_1) > 0$, and let w' be any extension of w to $K(e_1, e_2, e_3)$. We have $w'(x - e_2) = 0$ and $w'(x - e_3) = 0$ since otherwise one of $e_1 - e_2 = (x - e_2) - (x - e_1)$ or $e_1 - e_3$ would have positive w'-value, hence $w'(D) > 0$ and therefore $w(D) > 0$, contradicting the fact that $D \in R'^*$. Thus $w'(F) = w'((x - e_2)(x - e_3)) = 0$, hence $w(F) = 0$ and therefore $w(x - e_1) = 2w(\eta) - w(F)$ is even.

It follows that the principal divisor $(x - e_1)$ can be written as $2N$, for some divisor $N \in \mathrm{Cl}(R')[2]$. Let N be in the class of I_j, say $N = I_j + (z)$. This implies the equality of divisors $(x - e_1) = (m_j) + 2(z) = (m_j z^2)$, hence $x - e_1 = u m_j z^2$ for some $u \in R'^*$. Writing $u = \epsilon_1^{\alpha_1} \cdots \epsilon_t^{\alpha_t} u_1^2$ where $\alpha_i \in \{0, 1\}$, we have
$$x - e_1 = \epsilon_1^{\alpha_1} \cdots \epsilon_t^{\alpha_t} m_j (z u_1)^2.$$
Allowing for $P = O$ and the possibility $x = e_1$, we have proved that $|\operatorname{im}\phi_1| \leq 2 + 2^t h$. ■

There is an alternative to the Roquette approach to weak Mordell-Weil: that of Lang-Tate (*cf.* [Lan83, p.156]), and that is the approach taken in [Lan83] — of course for abelian varieties of any dimension and allowing characteristic 2; it is also the approach in [Sil86] — but there only for number fields. In the end, both approaches use the two finiteness theorems, so the Lang-Tate approach seems a little more round-about, at least in the elliptic curve case when $\operatorname{char} K \neq 2$.

3.3 Heights

The additive group $\mathbf{Q}^+$ is countable and has the weak Mordell-Weil property

$$\mathbf{Q}^+/[2]\mathbf{Q}^+ = 0,$$

yet is not finitely generated. The extra ingredient that allows us to prove that $E(K)$ is finitely generated for appropriate K is the existence of a height function on the group $E(K)$, in the sense of the following definition.

A **height function** on an abelian group A is a map

$$h : A \longrightarrow \mathbf{R}^{\geq} \quad \text{(the set of real numbers} \geq 0)$$

satisfying

H1 For every $\alpha \in \mathbf{R}^{\geq}$, there are only finitely many $P \in A$ with $h(P) < \alpha$.
H2 For every $Q \in A$ there exists $\beta(Q) \in \mathbf{R}$ such that

$$h([2]P - Q) > 2h(P) - \beta(Q), \quad \forall P \in A.$$

Here we are following Cassels [Cas66, p.258]. The number $h(P)$ is called the **height** of P.

Proposition 3.3.1 *Let A be an abelian group for which $A/[2]A$ is finite and there exists a height function h. Then A is finitely generated.*

Proof. Let $Q_1, \ldots, Q_n$ be representatives of the cosets of $[2]A$ in A, and let B denote the subgroup generated by the Q_i and the finitely many $P \in A$ for which $h(P) < \alpha := \max\{\beta(Q_i) : 1 \leq i \leq n\}$. The conclusion will follow from $B = A$. Suppose, however, there exist $P \in A - B$; by **H1** we can choose such a P with minimal height. Let $P = [2]P_1 - Q_i$. Then $P_1 \notin B$ and therefore

$$\begin{aligned} h(P_1) \geq h(P) &= h([2]P_1 - Q_i) \\ &> 2h(P_1) - \beta(Q_i) \geq 2h(P_1) - \alpha. \end{aligned}$$

This implies $h(P_1) < \alpha$, hence $P_1 \in B$, a contradiction. ■

Actually for our applications to elliptic curves, we will construct an h defined for all $x \in K$, and then we define

$$h(O) = 0, \quad \text{and for} \;\; P = (x, y) \in E(K),\; P \neq O, \quad h(P) = h(x).$$

For the cases of Mordell-Weil proved in this chapter, the h constructed satisfies

H1a For every $\alpha \in \mathbf{R}^{\geq}$, there are only finitely many $x \in K$ with $h(x) < \alpha$.

Thus **H1a** is a property of the field K. Clearly

$$\textbf{H1a} \implies \textbf{H1}$$

for the group $A = E(K)$, for every elliptic curve E defined over K.

Then we will prove that for any given elliptic curve E defined over K, h satisfies the following two conditions, where P, Q denote elements of $E(K)$.

H2a $$h([2]P) = 4h(P) + O(1), \quad \forall P,$$

where $O(1)$ denotes a function of P that is bounded above and below by constants that are independent of P (but of course may depend on E);

H2b for every Q there exists $\gamma(Q) \in \mathbf{R}$ such that

$$h(P + Q) < 2h(P) + \gamma(Q), \quad \forall P.$$

Notice that **H2a** and **H2b** imply **H2**: by replacing P by $[2]P - Q$ in **H2b** and then using half of **H2a**, we obtain

$$2h([2]P - Q) > h([2]P) - \gamma(Q) > 4h(P) - \gamma(Q) - C.$$

In the next subsection we construct h on number fields, and in the following subsection on function fields, which includes all remaining types of finitely generated fields. Then we prove that h satisfies **H2a** and **H2b** in all cases, and **H1a** in the cases of number fields and finitely generated fields of of positive characteristic. An explicit example of an $E(K)$ satisfying **H1** but not **H1a** will be given in §3.4.1.

3.3.1 Heights in number fields

Let K be a number field and v a valuation on K. If v lies over the p-adic valuation on $\mathbf{Q}$, then (see §2.10.2) we have the ramification index $e_v = v(p)$ and the inertial degree $f_v = [R_1/\mathcal{P}_v : \mathbf{F}_p]$, where R_1 is the ring of integers in K and $\mathcal{P}_v = \{a \in R_1 : v(a) > 0\}$. Thus $R_1/\mathcal{P}_v = \mathbf{F}_{p^{f_v}}$. We define the v-adic **absolute value** by

$$|0|_v = 0, \quad \text{and for} \;\; x \in K^*, \quad |x|_v = |R_1/\mathcal{P}_v|^{-v(x)} = p^{-f_v v(x)}.$$

This absolute value satisfies $|xy|_v = |x|_v|y|_v$ and the ultrametric triangle inequality

$$|x+y|_v \le \max\{|x|_v,\, |y|_v\},$$

which of course implies the ordinary, or archimedean triangle inequality

$$|x+y|_v \le |x|_v + |y|_x.$$

Another fact that follows from the definition is that if $x \in \mathbf{Q}$ and $|x|_p$ denotes the p-adic absolute value, *i.e.,* the definition applied to the case $K = \mathbf{Q}$, then

$$\text{for} \quad x \in \mathbf{Q}, \quad |x|_v = |x|_p^{e_v f_v}. \tag{¶}$$

There is also a relative version of (¶). The multiplicativity of e and f in towers implies that if L is a number field containing K and w is an extension of v to L, then, using the notation explained in §2.1.3,

$$e_w = e(w,v)e_v, \qquad f_w = f(w,v)f_v.$$

Consequently,

$$\text{for} \quad x \in K, \quad |x|_w = |x|_v^{e(w,v)f(w,v)}. \tag{¶$'$}$$

Each of the $[K:\mathbf{Q}]$ embeddings $K \hookrightarrow \mathbf{C}$ gives rise, by restricting the usual absolute value, to an archimedean absolute value on K. It is convenient to denote the embeddings by $v : K \hookrightarrow \mathbf{C}$, and call v an **archimedean valuation**, though of course it is not a valuation; the archimedean absolute value is defined by $|x|_v = |v(x)|$. The set of these $[K:\mathbf{Q}]$ archimedean absolute values is denoted $\mathcal{M}_\infty(K)$; for each prime number p, we let $\mathcal{M}_p(K)$ denote the set of valuations on K above p; and we define $\mathcal{M}(K)$ to be the union of all these sets:

$$\mathcal{M}(K) = \bigcup_{p \le \infty} \mathcal{M}_p(K).$$

For $v \in \mathcal{M}_\infty(K)$ we define $e_v = f_v = 1$, and if w extends v to a finite extension L of K, which is indicated notationally by $w|v$, we define[†] $e(w,v) = f(w,v) = 1$. Thus (¶) and (¶$'$) are true for all $v \in \mathcal{M}(K)$. Since each $v \in \mathcal{M}_\infty(K)$ has $[L:K]$ distinct extensions w, therefore

$$\sum_{w|v} e(w,v)f(w,v) = [L:K] \tag{♭}$$

is also true for all $v \in \mathcal{M}(K)$.

[†]This is non-standard: when a real embedding $v : K \hookrightarrow \mathbf{R}$ extends to a pair of conjugate complex embeddings $w, \overline{w} : L \hookrightarrow \mathbf{C}$, then $|x|_w = |\overline{x}|_w = |x|_{\overline{w}}$, and the usual procedure is to count only one w, put $e(w,v) = 2$, and say that v ramifies in L. According to our definition, non-real archimedean absolute values occur in pairs of identical twins.

A basic theorem of number theory is the **product formula**:

$$\text{for } x \in K^*, \qquad \prod_{v \in \mathcal{M}(K)} |x|_v = 1.$$

For example for $\mathbf{Q}(i)$, since the factorizations of 2, 3, 5 are $i(1+i)^2$, 3, $(1+2i)(1-2i)$ respectively, and there is one conjugate pair of embeddings $\mathbf{Q}(i) \hookrightarrow \mathbf{C}$,

$$\prod \left| \frac{1+i}{15} \right|_v = 2^{-1} 3^2 5^1 5^1 \sqrt{\frac{2}{15^2}} \sqrt{\frac{2}{15^2}} = 1.$$

We define the **Weil height** on K:

$$\text{for } x \in K, \quad h(x) = \sum_{v \in \mathcal{M}(K)} \log \max\{|x|_v, 1\}, \quad \text{thus } h(0) = 0,$$

and then for an elliptic curve $E_{/K}$ we define

$$h(O) = 0, \quad \text{and for } P \neq O, \quad h(P) = h(x(P)).$$

It is clear that $h(P) \geq 0$.

If $x = r/s$ where $r, s \in K$, $s \neq 0$, then adding $\sum_{v \in \mathcal{M}(K)} \log |s|_v$, which is 0 by the product formula, allows $h(x)$ to be written as

$$h(r/s) = \sum_{v \in \mathcal{M}(K)} \log \max\{|r|_v, |s|_v\}.$$

In the case $K = \mathbf{Q}$, the formula simplifies dramatically: if r and s are coprime positive integers, then

$$h(\pm r/s) = \log \max\{r, s\}.$$

This makes the **H1a** property obvious for $\mathbf{Q}$. In this form, h is often called the *naive height* on $\mathbf{Q}$; the exponentiated form $H(\pm r/s) = \max\{r, s\}$ was what Mordell used in his original proof [Mor22].

We write h_K for h when it is necessary to identify the field.

Proposition 3.3.2 *Let K be a number field.*

(a) *For all $x, y \in K$,*

$$h(xy) \leq h(x) + h(y),$$

and

$$h(x+y) \leq h(x) + h(y) + \nu \log 2$$

where ν is the number of archimedean $v \in \mathcal{M}(K)$.

(b) *Let L be a finite extension of K. Then*

$$\text{for } x \in K, \quad h_L(x) = [L:K] h_K(x).$$

Proof. (a) As a temporary notation, put $H_v(x) = \max\{|x|_v, 1\}$, hence $h(x) = \log \prod_v H_v(x)$. Since $|xy|_v = |x|_v|y|_v$, therefore $H_v(xy) \le H_v(x)H_v(y)$. Adding over v and taking the logarithm gives the estimate for $h(xy)$.

For $v \in \mathcal{M}_\infty(K)$,

$$H_v(x+y) \le \max\{|x|_v + |y|_v, 1\} \le 2H_v(x)H_v(y)$$

as one sees case by case. For example, when $|x|_v \le 1$, $|y|_v \le 1$, and $|x|_v + |y|_v > 1$, the inequality is $|x|_v + |y|_v \le 2$; when $|x|_v \ge 1$ and $|y|_v \ge 1$, the inequality is $|x|_v + |y|_v \le |x|_v|y|_v + |x|_v|y|_v$; and so on. We can drop the factor 2 in non-archimedean cases because of the ultrametric triangle inequality:

$$H_v(x+y) \le \max\{\max\{|x|_v, |y|_v\}, 1\} = \max\{H_v(x), H_v(y)\} \le H_v(x)H_v(y),$$

the last inequality again being checked case by case.

(b) For $v \in \mathcal{M}(K)$, let $w_1, \ldots, w_g$ be the extensions to L. Then by (¶′),

$$\begin{aligned}
\sum_{i=1}^{g} \log\max\{|x|_{w_i}, 1\} &= \sum_{i=1}^{g} \log\max\{|x|_v^{e(w_i,v)f(w_i,v)}, 1\} \\
&= \log\max\{|x|_v^{[L:K]}, 1\} \quad \text{by } (\flat) \\
&= [L:K]\log\max\{|x|_v, 1\}. \quad \blacksquare
\end{aligned}$$

3.3.2 Heights in function fields

Let K_1 be a field. A **function field** over K_1 is a field K containing K_1 as a subfield such that

- K_1 is algebraically closed in K, and
- K is finitely separably generated over K_1: there exists a finite (separating) transcendence basis $t_1, \ldots, t_n$ with $n > 0$ such that K is a finite separable extension of $K_2 = K_1(t_1, \ldots, t_n)$.

This is indicated notationally by K/K_1, and K_1 is called the **constant field**.

In example 3 of §1 we saw how every finitely generated field K, if not finite or a number field, is a function field over the algebraic closure K_1 of the prime subfield in K.

We now define a height function h on an arbitrary function field K/K_1. With $K_2 = K_1(t_1, \ldots, t_n)$ as in the definition, we write $K_2 = Q(R_2)$, where $R_2 = K_1[t_1, \ldots, t_n]$, and let R denote the integral closure of R_2 in K; thus $Q(R) = K$.

The essential valuations of the Krull domain R_2 can be grouped into the sets S_i, $1 \le i \le n$, described as follows. If

$$F_i = K_1(t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n),$$

then $v \in S_i$ means that v is the valuation on $K = F_i(t_i)$ that is trivial on F_i and is associated with a monic irreducible polynomial $p \in F_i[t_i]$; we let $\deg v$ denote the degree of that polynomial in the variable t_i, that is, the degree of the residue field of v over F_i. Not included in S_i is the $1/t_i$-adic valuation v_{1/t_i}; see §2.1.1, example 2. The residue field of v_{1/t_i} is F_i, hence $\deg v_{1/t_i} = 1$. We define

$$\mathcal{M}_i(K) = S_i \cup \{v_{1/t_i}\},$$

and then $\mathcal{M}_i(K)$ denotes the set of all valuations w on K extending some $v \in \mathcal{M}_i(K_2)$. Also define

$$\deg w = f(w, v) \deg v,$$

where $f(w, v)$, as usual, denotes the degree of the residue field of w over that of v. We choose a real number c satisfying $0 < c < 1$ and define the absolute values

$$|0|_v = 0 \quad \text{and for} \quad x \neq 0, \quad |x|_v = c^{v(x) \deg v},$$

where $v \in \mathcal{M}_i$ and $\mathcal{M}_i$ denotes either $\mathcal{M}_i(K_2)$ or $\mathcal{M}_i(K)$.

If K' is a finite separable extension of K, then from the definitions

$$\text{for} \quad x \in K, \quad |x|_w = |x|_v^{e(w,v)f(w,v)}, \tag{¶$'$}$$

just as we had in the number field case.

There is a product formula for each i:

$$\prod_{v \in \mathcal{M}_i} |x|_v = 1 \quad \forall x \neq 0.$$

This is quite obvious in the case R_2. For example consider

$$x = t_1^3 + 1/t_2.$$

As a polynomial in t_1, x is irreducible; we denote the corresponding valuation $v_x \in \mathcal{M}_1(K_2)$. The product formula for $i = 1$ written additively is

$$\sum v(x) \deg v = v_x(x) \deg v_x + v_{1/t_1}(x) \deg v_{1/t_1} = 1 \cdot 3 + (-3) \cdot 1 = 0.$$

The factorization of x into a "constant" in F_2 times monic polynomials in t_2 is

$$x = t_1^3(t_2 + 1/t_1^3)t_2^{-1},$$

and the formula when $i = 2$ is $1 - 1 = 0$. In general for $x \in F_i[t_i]$, as p ranges over the monic irreducible factors of x, $\sum v_p(x) \deg v_p$ is the degree of the polynomial x in t_i, which is cancelled by the term $v_{1/t_i}(x)$. This extends to the quotient field K_2: for $x, y \in F_i[t_i]$, $y \neq 0$,

$$\sum_{v \in \mathcal{M}_i(K_2)} v(x/y) \deg v = \sum v(x) \deg v - \sum v(y) \deg v = 0 - 0 = 0.$$

The proof of the product formula for $\mathcal{M}_i(K)$ goes as follows. Let $v \in \mathcal{M}_i(K_2)$. Because K is a finite separable extension

$$\sum_{w \in \mathcal{M}_i(K),\, w|v} e(w,v)f(w,v) = [K : K_2]. \tag{1}$$

Let N denote the norm map from K to K_2, and let $x \in K^*$. Then

$$\sum_{w \in \mathcal{M}_i(K),\, w|v} w(x)f(w,v) = v(N(x)). \tag{2}$$

(1) and (2) are standard results (see [BAC6, p.148–9] or [Lan83, p.14–19]) and so will be assumed as background material; (1) was already mentioned in §2.1.3. Multiplying (2) by $\deg v$ and adding over v, we see that the product formula for $\mathcal{M}_i(K)$ follows from that of $\mathcal{M}_i(K_2)$.

Define $\mathcal{M}(K) = \mathcal{M}_1(K) \cup \cdots \cup \mathcal{M}_n(K)$. Thus for $x \in K^*$, we have the product formula in additive notation

$$\sum_{v \in \mathcal{M}(K)} v(x) \deg v = 0.$$

We define the **height** on K by

$$h(0) = 0, \quad \text{and for } \ x \in K^*, \quad h(x) = \sum_{v \in \mathcal{M}(K)} \max\{-v(x)\deg v, 0\}.$$

Thus $h(x) \geq 0$ for all x. By the product formula, for any $y \in K^*$,

$$h(x) = \sum_v \max\{-v(xy)\deg v, -v(y)\deg v\}.$$

In particular, if $x \in K^*$ is written as a fraction r/s, $r, s \in K^*$, then

$$\begin{aligned} h(r/s) &= \sum_v \max\{v(r), v(s)\} \deg v \\ &= \sum_v \max\{-v(r), -v(s)\} \deg v. \end{aligned}$$

The last equality would look most peculiar if we were not in possession of the product formula!

As in the case $K = \mathbf{Q}$, the product formula yields a dramatic simplification when $K = K_1(t_1, \ldots, t_n)$ is a purely transcendental function field. Then if $x = r/s$ is a quotient of polynomials,

$$h(r/s) = \sum_{i=1}^{n} \max\{\deg_i(r), \deg_i(s)\},$$

where $\deg_i$ denotes the degree in the variable t_i. Notice that the **H1a** property is obvious when K_1 is finite: there are only finitely many polynomials of bounded total degree. However, since the elements of K_1 all have height 0, **H1a** fails when K_1 is infinite.

Proposition 3.3.2 is also true for function fields; of course now $\nu = 0$:

Proposition 3.3.3 *Let K/K_1 be a function field.*

(a) *For all $x, y \in K$,*

$$h(xy) \leq h(x) + h(y),$$

and

$$h(x+y) \leq h(x) + h(y).$$

(b) *Let L be a finite separable extension of K. Then*

$$\text{for } \ x \in K, \quad h_L(x) = [L:K]h_K(x).$$

The relation ($\flat$) of the previous section, and ($\P'$) are true, and the proof is unchanged from that of the previous proposition.

Note that the constant field of L is a finite separable extension of K_1.

3.4 Completion of the proof of Mordell-Weil

Proposition 3.4.1 *Let K be a either a number field or an arbitrary function field. Let h be the height on K as defined above, and let E be an elliptic curve defined over K.*

(a) *Then h applied to the abelian group $E(K)$ satisfies* **H2a** *and* **H2b**.

(b) *Assume now that K is either a number field or a finitely generated field of positive characteristic. Then h also satisfies* **H1a**.

Consequently, Mordell-Weil is proved for elliptic curves over number fields and over finitely generated fields of characteristic > 2.

Proof.

(a) If K is a number field, let R denote the integral closure of $\mathbf{Z}$ in K; if K is a function field then, as before, let R denote the intergral closure of $R_2 = K_1[t_1, \ldots, t_n]$ in K. Thus in both cases, $K = Q(R)$. We require a lemma.

Lemma 3.4.2 *Let $f(X)$ and $g(X)$ be coprime polynomials over K, and let $d = \max\{\deg f, \deg g\}$. Then for all $x \in K$ such that $g(x) \neq 0$,*

$$h(f(x)/g(x)) = d\,h(x) + O(1)$$

where $O(1)$ denotes a function of x that is bounded above and below by constants that depend only on f and g, and not on x.

Proof. By multiplying f and g by an appropriate nonzero member of R, we can assume that their coefficients are in R. Also choose $r, s \in R$ so that $x = r/s$. If $f(X) = a_0 + a_1 X + \cdots + a_d X^d$, where some of the higher a_i are 0 if $\deg f < d$, we define the homogeneous form

$$f(X_0, X_1) = X_0^d f(X_1/X_0) = a_0 X_0^d + a_1 X_0^{d-1} X_1 + \cdots,$$

and similarly $g(X_0, X_1) = X_0^d g(X_1/X_0)$. Thus

$$f(x)/g(x) = f(s,r)/g(s,r), \quad \text{where} \quad f(s,r),\, g(s,r) \in R.$$

We first prove the easier half

$$h(f(x)/g(x)) = \sum_v \log \max\{|f(s,r)|_v,\, |g(s,r)|_v\} \le d\, h(x) + C$$

for some constant C. Now

$$\begin{aligned} |f(s,r)|_v &\le \left(|a_0|_v |s^d|_v + \cdots + |a_d|_v |r|_v^d\right) \\ &\le (|a_0|_v + \cdots + |a_d|_v) \max\{|r|_v, |s|_v\}^d \\ &\le C_1 \max\{|r|_v, |s|_v\}^d, \end{aligned}$$

where $$C_1 = \max_v \{|a_0|_v + \cdots + |a_d|_v\}.$$

(There are only finitely many v for which some $|a_i|_v \neq 0$ or 1, and so this maximum exists.) Similarly

$$|g(s,r)|_v \le C_2 \max\{|r|_v, |s|_v\}^d, \quad \forall v,$$

and therefore the inequality is satisfied with $C = \log \max\{C_1, C_2\}$.

Since f and g are coprime, by Euclid's algorithm there exist polynomials a and b such that $a(X)f(X) + b(X)g(X) = 1$. Multiplying by an appropriate nonzero $k \in R$, we can assume that $a, b \in R[X]$, and

$$a(X)f(X) + b(X)g(X) = k.$$

If $e = \max\{\deg a, \deg b\}$, we also define $a(X_0, X_1) = X_0^e a(X_1/X_0)$ and $b(X_0, X_1) = X_0^e b(X_1/X_0)$. Thus

$$a(X_0, X_1) f(X_0, X_1) + b(X_0, X_1) g(X_0, X_1) = k X_0^{d+e}. \tag{1}$$

Since $\max\{\deg f, \deg g\} = d$, the reciprocal polynomials

$$\widetilde{f}(X) = X^d f(1/X) \quad \text{and} \quad \widetilde{g}(X) = X^d g(1/X)$$

are also coprime; say $a'(X)\widetilde{f}(X)+b'(X)\widetilde{g}(X) = k' \in R-\{0\}$, with $a', b' \in R[X]$. Interchanging the roles of X_0 and X_1, we have

$$X_1^d \widetilde{f}(X_0/X_1) = f(X_0, X_1), \quad X_1^d \widetilde{g}(X_0/X_1) = g(X_0, X_1),$$

and

$$a'(X_0, X_1)f(X_0, X_1) + b'(X_0, X_1)g(X_0, X_1) = k' X_1^{d+e}, \tag{2}$$

where $\qquad a'(X_0, X_1) = X_1^e a'(X_0/X_1), \quad b'(X_0, X_1) = X_1^e a'(X_0/X_1).$

By the half of the lemma already proved, we can choose a common constant C so that if c is any one of a, b, a', b', then

$$|c(s, r)|_v \le C \max\{|r|_v, |s|_v\}^e, \quad \forall v.$$

Then (1) implies

$$\begin{aligned} |r|_v^{d+e} &= |k|_v^{-1}|a(s, r)f(s, r) + b(s, r)g(s, r)|_v \\ &\le C \max\{|r|_v, |s|_v\}^e(|f(s, r)|_v + |g(s, r)|_v) \\ &\le 2C \max\{|r|_v, |s|_v\}^e \max\{|f(s, r)|_v, |g(s, r)|_v.\} \end{aligned}$$

(2) yields a similar upper bound for $|s|^{d+e}$, say with constant C'. Hence, with $C'' = 2\max\{C, C'\}$,

$$\max\{|r|_v, |s|_v\}^{d+e} \le C'' \max\{|r|_v, |s|_v\}^e \max\{|f(s, r)|_v, |g(s, r)|_v.\} \quad \forall v.$$

This implies

$$\begin{aligned} d\,h(x) &= d \textstyle\sum_v \log \max\{|r|_v, |s|_v\} \\ &\le h(f(s, r)/g(s, r)) + \log(C'') = h(f(x)/g(x)) + \log(C''). \end{aligned}$$

Proof of **H2a**: From Proposition 1.7.1, if $P = (x, y)$ is in $E(K)$, and not in $E(K)[2]$, then $x([2]P) = f(x)/g(x)$ where f and g are polynomials of degrees 4 and ≤ 3 respectively. (When char $K = 2$, $g = a_1^2x^2 + a_3^2$.) Also $\gcd(f, g) = 1$ by Proposition 1.7.9. By the lemma,

$$h([2]P) = 4h(P) + O(1).$$

Proof of **H2b**: We can assume that neither of P, Q is O, and that $P \ne \pm Q$. From the formula in Proposition 1.7.1, with $P = (x_1, y_1)$, $Q = (x_2, y_2)$ and $P + Q = (x_3, y_3)$,

$$x_3 = -x_1 - x_2 - a_2 + a_1\lambda + \lambda^2 \quad \text{where} \quad \lambda = \frac{y_2 - y_1}{x_2 - x_1}.$$

When we put all terms over the denominator $(x_2 - x_1)^2$ and substitute $y_1^2 = -a_1x_1y_1 - a_3y_1 + x_1^3 + \cdots$ from the Weierstrass equation, we obtain

$$x_3 = \frac{A_1x_1^2 + A_2x_1 + A_3 + A_4y_1}{(x_2 - x_1)^2},$$

where the A_i are polynomials in x_2, y_2 and the Weierstrass coefficients. (If it were not for the term in y_1, the lemma would immediately complete the proof.) Let $x_1 = r/s^2$, $y_1 = t/s^3$ where $r, s, t \in R$. From the b-form of the Weierstrass equation $y_1^2 = x_1^3 + (b_2/4)x_1^2 + \cdots$, we obtain the estimate

$$2\log\max\{|t|_v, |s|_v^3\} \le 3\log\max\{|r|_v, |s|_v^2\} + O(1),$$

and then, as in the easy half of the lemma, we deduce that

$$h(x_3) < 2h(x_1) + \gamma$$

where γ depends only on E and Q.

(b) To prove **H1a**, we can replace K by a finite separable extension L. For by Propositions 3.3.2 and 3.3.3,

$$\{x \in K : h_K(x) < \alpha\} \subset \{x \in L : h_L(x) < [L : K]\alpha\}.$$

Thus we can assume that K is Galois over F, where F is either $\mathbf{Q}$ or $\mathbf{F}_q(t_1, \ldots, t_n)$; say the Galois group is G. Letting G act on K on the right, the left action on $\mathcal{M}(K)$ is defined by $(\sigma w)(x) = w(x^\sigma)$. Since h_K is a sum over all $w \in \mathcal{M}(K)$,

$$\text{for } \ x \in K \ \text{ and } \ \sigma \in G, \quad h_K(x^\sigma) = h_K(x).$$

For a given positive α, let $x \in K^*$ satisfy $h(x) < \alpha$, and therefore $h(x^\sigma) < \alpha$, $\forall\, \sigma \in G$. Let $M(X) = X^N + s_1X^{N-1} + \cdots + s_N$ be the minimum polynomial of x over F. The coefficients s_j are symmetric functions in the conjugates x^σ of x over F. By the propositions just quoted,

$$\begin{aligned} h_K(s_j) &= h_K\left(\sum x^{\sigma_1}\cdots x^{\sigma_j}\right) \\ &< \binom{N}{j} j\alpha + \left[\binom{N}{j} - 1\right]\nu\log 2. \end{aligned}$$

It follows that for all j, $h_F(s_j) < \alpha'$ for an appropriate α', hence x is a root of $M(X)$ of degee at most $[K : F]$ and whose coefficients have bounded height.

Thus the proof of **H1a** is reduced to the cases $K = \mathbf{Q}$ and $K = \mathbf{F}_q(t_1, \ldots, t_n)$: those cases imply that the set of $M(X)$ is finite, hence the set of x is finite. And both these cases are obviously true as remarked in §3.3.1 and §3.3.2 respectively. ■

The standard theory of finitely generated abelian groups and Corollary 1.7.9 yield

Corollary 3.4.3 *Let E be an elliptic curve defined over K where K is either a number field or a finitely generated field of characteristic > 2. Then $E(K)$ has the form*

$$E(K) = \mathbf{Z}^r \oplus \mathcal{T}, \quad r \geq 0,$$

where $\mathcal{T}$ is the finite torsion subgroup and has the form

$$\mathcal{T} = C_{m_1} \oplus C_{m_2} \quad \textit{where} \quad m_1 \,|m_2 \,.$$

Also, $\operatorname{char} K$*, when positive, does not divide m_1.*

Hence for a positive integer m,

$$E(K)/[m]E(K) = (\mathbf{Z}/[m]\mathbf{Z})^r \oplus C_{d_1} \oplus C_{d_2} \quad \textit{where} \;\; d_i := \gcd(m, m_i).$$

When $\mathcal{T}$ is cyclic then $m_1 = 1$; and when $\mathcal{T} = 0$ then $m_1 = m_2 = 1$.

3.4.1 Function fields in characteristic 0

We add a few remarks concerning the unfinished proof of Mordell-Weil for function fields whose constant fields are number fields.

There would appear to be two ways to bridge the gap:

[A] Accept a less elementary proof which uses tools from algebraic geometry.

[B] Add an archimedean component h_a to h which satisfies **H2a** and **H2b**, hence the augmented height $h' = h + h_a$ still satisfies **H2a**, **H2b**, and such that h' also satisfies **H1a**.

Or, we might settle for a partial result:

[C] Prove Mordell-Weil for E defined over function fields with constant field a number field K_1 for which $j \notin K_1$.

$\boxed{\text{A}}$ We may as well consider the case of an arbitrary function field K/K_1 (in for a penny, in for a pound), and an elliptic curve defined over K. There are two cases: $j \notin K_1$ and $j \in K_1$.

In the first case, **H1** *is true* (and **H1a** may be false), but the proof of this is much more difficult than that of Proposition 3.4.1(b), and would seem to require considerable input from algebraic geometry (see [Lan83, ch.6] and [Sil94, ch.3]). Consider Example 1 from §1.7: over $K = \mathbf{Q}(t)$,

$$E_1 : y^2 = x^3 + tx^2 - tx, \quad \text{has} \quad j = 2^8(t+3)^3/(t+4) \notin \mathbf{Q} = K_1.$$

The point $(0, 0)$ has order 2 (but the points defined over $\mathbf{Q}$ do not form a subgroup because E_1 is not defined over $\mathbf{Q}$), and the point $P = (1, 1)$ has infinite order (since, for example, $[3]P = (a^2/d^2, ab/d^3)$ where $d = (t+2)(t+3)$ cannot be a torsion point by Proposition 2.10.3 — Nagell-Lutz). **H1** says that

the set of $P \in E_1(K)$ with $x(P) = r/s$, $r, s \in \mathbf{Q}[t]$ and $\max\{\deg(r), \deg(s)\} < \alpha$ is finite. But **H1a** is false for $K = \mathbf{Q}(t)$ since $h_K(x) = 0\ \forall x \in \mathbf{Q}$.

In the second case, where $j \in K_1$, there is a finite extension K' of K, in fact a quadratic extension except possibly when $j = 0$ or $j = 1728$ (this will be explained in detail in the next chapter) such that E is isomorphic over K' (in a sense defined in the next chapter) to an elliptic curve E' defined over K_1. The result is: $E'(K')/E'(K_1)$ is finitely generated. When $E'(K_1)$ is known to be finitely generated, *e.g.* when K_1 is a number field, this implies that $E'(K')$ itself is finitely generated.

First let us record an almost trivial instance, except that the proof quotes two results from elementary algebraic geometry that will be discussed only in Chapter 6.

Proposition 3.4.4 *Let E be an elliptic curve defined over the field K_1, and let K denote the purely transcendental extension $K_1(t_1, \ldots, t_n)$. Then*

$$E(K) = E(K_1).$$

Remark. Of course this does not extend to general function fields. For example if E is given by the equation $y^2 = x^3 + Bx + C$, $B, C \in K_1$, and $K = K_1(t)(\theta = \sqrt{t^3 + Bt + C})$, then $E(K)$ contains the point (t, θ) not in $E(K_1)$.

Proof. It is sufficient to treat the one variable case since that implies in succession

$$\begin{aligned} E(K_1(t_1, \ldots, t_n)) &= E(K_1(t_1, \ldots, t_{n-1})) \\ &= E(K_1(t_1, \ldots, t_{n-2})), \quad \textit{etc.} \end{aligned}$$

Thus suppose $E(K_1(t))$ contains a point $(x, y) \notin E(K_1)$. Since K_1 is algebraically closed in $K_1(t)$ and E is defined over K_1, therefore both x and y are transcendental over K_1. Thus $K_1(x, y)$ is a function field of genus 1. This contradicts Lüroth's theorem which says that $K_1(x, y) = K_1(t')$ for some $t' \in K_1(t)$, which has genus 0. ■

Over $K = \mathbf{Q}(t)$ consider the curves

$$E_2 : y^2 = x^3 - 4t^2x + 4t^3, \quad \text{and} \quad E_3 : y^2 = x^3 - 4x + 4$$

both with $j = -2^{10}3^3/11$. By methods to be developed, one finds that $E_3(\mathbf{Q})$ is infinite cyclic, generated by $P = (2, 2)$. (Nagell-Lutz discovers P and shows that the torsion subgroup is trivial; thus at this point it is certain that $E_3(\mathbf{Q})$ has the form $\mathbf{Z}^r$ for some integer $r \geq 1$.) Since these infinitely many points all have K-height 0, **H1** is false for $E_3(K)$.

Over the quadratic extension $K' = \mathbf{Q}(\sqrt{t})$ we have the group isomorphism

$$E_3(K') \longrightarrow E_2(K') \quad \text{where} \quad (a, b) \longmapsto (at, bt\sqrt{t})$$

and the result says that $E_3(K')$ is finitely generated. We can actually prove this because in this example it happens that K' is purely transcendental over $\mathbf{Q}$, and therefore by the proposition,

$$E_3(K') = E_3(\mathbf{Q}) = \langle (2,2) \rangle \approx E_2(K') = \langle (2t, 2t\sqrt{t}) \rangle.$$

$\boxed{\text{B}}$ I do not know if such an h' exists, although the main result of Altman in [Alt72] perhaps suggests that it should. However the obvious candidates do not seem to work in Lemma 3.4.2. To simplify notation, let us work over $K = \mathbf{Q}(t) = Q(\mathbf{Z}[t])$. Then for $r/s \in K$, where r, s are coprime polynomials in $\mathbf{Z}[t]$, we can try

$$h_a(r/s) = \log \max\{|r|, |s|\}$$

where $|-|$ extends the absolute value on $\mathbf{Q}$ in some manner to K.

Inspecting the proof of that lemma, it seems that $|-|$ is required to have the following four properties for $r, s \in \mathbf{Z}[t]$:

(o) For every α, the set $\{r : |r| + \deg(r) < \alpha\}$ is finite;
(i) $|r + s| \leq |r| + |s|$ (or conceivably $\leq C_1(|r| + |s|)$);
(ii) $|rs| \leq C_2|r||s|$;
(iii) $|r^2| \geq C_3|r|^2$ for some $C_3 > 0$.

Let $r = a_0 + a_1 t + \cdots + a_m t^m$. If $\tau \in \mathbf{C}$ is a transcendental number, then $f(r) = a_0 + a_1\tau + \cdots$ is a ring embedding $\mathbf{Z}[t] \hookrightarrow \mathbf{C}$, hence $|r| = |f(r)|$ satisfies (i) to (iii) — but not (o).

The following candidates all satisfy (o).

- $|r| = \max\{|a_i|\}$ satisfies (i) but not (ii): when all $a_i = 1$, $|r| = 1$ but $|r^2| = m + 1$;
- Lang's 'size' function [Lan66, p.49]) is the log of

  $$|r| = \max\{|a_0|, \ldots, |a_m|, e^m\}.$$

  He observes that

  $$|rs| \leq (|r||s|)^3;$$

  it is easy to reduce the exponent to 2, but it cannot be reduced to 1.
- regarding r as a function of the complex variable t,

  $$|r| = \exp \int_0^1 \log |r(e^{2\pi i\tau})| \, d\tau$$

  violates (i): improving a result of Mahler, Duncan [Dun66] shows that for polynomials r, s of degree $\leq m$,

  $$|r + s| \leq \binom{2m}{m}^{1/2} (|r| + |s|),$$

  but it seems (?) one cannot prove (i) with C_1 independent of m.

- $|r| = |a_0| + \cdots + |a_m|$ satisfies (i) and (ii) but randomized computer examples with large m indicate that $\inf |r^2|/|r|^2 = 0$, which also rules out modifications of this $|r|$ such as $|r|2^{\deg r}$ and $\max\{|r|, 2^{\deg r}\}$.

- $|r| = \left(\sum |a_i|^2\right)^{1/2}$ satisfies (i) and (iii) with $C_3 = 1$. However randomized examples indicate that $\sup |r^2|/|r|^2 = \infty$, which also rules out modifications like those in the last example. And the same story for $|r| = \left(\sum |a_i|^p\right)^{1/p}$ seems to be true for all $p > 1$.

$\boxed{\text{C}}$ If E is defined over K where K is a finite extension of the purely transcendental extension $K_1(t_1, \ldots, t_n)$ of the number field K_1 and $j \notin K_1$, then we can choose $t =$ one of the t_i such that $j \notin$ the algebraic closure K_2 of $K_1(t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$ in K. We can choose an embedding $K_2 \hookrightarrow \mathbf{C}$ and we have

$$E \text{ is defined over a finite extension } K' \text{ of } \mathbf{C}(t) \text{ and } j \notin \mathbf{C}.$$

In particular, $j \neq 0$ or 1728, and (as will be explained in the next chapter), replacing K' by a quadratic extension E is isomorphic with the generic-j curve

$$E' : y^2 + xy = x^3 - \frac{36}{j - 1728}x - \frac{1}{j - 1728},$$

and statement [C] would follow from

$$E'(K') \text{ is finitely generated.}$$

It may be feasible to give an elementary proof of this by showing that the height function h as defined in §3.3.2 satisfies **H1** (though not **H1a**).

3.5 The canonical height

For any field K, $\overline{K}$ denotes an algebraic closure, and $\overline{K}^s$ denotes a separable algebraic closure.

Let $x \in \overline{\mathbf{Q}}$, and let K be a number field containing x. By Proposition 3.3.2, the **absolute height**

$$h_{\text{abs}}(x) = \frac{1}{[K : \mathbf{Q}]} h_K(x)$$

is independent of the choice of K, and so is well-defined. A similar definition can be made in the function field case, based on Proposition 3.3.3 with $K_1(t_1, \ldots, t_n)$ playing the rôle of $\mathbf{Q}$, but now h_{abs} is "less absolute" since it depends on the choice of the separating transcendence basis. Thus if K is either a number field

or a function field with chosen transcendence basis, and E is an elliptic curve defined over K, then for $P \in E(\overline{K}^s)$, $h_{\text{abs}}(P)$ is unambiguous.

For any finite separable extension L of K, the absolute height satisfies **H2a** and **H2b** on the group $E(L)$ since h_L does. For such $E_{/K}$ the **canonical height** (of Néron and Tate) is defined by

$$\hat{h}(P) = \lim_{N\to\infty} \frac{h_{\text{abs}}([2^N]P)}{4^N}.$$

That the limit exists is the first thing proved in the next proposition.

The **Néron-Tate height pairing** is defined by

$$\langle P, Q\rangle = \frac{1}{2}\Big(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)\Big) \quad \text{for} \quad P, Q \in E(\overline{K}^s).$$

CAUTION: There are two normalizations of the canonical height: the *larger* one adopted here (and in [Cre92], [Kna92]), and the *smaller* one

$$\hat{h}_{\text{sm}} = \frac{1}{2}\hat{h}, \quad \text{so} \quad \langle P, Q\rangle = \hat{h}_{\text{sm}}(P+Q) - \hat{h}_{\text{sm}}(P) - \hat{h}_{\text{sm}}(Q),$$

as in [Hus87], [Sil88], and [Sil90].†. In [Sil86], the larger $\hat{h}$ is used, but $\langle P, Q\rangle$ is defined without the factor 1/2. ‡

Proposition 3.5.1 *Let E be a number field or a function field, and E an elliptic curve defined over K. In the following statements, P, Q, R denote arbitrary points in $E(\overline{K}^s)$ unless indicated otherwise.*

(a) *The limit defining $\hat{h}(P)$ exists; $\hat{h}(P) \geq 0$, $\forall P$, and $\hat{h}(O) = 0$.*

(b) *Let L be a finite separable extension of K. Then*

$$\forall\, P \in E(L), \quad \hat{h}(P) = h_{\text{abs}}(P) + O(1),$$

where $O(1)$ denotes a function of P that is bounded above and below by constants that depend only on E and L, and not on P. Hence the axiom **H1** *can be expressed in terms of the canonical height:*

H1 $\qquad \forall\, \alpha, \quad \{P \in E(K) : \hat{h}(P) < \alpha\}$ *is finite.*

(c) $\hat{h}([2]P) = 4\hat{h}(P)$, *hence*

$$\hat{h}(P) = \langle P, P\rangle.$$

† aPecs calculates ht= $\hat{h}_{\text{sm}}$.

‡ With this mixed notation the regulator (see below) should be defined as $\det(2^{-1}\langle P_i, P_j\rangle)$.

(d) *If* $h' : E(\overline{K}^s) \longrightarrow \mathbf{R}$ *satisfies* (b) *and* (c), *then* $h' = \hat{h}$.
(e) *The parallelogram law:* $\langle P, -Q \rangle = -\langle P, Q \rangle$, *equivalently,*

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

(f) $\hat{h}$ *is a quadratic form on* $E(\overline{K}^s)$, i.e.,

(1) $\hat{h}$ *is even:* $\hat{h}(-P) = \hat{h}(P)$, *and*

(2) *the Néron-Tate pairing is symmetric and bilinear.*

Consequently,

- *for all* $m \in \mathbf{Z}$,
$$\hat{h}([m]P) = m^2\hat{h}(P),$$
hence
$$\hat{h}(P) = \lim_{N\to\infty} \frac{h_{\mathrm{abs}}([N]P)}{N^2};$$
- *if* T *is a torsion point then for all* P, $\hat{h}(P+T) = \hat{h}(P)$. *In particular,* $\hat{h}(T) = 0$, *and therefore* $\langle P, T \rangle = 0$;
- *for any* $P_1, \ldots, P_m$, *let* $(\langle P_i, P_j \rangle)$ *denote the* $m \times m$ *symmetric height pairing matrix ; then for any row vector of integers* $N = (n_1, \ldots, n_m)$, *using matrix multiplication,*
$$\hat{h}(n_1P_1 + \cdots + n_mP_m) = N\left(\langle P_i, P_j \rangle\right) N^{\mathrm{tr}},$$
where tr *denotes the transpose. More generally, if* $N = (n_{jk})$ *is an* $s \times m$ *matrix of integers and* $P'_j = \sum_k n_{jk}P_k$, *then*
$$\left(\langle P'_j, P'_k \rangle\right) = N\left(\langle P_i, P_j \rangle\right) N^{\mathrm{tr}}.$$
If the P_i *are linearly dependent mod torsion,* i.e., *there are integers* $n_1, \ldots, n_m$, *not all* 0, *and a torsion point* T *such that* $n_1P_1 + \cdots + n_mP_m = T$, *then*
$$\det\left(\langle P_i, P_j \rangle\right) = 0.$$

(g) *When extended to the real vector space* $E(\overline{K}^s) \otimes_{\mathbf{Z}} \mathbf{R}$, $\hat{h}$ *is positive semi-definite; therefore we have*
— *the Cauchy-Schwartz inequality:*

$$\langle P_1, P_2 \rangle^2 \leq \hat{h}(P_1)\hat{h}(P_2),$$

— *hence the triangle inequality:*

$$\sqrt{\hat{h}(P_1+P_2)} \leq \sqrt{\hat{h}(P_1)} + \sqrt{\hat{h}(P_2)};$$

Remark. For converses to various statements in (f), and cases when $\hat{h}$ is positive definite, see the next proposition. In [Ser89, p.43], Serre poses the question: assuming $K = \mathbf{Q}$, does $\langle P, Q \rangle = 0$ imply that at least one of P, Q is a torsion point? As far as I know, this question remains unanswered.

Proof. For convenience, throughout this proof h stands for h_{abs}.

(a) Let P be defined over the finite separable extension L, and set $a_N = h([2^N]P)/4^N$. By **H2a** for $E(L)$, there exists a constant C such that for all $N \geq 1$, $|a_N - a_{N-1}| < C/4^N$. Hence for $0 \leq M < N$,

$$|a_N - a_M| \leq |a_N - a_{N-1}| + |a_{N-1} - a_{N-2}| + \cdots + |a_{M+1} - a_M| < C/(3 \cdot 4^M),$$

which proves that the sequence is Cauchy. Since $h([2^N]P) \geq 0$ with equality when $P = O$, the other comments follow.

(b)

$$|\hat{h}(P) - h(P)| = \lim_{N \to \infty} |a_N - a_0| \leq C/3.$$

(c)

$$\begin{aligned} \hat{h}([2]P) &= \lim \left(h([2^{N+1}]P)/4^N\right) \\ &= \lim \left((4h([2^N]P) + O(1))/4^N\right) = 4\hat{h}(P). \end{aligned}$$

(d) follows from $4^N h'(P) = h'([2^N]P) = h([2^N]P) + O(1)$ and the definition of $\hat{h}$.

(e)

The "proof" of (e) that appeared here was nonsense, as a student at Rutgers found out. For the time being I must simply refer to the proof given in [Sil86, Theorem 6.2, p. 216].

(f) (1) follows from $x(-P) = x(P)$. The symmetry of the pairing is evident. We wish to show that

$$\begin{aligned} T(P, Q, R) &:= \langle P + Q, R \rangle - \langle P, R \rangle + \langle Q, R \rangle, \\ &= \hat{h}(P + Q + R) - \hat{h}(P + Q) - \hat{h}(P + R) - \hat{h}(Q + R) \\ &\qquad + \hat{h}(P) + \hat{h}(Q) + \hat{h}(R) \end{aligned}$$

is identically 0. From the last expression, T is a symmetric function of the three variables, and from the previous expression and the first form of the parallelogram law, $T(P,Q,-R) = -T(P,Q,R)$. Thus

$$T(-P,-Q,-R) = (-1)^3 T(P,Q,R).$$

But $\hat{h}(-S) = \hat{h}(S)\ \forall S$ implies the opposite: $T(-P,-Q,-R) = T(P,Q,R)$.

Consequently, $\hat{h}([m]P) = \langle [m]P, [m]P\rangle = m^2\langle P,P\rangle = m^2\hat{h}(P)$. Hence, with all limits for $N \to \infty$,

$$\begin{aligned}\lim\,(h([N]P)/N^2) &= \lim\,((\hat{h}([N]P) + O(1))/N^2 \\ &= \lim\,(\hat{h}(P) + O(1)/N^2) = \hat{h}(P).\end{aligned}$$

If T has order m, then

$$\hat{h}(P+T) = \hat{h}([m](P+T))/m^2 = \hat{h}([m]P)/m^2 = \hat{h}(P).$$

Let $A = (\langle P_i, P_j\rangle)$. The fact that the pairing is bilinear implies that

$$\big(\langle P'_j, P'_k\rangle\big) = NAN^{\mathrm{tr}},$$

which reduces to $\hat{h}(n_1P_1 + \cdots)$ when N is $1 \times m$.

Let $F = F(n_1,\ldots,n_m)$ denote the m-ary quadratic form with coefficient matrix A. We will show in a moment that F is positive semi-definite, *i.e.,* the eigenvalues of A are all ≥ 0. Assuming this, when there is linear dependence mod torsion among the P_i, so that $F(n_1,\ldots) = \hat{h}(T) = 0$ with not all $n_i = 0$, then F is not positive definite: at least one eigenvalue is 0, and the determinant is 0.

(g) On the contrary, there would be real numbers $\alpha_1,\ldots,\alpha_m$ and points $P_1,\ldots,P_m$ such that for the extended $\hat{h}$,

$$\hat{h}(\alpha_1P_1 + \cdots + \alpha_mP_m) = \sum_{i,j=1}^{m} \alpha_i\alpha_j\langle P_i, P_j\rangle < 0.$$

With $P_1,\ldots,P_m$ fixed, $\hat{h}$ is a continuous (quadratic) function of the real variables $\alpha_1,\ldots,\alpha_m$, and therefore the above inequality is true with the α_i replaced by sufficiently close rational approximations. Then, multiplying through by a common denominator, we can assume that the α_i are integers. This contradicts $\hat{h}(P) \geq 0\,\forall P \in E(\overline{K}^s)$.

We recall a proof of Cauchy-Schwartz, independent of the foregoing. For all integers m and n,

$$\begin{aligned}0 \leq \hat{h}(mP_1 + nP_2) &= \langle mP_1 + nP_2, mP_1 + nP_2\rangle \\ &= m^2\hat{h}(P_1) + 2mn\langle P_1, P_2\rangle + n^2\hat{h}(P_2).\end{aligned}$$

Since this homogeneous quadratic in m and n is never negative, its discriminant satisfies

$$\langle P_1, P_2\rangle^2 - \hat{h}(P_1)\hat{h}(P_2) \le 0.$$

This gives Cauchy-Schwartz, which we can write as

$$|\langle P_1, P_2\rangle| \le \sqrt{\hat{h}(P_1)}\sqrt{\hat{h}(P_2)}.$$

Multiplying this by 2 and adding $\hat{h}(P_1) + \hat{h}(P_2)$ to both sides gives the triangle inequality. ■

Exercise: Let $E_{/K}$ be as in the proposition and let $P_1, P_2, \ldots \in E(K)$. From the parallelogram law we deduce

$$\hat{h}(P_1 + P_2) \le 2\hat{h}(P_1) + 2\hat{h}(P_2),$$

hence

$$\hat{h}(P_1 + P_2 + P_3) \le 2\hat{h}(P_1 + P_2) + 2\hat{h}(P_3) \le 4\hat{h}(P_1) + 4\hat{h}(P_2) + 2\hat{h}(P_3),$$

the latter exhibiting a peculiar asymmetry. Prove that for real numbers t_i

$$\hat{h}(P_1 + \cdots + P_n) \le t_1\hat{h}(P_1) + \cdots + t_n\hat{h}(P_n)$$

is a valid rule (for all such $E_{/K}$ and all $P_i \in E(K)$) iff the matrix

$$\begin{pmatrix} t_1 - 1 & -1 & \ldots & -1 \\ -1 & t_2 - 1 & \ldots & -1 \\ \vdots & & & \vdots \\ -1 & -1 & \ldots & t_n - 1 \end{pmatrix}$$

is positive semi-definite. For example,

$$\hat{h}(P_1 + P_2) \le 3\hat{h}(P_1) + \frac{3}{2}\hat{h}(P_2).$$

(*Hint:* The determinant of the matrix is

$$t_1 \cdots t_n \left(1 - t_1^{-1} - \cdots - t_n^{-1}\right).$$

Let $\delta(1, \ldots, n)$ denote this determinant and for any subsequence $i_1, \ldots, i_m$ of $1, \ldots, n$, let $\delta(i_1, \ldots, i_m)$ denote the determinant of the $m \times m$ matrix associated to the numbers $t_{i_1}, \ldots, t_{i_m}$. Then from [Gan60, vol.1, p.307], the condition is that all these principal subminors $\delta(i_1, \ldots, i_m) \ge 0$.) ■

The next proposition is concerned with $E(K)$ for which we have proved Mordell-Weil: K is a number field or a finitely generated field of characteristic > 2. Thus $E(K)$ can be written as a direct sum $\mathcal{T} \oplus \mathbf{Z}^r$ where $\mathcal{T}$ is the torsion

subgroup and r is the rank (over K). A **Mordell-Weil basis** of $E(K)$ is a $\mathbf{Z}$-basis $Q_1, \ldots, Q_r$ of the free part; thus every $P \in E(K)$ is uniquely expressible in the form

$$P = n_1 Q_1 + \cdots + n_r Q_r + T, \quad n_i \in \mathbf{Z} \text{ and } T \in \mathcal{T}.$$

The **regulator** of $E_{/K}$ is the $r \times r$ determinant

$$\mathrm{Reg}(E_{/K}) = \det(\langle Q_i, Q_j \rangle) \quad (= 1 \text{ when } r = 0).$$

It will be proved in the next proposition that $\mathrm{Reg}(E_{/K})$ is a well-defined positive number.

Proposition 3.5.2 *Let K be either a number field or a finitely generated field of characteristic > 2, and let E be an elliptic curve defined over K. Thus $E(K)$ is finitely generated. Let $\mathcal{T}$ denote the torsion subgroup and let $Q_1, \ldots, Q_r$ be a Mordell-Weil basis.*

(a) *For $P \in E(\overline{K}^s)$,*

$$\hat{h}(P) = 0 \iff P \in \mathcal{T}.$$

(b) *$\hat{h}$ extended to the real vector space $E(\overline{K}^s) \otimes \mathbf{R}$ is positive definite.*

For the remainder of the proposition, $P_1, \ldots$ denote points in $E(K)$.[‡]

(c) *$P_1, \ldots, P_m$ are dependent* $\bmod \mathcal{T}$, i.e., *$n_1 P_1 + \cdots + n_m P_m \in \mathcal{T}$ for some integers $n_1, \ldots, n_m$ not all 0, iff*

$$\det(\langle P_i, P_j \rangle) = 0.$$

Otherwise this determinant is positive. In particular

$$\mathrm{Reg}(E_{/K}) > 0.$$

(d) *Suppose $P_1, \ldots, P_r$ are independent* $\bmod \mathcal{T}$, *and let q denote the index in $E(K)$ of the subgroup generated by the P_i together with $\mathcal{T}$. Then*

$$\det(\langle P_i, P_j \rangle) = q^2 \,\mathrm{Reg}(E_{/K}).$$

Hence the regulator is characterized as the smallest height pairing determinant of r independent points; and r points in $E(K)$ constitute a Mordell-Weil basis iff their height pairing determinant equals the regulator.

(e) *Suppose for some $\alpha > 0$ and for some integer $m > 1$, the image of the finite set $S = \{P : \hat{h}(P) \le \alpha\}$ in the group $G := E(K)/[m]E(K)$ generates G (for example if S surjects onto G). Then S contains a Mordell-Weil basis.*

[‡] K may be replaced by any finite separable extension in order to accommodate points $P_1, \ldots, P_i \in E(\overline{K}^s)$.

Proof. (a) In the previous proposition we saw that $\hat{h}(P) = 0$ for $P \in \mathcal{T}$. The converse follows from **H1**: if $\hat{h}(P) = 0$ then $\hat{h}([m]P) = m^2\hat{h}(P) = 0$, hence the set $\{P, [2]P, [3]P, \dots\} \subset \{P' : \hat{h}(P') < 1 \text{ (say)}\}$ is finite.

(b) By **H1**, when $r > 0$, $\hat{h}$ assumes a minimum positive value μ_1 on $E(K)$, or, what amounts to the same thing, on $E(K)/\mathcal{T} = \mathbf{Z}^r = \mathbf{Z}Q_1 + \cdots + \mathbf{Z}Q_r$. By (a), $\hat{h}$ is positive on the latter group: $\hat{h}(n_1Q_1 + \cdots + n_rQ_r) = 0 \Longrightarrow n_1 = \cdots = n_r = 0$. It is still conceivable that $h(\alpha_1Q_1 + \cdots) = 0$ for some $\alpha_1, \dots \in \mathbf{R}$ not all 0. Suppose this is the case, *i.e.,* suppose $\hat{h}$ is not definite. Then, using the notation of the remark above, the diagonalized form of $\hat{h}$ is $\lambda_1 x_1'^2 + \cdots + \lambda_s x_s'^2$, where the λ_i are positive, but $s < r$. Then

$$B := \{(x_1', \dots) \in \mathbf{R}^r : \lambda_1 x_1'^2 + \cdots + \lambda_s x_s'^2 \le \mu_1/2\}$$

is convex, symmetric with respect to the origin, and has infinite volume since $s < r$. Translating back by the orthogonal transformation, $B' = BM$ is also convex and symmetric, and has infinite volume. By a standard result of the geometry of numbers,† B' contains a nonzero integral point $P = (n_1, \dots, n_r)$. Then $0 < \hat{h}(P) \le \mu_1/2$, contradicting the minimality of μ_1.

(c) and (d) We proved in the previous proposition that the height pairing determinant is 0 when the P_i are dependent. Now if $Q_1, \dots, Q_r$ is any Mordell-Weil basis, then $(\langle Q_i, Q_j \rangle)$ is the symmetric matrix of a concrete realization of the positive definite form $\hat{h}$. Consequently the determinant of this matrix is positive. Next, if $P_1, \dots, P_r$ are independent mod $\mathcal{T}$, then by basic abelian group theory, there exists a Mordell-Weil basis $Q_1', \dots$ such that for $1 \le i \le r$, $P_i = d_iQ_i'$ where the d_i are positive integers, each a multiple of the preceding: $d_{i-1} | d_i$ for $2 \le i \le r$. Moreover, $q = d_1 \cdots d_r$. The displayed formula in (d) follows from the bilinearity of height pairing and elementary determinant theory.

An independent set $P_1, \dots, P_i$ of fewer than r points can be augmented to an independent set $P_1, \dots, P_r$. The height pairing determinant of the augmented list is positive, *i.e.,* $\hat{h}(n_1P_1 + \cdots + n_rP_r) > 0$ unless all $n_j = 0$. Hence $\hat{h}(n_1P_1 + \cdots + n_iP_i) > 0$ unless $n_1 = \cdots = n_i = 0$, which means that the $i \times i$ height pairing determinant of the original list of points is positive.

(e) This result, attributed to Zagier in [Sil90, Prop. 7.2], sharpens earlier statements such as the "descent lemma" in [Ser89, p. 53].

Since $\hat{h}(T) = 0 \, \forall T \in \mathcal{T}$, therefore $S \supset \mathcal{T}$, hence S contains in fact a set of generators for the whole group $E(K)$. Here is a proof by contradiction.

Suppose $Q \notin \langle S \rangle$, the subgroup generated by S; we can choose Q of minimal height (**H1** again). Let $Q = P + [m]R$ where $P \in \langle S \rangle$, hence $P + Q$ and R are not in $\langle S \rangle$. By minimality,

$$\hat{h}(P + Q) = \hat{h}(P) + \hat{h}(Q) + 2\langle P, Q \rangle \ge \hat{h}(Q),$$

† [Har-Wr54, theorem 446] is often quoted; see also [Cas91, ch.4] for the generalization by van der Corput.

hence $2\langle P, Q\rangle \geq -\hat{h}(P)$, and therefore

$$m^2\hat{h}(R) = \hat{h}(Q-P) = \hat{h}(Q) + \hat{h}(P) - 2\langle P, Q\rangle \leq \hat{h}(Q) + 2\hat{h}(P) < 3\hat{h}(Q).$$

Thus

$$\hat{h}(R) < \frac{3}{m^2}\hat{h}(Q) \leq \frac{3}{4}\hat{h}(Q),$$

which contradicts the minimality of $\hat{h}(Q)$. ■

Corollary 3.5.3 *Let $E_{/K}$ be as in the theorem and suppose that for some $\alpha > 0$ the set $S := \{P \in E(K) : \hat{h}(P) \leq \alpha\}$ contains r independent points $P_1, \ldots, P_r$ such that the subgroup H of $E(K)$ generated by $\{P_1, \ldots, P_r\} \cup \mathcal{T}$ contains S. Then $H = E(K)$, that is, $P_1, \ldots, P_r$ constitute a Mordell-Weil basis.*

Proof. Choose a prime p not dividing either $|\mathcal{T}|$ or the index $i := [E(K) : H]$. The first assumption on p implies that $V := E(K)/[p]E(K)$ is a vector space over $\mathbf{F}_p$ of dimension r. To deduce the corollary from statement (e) of the proposition, it remains to show that the images of the P_i are linearly independent in V. If not, there would be a relation

$$[p]P = [n_1]P_1 + \cdots + [n_r]P_r$$

for some $P \in E(K)$ with not all $n_j \equiv 0 \bmod p$. But the existence of such a P implies that $p|i$, which contradicts the second assumption on p. ■

3.5.1 Calculating the canonical height: a first look

$\boxed{h \text{ denotes } h_{\text{abs}}}$

Because of poor convergence, in the case of number fields, $\lim h([N]P)/N^2$ does not afford a good method of calculating $\hat{h}(P)$. There is a practical algorithm for this due to Tate, as refined by Silverman [Sil88] (the bibliography there contains references to the work of others), but we must postpone its description because it involves a number of matters not yet discussed. Thus for the time being, we suppose we can calculate $\hat{h}(P)$ to any desired degree of accuracy, if necessary using the limit formula and persistence. Here is a numerical example over $K = \mathbf{Q}$.

The first curve in the lists [AntIV] and [Cre92] with a point P of infinite order is

$$y^2 + y = x^3 - x, \qquad P = (0, 0). \qquad \textbf{A37}$$

Calculation produces the following results:

n	$h([2^n]P)/4^n$	$h([3^n]P)/9^n$
1	0	0
2	0	0
3	.0433	.0480
4	.05029	.05119
5	.0511006	.05106
6	.0511008	.0511134
7	.0511014	.05111151
8	.0511034	
9	.05111065	
10	.05111140815	

The actual value to 11 decimal places, as calculated by the Tate-Silverman algorithm is[†]

$$\hat{h}(P) = .05111140824 \ .$$

Thus the results in the table give us several digits of accuracy, but at great cost: the numerator and denominator of $x([2^{10}]P)$ each contain more than 5800 digits, while those of $x([3^7]P)$, over 11,000. And we must maintain the exact rational values of the coordinates because we need to know $x([N]P)$ in lowest terms, after cancellation of common factors.

Also, to make effective use of Proposition 3.5.2(e) one needs an explicit value of U in

$$h(P) - \hat{h}(P) < U, \quad \forall P \in E(K).$$

($\{P : h(P) < \alpha + U\}$ contains the set S of statement (e).) At this point we state two estimates for U; the first due to Silverman [Sil90] is widely used, the second due to Siksek [Sik95] is very recent. The proofs depend on facts that we have not yet sufficiently discussed and must be postponed; the proof of Siksek's estimate will be given in Chapter 7, and that of Silverman somewhat later.

In all the examples I have seen, Siksek's estimate is better, *i.e.,* smaller, and in many cases substantially smaller; we present the example **A37**(**Q**) after stating the two estimates. However, their approaches to the problem are quite different making a general comparison difficult. Silverman also gives a lower bound $L < h - \hat{h}$, which we also quote.

If $|x|'_v$ and $\hat{h}_{\mathrm{sm}}$ denote the absolute values and canonical height as defined in [Sil90], the relation with our notation is

$$|x|_v = \begin{cases} |x|'_v & \text{if } \ v \in \mathcal{M}_\infty(K), \\ |x|'^{e_v f_v} & \text{if } \ v \in \mathcal{M}_p(K), \ \ p < \infty, \end{cases}$$

and
$$\hat{h}(P) = 2\hat{h}_{\mathrm{sm}}(P).$$

[†]This is `2 * ht(0)` as calculated by **aPecs**.

Note also that there is only one $|x|'_v$ for a conjugate pair of embeddings $K \hookrightarrow \mathbf{C}$, whereas there are two identical $|x|_v$. Silverman's result, incorporating the improvement in a special case due to Tate ([Sil90, Theorem 4.1]), in our notation is:

Let E be an elliptic curve defined over a number field K. Define

$$h_\infty(x) = \frac{1}{[K:\mathbf{Q}]} \sum_{v \in \mathcal{M}_\infty(K)} \max\{1, \log |x|_v\},$$

and $$M = \frac{1}{6}h(\Delta) + \frac{1}{6}h_\infty(j) + h_\infty(b_2/12) + \begin{cases} 0 & \text{if } b_2 = 0, \\ \log 2 & \text{if } b_2 \neq 0, \end{cases}$$

Then one can take

$$L = -2.14 - M,$$

$$U = 1.946 + M + \frac{1}{12}h(j) - \frac{1}{4}\sum{}^{*} \log \max\{1, |j|_v\},$$

where the sum is over those valuations on K for which $v(j) = -1$ and $v(c_4) = 0$.

Siksek's estimate for U is as follows. Let $v \in \mathcal{M}(K)$ and let K_v denote the completion as usual. The precise definition of the quantities σ_v that appear in the estimate† can only be given in Chapter 7; for now we must be content with the bound $\sigma_v \leq 1/3$. Define the polynomials

$$\begin{aligned} f(x) &= 4x^3 + b_2x^2 + 2b_4x + b_6, \qquad & f_1(x) = x^4 f(1/x), \\ g(x) &= x^4 - b_4x^2 - 2b_6x - b_8, \qquad & g_1(x) = x^4 g(1/x). \end{aligned}$$

Thus from §1.7.2, if $P = (x, y) \in E(K_v)$, then the x-coordinate of $[2]P$ is $g(x)/f(x) = \phi_2/\psi_2^2$. From the identity in Proposition 1.7.12(b), $f(x)$ and $g(x)$ cannot simultaneously be 0; since $g_1(0) = 1$, it follows that $f_1(x)$ and $g_1(x)$ cannot simultaneously be 0.

Next define

$$\begin{aligned} D_v &= \{x \in K_v : |x|_v \leq 1 \text{ and } f(x) \in K_v^2\}, \\ D'_v &= \{0\} \cup \{x \in K_v^* : |x_v|_v \leq 1 \text{ and } f_1(x) \in K_v^2\}, \\ d_v &= \inf_{x \in D_v} \max\{|f(x)|_v, |g(x)|_v\}, \\ d'_v &= \inf_{x \in D'_v} \max\{|f_1(x)|_v, |g_1(x)|_v\}. \end{aligned}$$

†These quantities are called μ_v in [Sik95], but we had already appropriated the symbol μ for successive minima; see the next section.

Since the two D are compact subsets of K_v (in the v-adic topology), the two inf are attained; and since the two polynomials cannot both be 0 in either case, we have $d_v > 0$ and $d'_v > 0$. Finally, define

$$\epsilon_v = 1/\min\{d_v, d'_v\}.$$

Let E be an elliptic curve defined over a number field K. Then $\forall\, Q \in E(K)$,

$$h(Q) - \hat{h}(Q) \le U = \frac{1}{[K:\mathbf{Q}]} \sum_{v \in \mathcal{M}(K)} \sigma_v n_v \log(\epsilon_v),$$

where $n_v = e_v f_v$ ($= 1$ when v is archimedean with our conventions concerning $\mathcal{M}_\infty$). Moreover, for all v, $0 \le \sigma_v \le 1/3$, and for nonarchimedean v with residue characteristic p,

$$0 \le n_v \log(\epsilon_v) \le 2 \left\lfloor \frac{v(\Delta)}{2} \right\rfloor \log p.$$

In [Sik95, Lemma 2.3(5)] the upper bound for $n_v \log(\epsilon_v)$ has 4Δ in place of Δ. The present version, which is an improvement when $p = 2$, was not noticed before the paper was printed. In working out numerical examples one usually uses better bounds, especially for the σ_v — as will be explained in Chapter 7.

For $E = \mathbf{A37}$, $b_2 = 0$, $c_4 = 48$, $\Delta = 37$, $j = 2^{12}3^3/37$, hence

$$\text{Silverman:} \qquad -4.08 < h(Q) - \hat{h}(Q) < 3.947 \quad \forall Q \in E(\mathbf{Q}).$$

In Siksek's estimate, with an obvious simplification of notation, $\epsilon_p = 0$ for all primes p, and the only contribution to U is ϵ_∞:

$$f(x) = 4x^3 - 4x + 1, \quad g(x) = x^4 + 2x^2 - 2x + 1,$$

$$f_1(x) = x^4 - 4x^3 + 4x, \quad g_1(x) = x^4 - 2x^3 + 2x^2 + 1.$$

One finds that $d_\infty = g(\lambda_2) = .6115^-$, where λ_2 is the middle root of f, and that $d'_\infty = g_1(0) = 1$. Thus $\epsilon_\infty = 1/.6115^- = 1.635$, and

$$h(Q) - \hat{h}(Q) < \frac{1}{3}\log(1.635) = .164\ .$$

From above we have $\hat{h}(P) = \hat{h}(0,0) = .051$, hence for $Q \in E(\mathbf{Q})$ (*cf.* Proposition 2.2.2),

$$\begin{aligned}\{Q : \hat{h}(Q) < .051\} &\subset \{Q : h(Q) < .051 + .164 = .215\} \\ &= \left\{\left(\frac{m}{e^2}, \frac{n}{e^3}\right) \in E(\mathbf{Q}) : \max\{|m|, e^2\} < \exp(.215) = 1.24\right\}.\end{aligned}$$

The last set consists of

$$P = (0,0), \quad -P = (0,-1), \quad [2]P = (1,0),$$

$$[-2]P = (1,-1), \quad [3]P = (-1,-1), \quad [-3]P = (-1,0).$$

It follows that $\hat{h}(P) = .051\ldots$ is the first successive minimum, as defined in the next section (taking the upper bound .164 on faith — a proof for infidels is given in Corollary 3.7.6, where it is also proved that the rank of $E(\mathbf{Q})$ is 1).

Let us now consider a function field example in positive characteristic:

$$P_0 = (1,1) \quad \text{on} \quad E : y^2 = x^3 + tx^2 - tx \quad \text{over} \ \ \mathbf{F}_p(t),$$

where p is an odd prime and t is a transcendental. Let $P_\nu = [2^\nu]P_0$ and let $x(P_\nu) = m_\nu/e_\nu^2$, where m_ν and e_ν are relatively prime polynomials. By Corollary 1.7.2,

$$x(P_{\nu+1}) = \frac{(m_\nu^2 + te_\nu^4)^2}{4m_\nu e_\nu^2(m_\nu^2 + tm_\nu e_\nu^2 - te_\nu^4)}.$$

By an induction left to the reader, t does not divide m_ν or e_ν, and there is no cancellation in the above fraction. Hence

$$\deg m_\nu = 2 \cdot 4^{\nu-1}, \quad \deg e_\nu = 4^{\nu-1} - 1,$$

$$\hat{h}(P_0) = \lim_{\nu\to\infty} 4^{-\nu} \max\{\deg m_\nu, \deg e_\nu^2\} = \frac{1}{2},$$

a rational number! In fact in the function field case, canonical heights are always rational numbers, being expressible in terms of "intersection numbers" — but that is another long chapter in algebraic geometry. (One expects canonical heights in the number field case to be either 0 or transcendental, but I don't know of any results in that direction.)

3.5.2 The successive minima

Notation: for $P_1, \ldots \in E(K)$, $r(P_1, \ldots, P_i)$ denotes the rank of the subgroup of $E(K)$ generated by the set $\{P_1, \ldots P_i\}$.

Property **H1** as stated in Proposition 3.5.1(b) permits us to define the **successive minima** of $E(K)$ as the increasing sequence of positive real numbers $\mu_1 \le \mu_2 \le \cdots \le \mu_r$ for which there exist $P_i \in E(K)$ with $\hat{h}(P_i) = \mu_i$ such that for all $P \in E(K)$,

$$\hat{h}(P) \le \mu_i \implies r(P_1, \ldots, P_i, P) = i.$$

In other words, assuming $r > 0$, among all P with $\hat{h}(P) > 0$, some P_1 is chosen with minimum positive canonical height. Then inductively, having chosen $P_1, \ldots, P_i$, as long as $i < r$, among all P that increase the rank, *i.e.,* $r(P_1, \ldots, P_i, P) = i + 1$, P_{i+1} is chosen with minimum canonical height.

(This definition raises the question, at least when K is a number field: Can $\mu_i = \mu_{i+1}$? I guess that this never happens when $K = \mathbf{Q}$, but can happen for other number fields. The latter is suggested by this: if E is defined over a subfield K_0 of K, then K_0-conjugate points in $E(K)$ have the same canonical height. For example when E is $y^2 = x^3 + x^2 - 1$ the points $P = (1+i, 1+2i)$ and $\bar{P} = (1-i, 1-2i)$ have the same height $\hat{h}(P) = \hat{h}(\bar{P}) = .4436$ by Silverman's algorithm mentioned in the previous section. One can check that they are independent, however they do not provide an example of $\mu_1 = \mu_2$ since $\hat{h}(P + \bar{P}) = \hat{h}(1, -1) = .3539$).

Proposition 3.5.4 *Let K be a number field or a finitely generated field of characteristic > 2, let $\mu_1, \dots, \mu_r$ be the successive minima of $E_{/K}$, let $P_1, \dots, P_r$ be independent points in $E(K)$ such that $\hat{h}(P_i) = \mu_i$, and let H be the subgroup generated by $\{P_1, \dots, P_r\} \cup \mathcal{T}$.*

(a)
$$\mu_1 \cdots \mu_r \le \Gamma_r \mathrm{Reg}(E_{/K}),$$

where $\Gamma_1 = 1$, $\Gamma_2 = 4/3$, $\Gamma_3 = 2$, $\Gamma_4 = 4$, $\Gamma_5 = 8$, $\Gamma_6 = 64/3$, $\Gamma_7 = 64$, $\Gamma_8 = 256$, and for $r \ge 9$,

$$\Gamma_r = \left(\frac{4}{\pi}\right)^r \Gamma\left(\frac{r}{2} + 1\right)^2.$$

(b)
$$[G : H]^2 \le \frac{\Gamma_r \det(\langle P_i, P_j \rangle)}{\mu_1 \cdots \mu_r},$$

hence if the right side is < 4 then $P_1, \dots, P_r$ form a Mordell-Weil basis.

(c) (Minkowski) *In any case if $r \le 4$ then $P_1, \dots, P_r$ form a Mordell-Weil basis.*

Remark. It may be that the restriction $r \le 4$ in statement (c) can be lifted, at least when K is a number field. Let us elucidate the situation.

Let $F(x_1, \dots, x_r)$ be a real positive definite quadratic form. Then F satisfies

H1 $\quad \forall\, \alpha > 0, \quad \{N = (n_1, \dots) \in \mathbf{Z}^r : F(n_1, \dots) < \alpha\}$ is finite.

(For let A be the symmetric matrix of F, so $F(x_1, \dots) = XAX^{\mathrm{tr}}$ where $X = (x_1, \dots)$, and let M be a diagonalizing orthogonal matrix, *i.e.*, the substitution $X = X'M$ gives $F = \lambda_1 x_1'^2 + \cdots + \lambda_r x_r'^2$, where the eigenvalues λ_i are all positive. Then $F < \alpha$ implies that the squared length $X'X'^{\mathrm{tr}}$ is bounded. Since M is orthogonal, $x_1^2 + \cdots = XX^{\mathrm{tr}} = X'X'^{\mathrm{tr}}$ is bounded.)

Thus the successive minima $0 < \mu_1 \le \cdots \le \mu_r$ can be defined: choose a nonzero $P_1 \in \mathbf{Z}^r$ with minimum "height" $F(P_1) = \mu_1$. Inductively, with

$P_1, \ldots, P_i$ chosen and $i < r$, choose $P_{i+1} \in \mathbf{Z}^r$ with minimum $F(P_{i+1}) = \mu_{i+1}$ subject to increasing the dimension, *i.e.,* P_{i+1} not in the subspace of $\mathbf{R}^r$ spanned by $P_1, \ldots, P_i$. We adapt a proof of Minkowski, following [Cas78, p.257], to yield: if $r \leq 4$, then independent points $P_1, \ldots, P_r$ chosen to give the successive minima form a basis of $\mathbf{Z}^r$.

The result actually given by Minkowski (Cassel's Lemma 1.2[‡]) is this: let $Q_1, \ldots, Q_r$ be a basis of $\mathbf{Z}^r$, ordered so that $M_i = F(Q_i)$ satisfy $0 < M_1 \leq \cdots \leq M_r$, and suppose that for $1 \leq J \leq r$,

$$Q = s_1 Q_1 + \cdots + s_J Q_J,\ s_j \in \{0, 1, -1\},\ s_J = 1 \implies F(Q) \geq M_J.$$

If $r \leq 4$, then the basis $Q_1, \ldots, Q_r$ is **Minkowski reduced**, *i.e.,* $\forall i$, Q_i has minimum height among all Q such that $Q_1, \ldots, Q_{i-1}, Q$ can be included in a basis of $\mathbf{Z}^r$. There are counterexamples when $r > 4$. Here $Q_1, \ldots$ is assumed to be a basis to begin with, whereas we want to prove that a certain sequence $P_1, \ldots$ is a basis. Nevertheless, a similar proof works. Cassels also gives references for the difficult analysis of cases with higher r.

But do the quadratic forms that occur as $\hat{h}$ have special properties not shared by all positive definite forms that allow (c) to be extended to higher r? Corollary 3.5.3 with $\alpha = \mu_r$ doesn't immediately apply since conceivably we could have situations such as $\hat{h}(P'_r) = \mu_r$, $[3]P'_r = P_1 + [5]P_r$.

Proof. (a) This is a standard result for positive definite quadratic forms. See [Cas78,p.262] and [Sie88,p.26]. For $r \leq 8$ the constants Γ_r are best possible; see [Cas59,p.332].

(b) follows from (a) and Proposition 3.5.2(d).

(c) We suppose by induction on J that for any choice of independent $P_1, \ldots, P_J$ whose heights are the successive minima, that there exists a Mordell-Weil basis $P_1, \ldots, P_J, Q_{J+1}, \ldots, Q_r$. This is true when $J = 0$.

Assuming the result for J, consider any

$$P = p_1 P_1 + \cdots + p_J P_J + q_{J+1} Q_{J+1} + \cdots + q_r Q_r$$

such that $P_1, \ldots, P_J, P$ are independent, with $\hat{h}(P) = \mu_{J+1}$; then some $q_i \neq 0$. In fact by elementary group theory, the Q_i can be chosen so that $q_{J+1} > 0$ and $q_i = 0$ for $i > J + 1$. Among all the possibilities for the Q_i and P with $\hat{h}(P) = \mu_{J+1}$, $q_{J+1} > 0$ and $q_i = 0$ for $i > J + 1$, choose $Q_{J+1}, \ldots, Q_r$ with $\hat{h}(Q_{J+1})$ minimal; and then among the P choose one with q_{J+1} positive and minimal. Since $P_1, \ldots, P_J, Q_{J+1}$ are independent, therefore $\hat{h}(Q_{J+1}) \geq \mu_{J+1}$. We will prove that $q_{J+1} = 1$, hence P can replace Q_{J+1} in the basis, and the induction proceeds with $P_{J+1} = P$.

For convenience write $Q_i = P_i$ and $q_i = p_i$, for $1 \leq i \leq J$, and now for $i \leq J$, replace Q_i by $-Q_i$ as necessary, so that in $P = q_1 Q_1 + \cdots + q_{J+1} Q_{J+1}$

[‡]In two of the displayed formulas in the bottom third of p.258, f_{nn} should be f_{kk}.

all $q_i \geq 0$. Let $S = \{i : q_i > 0\}$, thus $J+1 \in S$, and let $S' = S - \{J+1\}$. Define $Q = \sum_{i \in S} Q_i$. Also let $h_{ij} = \langle Q_i, Q_j \rangle$ for $i, j \leq r$.

By definition of μ_i and the fact that $Q_i + Q_j$ could replace Q_j in the basis, therefore $\forall\, i, j \leq J+1$, $\hat{h}(Q_i + Q_j) = h_{ii} + h_{jj} + 2h_{ij} \geq h_{jj}$. Thus

$$\forall\, i, j \leq J+1, \quad h_{ii} + 2h_{ij} \geq 0. \tag{¶}$$

Since $P_1, \ldots, P_J, Q, Q_{J+2}, \ldots, Q_r$ is a basis which gives the same q_{J+1} and $q_i = 0$ for $i > J+1$, by minimality we have

$$\hat{h}(Q) - \hat{h}(Q_{J+1}) = \sum_{i,j \in S} h_{ij} - h_{J+1,J+1} \geq 0.$$

Call this quantity A. An easy calculation shows that

$$\hat{h}(P) - \hat{h}(P - Q) = A + 2B \quad \text{where} \quad B = \sum_{i \in S'} (a_i - 1) \sum_{j \in S'} h_{ij}.$$

It remains to show that $B \geq 0$; for then $\hat{h}(P - Q) \leq \mu_{J+1}$, and the coefficient of Q_{J+1} in $P - Q$ is $q_{J+1} - 1$. If $q_{J+1} > 1$ then, by definition of the successive minima, $\hat{h}(P - Q) = \mu_{J+1}$ and the minimality of q_{J+1} is contradicted.

$B \geq 0$ follows from

$$\sum_{j \in S'} h_{ij} = (2 - |S|/2)h_{ii} + \frac{1}{2} \sum_{j \in S',\, j \neq i} (h_{ii} + 2h_{ij}) \geq 0,$$

which in turn follows from (¶) and $2 - |S|/2 \geq 2 - r/2 \geq 0$ — provided $r \leq 4$. ■

3.6 Algorithms for Mordell-Weil bases: a first look

Let E be an elliptic curve defined over the field K such that $E(K)$ is finitely generated. The explicit determination of

$$G := E(K) = \mathcal{T} \oplus \mathbf{Z}^r = \mathcal{T} \oplus \langle P_1, \ldots, P_r \rangle$$

can be attempted along the following lines; we say *attempted* because to date there is no algorithm that is guaranteed to work, even for $E_{/\mathbf{Q}}$.

0. Determine a suitable Weierstrass equation for $E_{/K}$.
1. Determine $\mathcal{T}$.
2. Find an upper bound r_1 for the rank.
3. Find a lower bound r_0 for the rank, and keep at this step and step 2 until $r_0 = r_1$. Then the common value is the rank r.

This step is almost always carried out by finding actual points $Q_1, \ldots, Q_r$ that are independent mod torsion, so that

$$H := \mathcal{T} \oplus \langle Q_1, \ldots, Q_r \rangle$$

is of finite index in G.

4. Estimate the index $[G : H]$ and refine $Q_1, \ldots, Q_r$ until $H = G$.

Throughout this section we consider only E defined over number fields K since the case of function fields has a decidedly different flavor. As in the previous section, h denotes h_{abs}, and U denotes an upper bound for $h(P) - \hat{h}(P)$ for all $P \in E(K)$.

The following elaborations of steps 0–4 are at times rather terse since they involve concepts that will be explained only later; for example we use the terms *twist* and *isogeny* which are defined only in Chapters 4 and 6 respectively. And a few new terms (L-function, root number,...) are employed without giving a forward reference.

$\boxed{\mathbf{0}}$ Starting with a Weierstrass equation over K, we obtain one over the ring of integers $\mathcal{O}$ of K by substituting x/d^2, y/d^3 for x, y for appropriate d. When possible (*e.g.* when the class number of $\mathcal{O}$ is 1) the Laska-Kraus algorithm, to be discussed in Chapter 5, transforms E to "global minimal form".

$\boxed{\mathbf{1}}$ Five inputs to this part of the algorithm, not necessarily in the order in which they should be applied, are

(i) $$\mathcal{T} = \{P \in G : \hat{h}(P) = 0\} \subset \{P \in G : h(P) < U\}.$$

The implementation of this and of step 3 below requires a "search engine" to find all $P \in E(K)$ with $h(P) <$ a given α — what is needed is an efficient procedure to find all $x \in K$ satisfying

$$h(x) < \alpha, \quad \text{and} \quad 4x^3 + b_2x^2 + 2b_4x + b_6 \in K^{*2}. \tag{¶}$$

Of course the engine first filters out x that do not satisfy some obviously necessary condition; *e.g* if $\sigma : K \hookrightarrow \mathbf{R}$, and if the image of the polynomial in (¶) has three real roots $\lambda_1 < \lambda_2 < \lambda_3$, then only x in the intervals $\lambda_1 \le x \le \lambda_2$ and $\lambda_3 \le x$ need be considered. But otherwise the engine works by brute force — it tests all filtered x — *unless* $E(K)$ contains a point of order 2 (and $\mathcal{O}$ has class number 1), and then the engine take can take advantage of the special form that a point $(x, y) \in E(K)$ must have, as will be explained in Proposition 3.6.4.

(ii) The reduction $\widetilde{E}$ of E mod a "good" prime ideal I of $\mathcal{O}$ generally induces an injection $\mathcal{T} \hookrightarrow \widetilde{E}(\mathcal{O}/I)$, hence limiting $\mathcal{T}$.

(iii) In some cases information about isogenies defined over K can be used. Occasionally information about twists of E can be informative; see Chapter 8.

(iv) Nagell-Lutz: indications are given in §2.10.

(v) Explicit bounds on $\mathcal{T}$: in a famous theorem, Mazur determined all the possible $\mathcal{T}$ when $K = \mathbf{Q}$; building on ideas of Mazur and Kamienny, Merel [Mer96] has proved the "strong boundedness conjecture": $|\mathcal{T}| \leq c_n$ where c_n depends only on $n := [K : \mathbf{Q}]$, and not on the particular field K. Darmon [Dar96] gives an excellent overview of all this work. For details see Chapter 8.

$\boxed{\mathbf{2}}$ First, one must not forget that E's which are isogenous over K have the same r and so one should choose the "simplest" representative.

The "easy" case is when $E(K)$ contains a point of order 2. Then simple 2-descent can (often) be applied, and may in fact determine r; see §3.6.1 below. Most of what has been done for the situation when there is no 2-torsion is concentrated on the case $K = \mathbf{Q}$. For example we give Billing's upper bound for r in §3.7, which however is practicable only for $E_{/\mathbf{Q}}$ with relatively small Δ. An upper bound for r due to Mazur (Theorem 9.9 in [Maz72]) that applies to certain $E_{/\mathbf{Q}}$ possessing a rational isogeny is sometimes useful; and many more special results will be presented later.

The best general method to attempt to find r of $E(\mathbf{Q})$ is Cremona's finely tuned version of the Birch, Swinnerton-Dyer enumeration of principal homogeneous spaces (torsors); this will be described in a later chapter.

Again for $E_{/\mathbf{Q}}$, assuming the Birch, Swinnerton-Dyer conjecture and the Riemann Hypothesis for the L-function of E, Mestre has given a conjectural upper bound for r whose integral part is often equal to r (provided Δ is not terribly big). Assuming "only" the Birch, Swinnerton-Dyer conjecture, formulas of Rohrlich for the root number determine at least the parity of r in many cases. One hopes that the missing cases of the formulas will be forthcoming, but in the meantime one can evaluate the L-function at two convenient points to determine the sign of the functional equation, hence (conjecturally) the parity of r, again only for reasonably sized Δ.

$\boxed{\mathbf{3}}$ In general one hopes to converge on the value of r by obtaining "theoretical" upper bounds in step 2 coupled with lower bounds obtained by finding points that are independent mod $\mathcal{T}$, usually by the search engine. (There are Heegner points and Monsky points that are calculated rather than found, but these seem to be restricted to special $E_{/\mathbf{Q}}$ with $r = 1$.) The independence of the points $Q_1, \ldots, Q_r$ is usually checked by $\det(\langle Q_i, Q_j \rangle) > 0$; of course when $r = 1$ all we need is $Q_1 \notin \mathcal{T}$.

$\boxed{\mathbf{4}}$ Let $h_i := \hat{h}(Q_i)$ and arrange the Q_i so that $h_1 \leq h_2 \leq \cdots \leq h_r$. Typically by using the search engine and the value of U one knows all points $P \in E(K)$ for which $\hat{h}(P) \leq$ some α which allows one to conclude that the first so many of the h_i are actually the successive minima, say

$$h_1 = \mu_1, \ldots, h_q = \mu_q \quad \text{and} \quad h_i \geq \mu_{q+1} \ \text{ for } \ q+1 \leq i \leq r.$$

For example, by searching up to $\hat{h}(P) \leq h_1$ one knows that $h_1 = \mu_1$ — if any P with $0 < \hat{h}(P) < h_1$ were found, P would be taken as a new Q_1 to get an

improved H. (The old Q_1 may become a new Q_i for some $i > 1$; in any case it, along with its canonical height, are not simply discarded.) When $r = 1$ one need only check that there are no $P \in E(K)$ with $0 < \hat{h}(P) \le h_1/4$ to verify that $h_1 = \mu_1$.

When

$$q = r \le 4,$$

no further work is needed: $Q_1, \dots, Q_r$ is a Mordell-Weil basis by Minkowski's theorem (Proposition 3.5.4(c)).

However it is frequently too costly to search far enough to have $q = r$, even when $r \le 4$. When we have at least $q \ge 1$, Proposition 3.5.4(b) and $\mu_i \ge \mu_q = h_q$ for $i > q$ give us

$$[G : H]^2 \le R := \frac{\Gamma_r \det(\langle Q_i, Q_j \rangle)}{h_1 \cdots h_{q-1} h_q^{r-q+1}}.$$

We wish to determine which primes $p \le \sqrt{R}$ actually divide $[G : H]$, *i.e.,* when there is there a point $P \in E(K)$ satisfying a relation

$$[p]P = [n_1]P_1 + \cdots + [n_r]P_r + T, \qquad T \in \mathcal{T},$$

with not all $n_j \equiv 0 \bmod p$. We quote the following efficient algorithm from [Sik95] which typically eliminates most p and for the remaining p imposes conditions on the n_j.

Siksek's sieve:

Choose $P_{r+1}, \dots, P_{r+s} \in E(K)$ whose images form a basis of the $\mathbf{F}_p$-vector space $\mathcal{T}/p\mathcal{T}$. We suppose for convenience that the p-primary part of $\mathcal{T}$ is cyclic, so $s = 0$ or 1.

Let $\mathbf{n} = (n_1, \dots, n_{r+s})$ denote an element of $\mathbf{Z}^{r+s}$ and $\bar{\mathbf{n}}$ its reduction mod p in ${\mathbf{F}_p}^{r+s}$. Define

$$V_p = \left\{ \bar{\mathbf{n}} \in {\mathbf{F}_p}^{r+s} : \text{ if } \mathbf{n} \in \mathbf{Z}^{r+s} \text{ and } \mathbf{n} \equiv \bar{\mathbf{n}} \bmod p \text{ then } \sum_{j=1}^{r+s} n_j P_j \in [p]E(K) \right\}.$$

Clearly V_p is an $\mathbf{F}_p$-subspace, and the index is divisible by p iff V_p is nonzero.

Next, let v be a non-archimedean valuation of K such that

(1) the Weierstrass equation of E is v-integral and $v(\Delta) = 0$, so as in §2.5.2 reduction mod p gives an elliptic curve $\widetilde{E}$ over the residue field k_v;† and

(2) the p-primary component of $\widetilde{E}(k_v)$ is cyclic and non-trivial. Thus $\widetilde{E}(k_v) = C_{p^i m} \oplus C_n$ where $i \ge 1$, $n \ge 1$ and $p \nmid mn$, so $|\widetilde{E}(k_v)| = lp$ where $l = p^{i-1}mn$.

Let $P'_j = [l]P_j$. If $\widetilde{P'_j} = O$ for $j = 1, \dots, r+s$, then sieving by v gives no result and we select a new v. However if $\widetilde{P'_1} = \cdots = \widetilde{P'_{i-1}} = O$ and $\widetilde{P'_i} \ne O$, then

†Later we will be able to state this more flexibly as *let E have good reduction $\widetilde{E}$ at v.*

— the subgroup $[l]\widetilde{E}(k_v)$ is cyclic of order p;
— it contains all $\widetilde{P_j}$; and
— it is generated by $\widetilde{P}_i$.

Let $\widetilde{P_j} = m_j \widetilde{P}_i$. It follows that every $\bar{\mathbf{a}} \in V_p$ must satisfy the linear relation

$$\sum m_j a_j \equiv \mathrm{mod} p.$$

By testing an appropriately large number of v, either we find $r+s$ independent relations of this sort, and then p is eliminated, or the dimension of V_p appears to stabilize at some positive value. This completes the sketch of Siksek's sieve. ■

It remains to deal with p that pass through the sieve: given Q, find a solution P of $[p]P = Q$ or determine that there is no solution in $E(K)$.

When $p = 2$ the preferred method is that of Washington (Proposition 1.7.5(a′)). For $p > 2$, I know of no such magic bullet. Cremona ([Cre92]) and Siksek ([Sik95]) suggest (in the number field case) using the complex parametrization by the Weierstrass $\wp$-function; but that must await a later chapter.

3.6.1 Simple 2-descent

Let K be a field of characteristic $\neq 2$. Then on the elliptic curve

$$E : y^2 = x(x^2 + ax + b), \quad \text{where } \Delta = 16b^2(a^2 - 4b) \neq 0,$$

$T = (0, 0)$ is a point of order 2; indeed the Weierstrass equation of any elliptic curve with a point of order 2 over a field K with $\mathrm{char}\, K \neq 2$ can be put in this form by a simple transformation of variables.

If we set

$$\overline{x} = (y/x)^2, \qquad \overline{y} = y(x^2 - b)/x^2,$$

then a calculation shows that

$$\overline{y}^2 = \overline{x}(\overline{x}^2 + \overline{a}\,\overline{x} + \overline{b}), \quad \text{where } \overline{a} = -2a,\ \overline{b} = a^2 - 4b.$$

Moreover, the discriminant of this Weierstrass equation is

$$\overline{\Delta} = 16\overline{b}^2(\overline{a}^2 - 4\overline{b}) = 16^2 b(a^2 - 4b)^2 \neq 0.$$

Let us denote this elliptic curve $\overline{E}$. Also recall the notation of Proposition 3.2.1 for $e_1 = 0$: $\phi_1 : E(K) \longrightarrow \Gamma_1 = K^*/K^{*2}$, where

$$\phi_1(O) = 1K^{*2}, \quad \phi_1(T) = e_2 e_3 K^{*2} = bK^{*2},$$

otherwise $$\phi_1(x, y) = xK^{*2}.$$

Similarly we have the homomorphism $\overline{\phi_1} : \overline{E}(K) \longrightarrow K^*/K^{*2}$.

Lemma 3.6.1 *With the above notation, define* $\alpha : E(K) \longrightarrow \overline{E}(K)$ *by*

$$\alpha(O) = \alpha(T) = O,$$

and for any other $(x, y) \in E(K)$*, by*

$$\alpha(x, y) = \left((y/x)^2, y(x^2 - b)/x^2\right).$$

Then α *is a group homomorphism with* $\ker \alpha = \{O, T\}$ *and* $\operatorname{im} \alpha = \ker \overline{\phi_1}$.

Remark. In Chapter 6 we will see all of this in a broader context: "$\overline{E}$ is E divided by the subgroup $\{O, T\}$, and $\alpha : E \longrightarrow \overline{E}$ is a 2-isogeny", the cardinal 2 referring to the size of $\ker \alpha$. Then the proof that α is a homomorphism will be subsumed in a more general result.

Proof. Let $P_i \in E(K)$, $i = 1, 2, 3$, be such that

$$P_1 + P_2 + P_3 = O,$$

and define $S = \alpha(P_1) + \alpha(P_2) + \alpha(P_3)$. We must prove that $S = O$.

First suppose that none of the P_i is O or T, equivalently no $\alpha(P_i) = O$, and that the P_i are distinct; we will refer to this as *the general case*. The line L in in $\mathbf{P}^2(K)$ containing the three P_i does not contain O as a fourth point, hence the equation of L has the form $Y = \lambda X + \nu Z$. Since also T is not a fourth point on L, therefore $\nu \neq 0$. One finds (or cribs from [Sil-Ta92, p.81]) that the three points $\alpha(P_i)$ lie on the line $Y = \overline{\lambda} X + \overline{\nu} Z$ where

$$\overline{\lambda} = \frac{\lambda\nu - b}{\nu}, \quad \overline{\nu} = \frac{\nu^2 - a\lambda\nu + b\lambda^2}{\nu}.$$

This proves $S = 0$ in the general case when the three points $\alpha(P_i)$ are distinct.

Still in the general case, suppose that $\alpha(P_1) = \alpha(P_2)$. Since $\alpha(P_3) \neq O$, therefore $\alpha(P_1) \notin \overline{E}(K)[2]$. This implies that $P_1 \notin E(K)[2]$ because of the rule

$$\alpha(-P) = -\alpha(P), \tag{1}$$

which follows from the definition of α and $-(x, y) = (x, -y)$. For $i = 1, 2$ replace P_i by an infinitesimal shift by t_i (see §2.5.1), where t_1, t_2 are independent transcendentals, and define $P_3' = -P_1' - P_2'$, $S' = \alpha(P_1') + \alpha(P_2') + \alpha(P_3')$. Since $y(P_1) \neq 0$, the definition of α shows that the three $\alpha(P_i')$ are distinct; for $[2]\alpha(P_1) = -\alpha(P_2)$, for instance, would imply the equality of non-constant power series in different variables. Thus $S' = 0$. The definition of α allows us to substitute $t_1 = 0$ and $t_2 = 0$ to obtain $S = O$.

There remain the cases where one or more of the $P_i \in \{O, T\}$. The result is clear by (1) when some $P_i = O$. Thus assume, say, $P_3 = T$. Suppose first $P_1 \notin E(K)[2]$. Let P_1' be an infinitesimal shift of P_1, and define $P_2' = P_2$, $P_3' = -P_1' - P_2'$, and S' as before. Then none of $P_i' \in \{O, T\}$, so by what we

have already proved, $S' = 0$, and again substituting 0 for the infinitesimal gives the result. The final case is when the P_i are the three points of order 2. Then $\alpha(P_1) = \alpha(P_2) = \overline{T}$, $\alpha(P_3) = \alpha(T) = O$, and again $S = O$.

This completes the proof that α is a homomorphism.

From the definition, $\ker \alpha = \{O, T\}$. To prove $\operatorname{im} \alpha = \ker \overline{\phi_1}$, we must show that

(i) $\overline{T} = (0,0) \in \operatorname{im} \alpha \Leftrightarrow \overline{b} \in K^{*2}$, and

(ii) for $(\overline{x}, \overline{y}) \in \overline{E}(K)$ with $\overline{x} \neq 0$, $(\overline{x}, \overline{y}) \in \operatorname{im} \alpha \Leftrightarrow \overline{x} \in K^{*2}$.

Proof of (i): Since $\alpha(O) = \alpha(T) = O \neq \overline{T}$,

$$\begin{aligned}(0,0) = \big((y/x)^2, y(x^2-b)/x^2\big) \quad &\Leftrightarrow \quad x \neq 0 \ \text{ and } \ y = 0 \\ &\Leftrightarrow \quad \exists \ \text{ a root } \ x \in K \ \text{ of } \ x^2 + ax + b \\ &\Leftrightarrow \quad \text{the discriminant } \ a^2 - 4b = \overline{b} \in K^{*2}.\end{aligned}$$

Proof of (ii): Here we follow [Sil-Ta92, p.84]. If $P = (x,y) \in E(K) - \{O, T\}$, then $\overline{\phi_1}\alpha(x,y) = (y/x)^2 K^{*2} = 1K^{*2}$. Conversely if $(\overline{x}, \overline{y}) \in \overline{E}(K)$ with $\overline{x} = w^2 \neq 0$, then for $i = 1, 2$ define

$$x_i = \frac{1}{2}\left(\overline{x} - a + (-1)^i \frac{\overline{y}}{w}\right), \qquad y_i = (-1)^i w x.$$

First we must verify that $(x_i, y_i) \in E(K)$. Using $\overline{y}^2 = \overline{x}(\overline{x}^2 + \overline{a}\,\overline{x} + \overline{b})$, one calculates $x_1 x_2 = b$, hence $x_i \neq 0$. What needs to be verified is

$$\frac{y_i^2}{x_i^2} = x_i + a + \frac{b}{x_i} \quad , i.e., \quad \overline{x} = x_1 + a + x_2,$$

and that is now obvious.

Again using $x_1 x_2 = b$, we calculate

$$\begin{aligned}\alpha(x_i, y_i) &= ((y_i/x_i)^2, y_i(x_i^2 - b)/x_i^2) \\ &= (w^2, w(x_2 - x_1)) = (\overline{x}, \overline{y}).\end{aligned}$$

■

Since $\overline{E}$ has the same type of Weierstrass equation as E, we can apply the same procedure: we have a homomorphism $\overline{\alpha} : \overline{E}(K) \longrightarrow \overline{\overline{E}}(K)$ where

$$\overline{\overline{E}} : \ \overline{\overline{y}}^2 = \overline{\overline{x}}\left(\overline{\overline{x}}^2 + \overline{\overline{a}}\,\overline{\overline{x}} + \overline{\overline{b}}\right),$$

$$\overline{\overline{a}} = -2\overline{a} = 4a, \qquad \overline{\overline{b}} = \overline{a}^2 - 4\overline{b} = 16b.$$

There is an obvious isomorphism $\tau : \overline{\overline{E}}(K) \longrightarrow E(K)$ given by $\tau\left(\overline{\overline{x}}, \overline{\overline{y}}\right) = (x, y) = \left(\overline{\overline{x}}/4, \overline{\overline{y}}/8\right)$. Composing this with $\overline{\alpha}$ gives a homomorphism

$$\beta = \tau\overline{\alpha} : \overline{E}(K) \longrightarrow E(K)$$

where $\beta(O) = \beta(\overline{T}) = O$, and otherwise

$$\beta(\overline{x}, \overline{y}) = \left(\frac{\overline{y}^2}{4\overline{x}^2}, \frac{\overline{y}\left(\overline{x}^2 - \overline{b}\right)}{8\overline{x}^2} \right).$$

Noting that the factor 1/4 is a square, the results of the previous lemma apply equally well to β:

$$\ker \beta = \{O, \overline{T}\}, \qquad \operatorname{im} \beta = \ker \phi_1.$$

Thus $\beta\alpha$ and $\alpha\beta$ are endomorphisms of the groups $E(K)$ and $\overline{E}(K)$ respectively.

Lemma 3.6.2 *Again with* $\operatorname{char} K \neq 2$ *and the above notation,*

$$\beta\alpha = [2]_E, \qquad \alpha\beta = [2]_{\overline{E}}.$$

Proof. First let us prove $\beta\alpha(P) = [2]P$, $\forall P \in E(K)$. This is clear if $P = O$ or T. If $P = (x, 0)$, $x \neq 0$, is another point of order 2, then $\beta\alpha(P) = \beta(0,0) = O = [2]P$.

Thus we can assume that both x and y are nonzero in $P = (x, y)$. By Corollary 1.7.2 and with a little help from the computer one now verifies that $[2]P$ and $\beta\alpha(P)$ both coincide with

$$\left(\frac{(x^2 - b)^2}{4y^2}, \frac{(x^2 - b)(x^4 + 2ax^3 + 6bx^2 + 2abx + b^2)}{8y^3} \right).$$

Second, $\alpha\beta(\overline{P}) = [2]\overline{P}$, $\forall \overline{P} \in \overline{E}(K)$ can be proved by a similar calculation, or, as pointed out in [Sil-Ta, p.82], we can argue as follows. From the description of $\operatorname{im} \alpha$ in the previous lemma, $\exists\, P \in E(K')$, where K' is at most a quadratic extension of K, such that $\alpha(P) = \overline{P}$. Then

$$\alpha\beta(\overline{P}) = \alpha\beta\alpha(P) = \alpha([2]P) = [2]\alpha(P) = [2]\overline{P}. \qquad \blacksquare$$

Proposition 3.6.3 *Let K be a field of characteristic $\neq 2$; let*

$$\begin{aligned} E &: \quad y^2 = x(x^2 + ax + b), \quad \Delta = 16b^2(a^2 - 4b) \neq 0, \\ \overline{E} &: \quad \overline{y}^2 = \overline{x}(\overline{x}^2 + \overline{a}\,\overline{x} + \overline{b}), \quad \overline{a} = -2a,\ \overline{b} = a^2 - 4b \end{aligned}$$

be elliptic curves defined over K; let $\phi_1 : E(K) \longrightarrow K^*/K^{*2}$ and $\overline{\phi_1} : \overline{E}(K) \longrightarrow K^*/K^{*2}$ be the homorphisms of the weak Mordell-Weil theorem associated to the points $T = (0,0)$ and $\overline{T} = (0,0)$ of order 2; and suppose that $E(K)$ is finitely generated with rank r. Then
(a) *$\overline{E}(K)$ is also finitely generated, with the same rank r, and the orders of the torsion subgroups are equal within a factor of 2: $|\mathcal{T}_E| = 2^i |\mathcal{T}_{\overline{E}}|$ where $i \in \{-1, 0, 1\}$.*

(b) $$2^r = \frac{|\operatorname{im} \phi_1||\operatorname{im} \overline{\phi_1}|}{4}.$$

Remarks. It will be convenient, especially for the examples, to have briefer notation for the terms in the numerator of the formula:

For the rest of this chapter, let φ and ϑ denote $\operatorname{im}\phi_1$ and $\operatorname{im}\overline{\phi_1}$ respectively.

Thus φ and ϑ are subgroups of K^*/K^{*2}, and the formula for 2^r implies that $|\varphi||\vartheta| \geq 4$. Since $\phi_1(T) = bK^{*2}$, therefore φ contains at least $\{1, b\}$ (the elements represent cosets of K^{*2}); however when b is a square this collapses to $\{1\}$. We will see examples of $\varphi = \{1\}$, and then it must be that $|\vartheta| \geq 4$. Similarly $\vartheta \supseteq \{1, \bar{b}\}$, where the latter may collapse to $\{1\}$.

Determining or estimating the rank by means of this formula in conjunction with some method of estimating $|\varphi|$ and $|\vartheta|$, we will refer to as **simple 2-descent**. **CAUTION:** this is non-standard; 2-*descent* usually refers to the homomorphism $[2] : E \longrightarrow E$ and our 2-descent would be called α-descent, with α as in Lemma 3.6.1.

Proof. We use α and β as described earlier in this section. Since $\beta : \overline{E}(K) \longrightarrow E(K)$ with kernel of order 2, therefore $\overline{E}(K)$ is finitely generated with rank $\overline{r} \leq r$, and with $|\mathcal{T}_{\overline{E}}| \leq 2|\mathcal{T}_E|$. The opposite inequalities provided by α prove that $\overline{r} = r$ and that $|\mathcal{T}_E|/|\mathcal{T}_{\overline{E}}|$ is 1/2, 1 or 2.

We can write $E(K)$ as

$$\mathcal{T}_2 \oplus \mathcal{T}_{\text{odd}} \oplus \mathbf{Z}^r$$

where $\mathcal{T}_2$ is the 2-primary part of the torsion subgroup. By Corollary 1.7.7, if there is only one point of order 2, *i.e.,* if $\bar{b} = a^2 - 4b$ is not a square, then $\mathcal{T}_2 = C_{2^n}$ for some $n > 0$, while if there are three points of order 2, *i.e.,* if $\bar{b}$ is a square, then $\mathcal{T}_2 = C_{2^n} \oplus C_{2^m}$ for appropriate n, m. Thus by elementary group theory, the index $[E(K) : [2]E(K)] = 2^{r+t}$ where $2^t = |E(K)[2]|$.

By Lemma 3.6.2, we have the subgroup inclusions

$$E(K) \supseteq \beta\overline{E}(K) \supseteq [2]E(K) = \beta\alpha E(K),$$

hence
$$2^{r+t} = [E(K) : \beta\overline{E}(K)][\beta\overline{E}(K) : \beta\alpha E(K)].$$

By the "isomorphism theorems",

$$\begin{aligned}
\frac{\beta\overline{E}(K)}{\beta\alpha E(K)} &\approx \frac{\overline{E}(K)}{\alpha E(K) + \ker\beta} \\
&\approx \frac{\overline{E}(K)/\alpha E(K)}{[\alpha E(K) + \ker\beta]/\alpha E(K)} \\
&\approx \frac{\overline{E}(K)/\alpha E(K)}{\ker\beta/[\ker\beta \cap \alpha E(K)]},
\end{aligned}$$

hence
$$2^{r+t} = \frac{[E(K) : \beta\overline{E}(K)][\overline{E}(K) : \alpha E(K)]}{[\ker\beta : \ker\beta \cap \alpha E(K)]}.$$

Since $\operatorname{im}\alpha = \ker\overline{\phi_1}$, therefore $[\overline{E}(K) : \alpha E(K)] = |\vartheta|$, and similarly $[E(K) : \beta\overline{E}(K)] = |\varphi|$. Thus it remains to verify

$$[\ker\beta : \ker\beta \cap \alpha E(K)] = 2^{2-t} = \begin{cases} 2 & \text{if } \overline{b} \text{ is not a square,} \\ 1 & \text{if } \overline{b} \text{ is a square.} \end{cases}$$

This follows from the fact that when $\overline{b} = d^2$, then $E(K)[2]$ contains the points $((-a \pm d)/2, 0)$ which map under α to $\overline{T}$, and when $\overline{b}$ is not a square, $\overline{T} \notin \operatorname{im}\alpha$. ■

3.6.2 Simple 2-descent over UFD's

In this section, R denotes a UFD of characteristic $\neq 2$. We suppose that a set $\{\pi\}$ of irreducible elements has been chosen, so that unique factorization takes the form $z = u\prod \pi^{v_\pi(z)}$, $u \in R^*$, and $\gcd(z, z') = \prod \pi^{\min\{v_\pi(z), v_\pi(z')\}}$ is uniquely defined.

We consider E and $\overline{E}$ as in the previous section with $a, b \in R$. As we will see in the next proposition, φ, and analogously ϑ, has a simple description in terms of quartic diophantine equations of the form

$$N^2 = b_1 M^4 + aM^2e^2 + b_2e^4, \quad \text{where } \ b_1, b_2 \in R,\ b_1b_2 = b, \tag{1}$$

where a solution $N, M, e \in R$ is sought with $e \neq 0$. Such a solution corresponds to a rational point $(u, v) = (M/e, N/e^2)$ on the curve

$$v^2 = b_1u^4 + au^2 + b_2. \tag{1'}$$

This curve is an example of a **torsor**.† The torsor $(1')$ is termed **elliptic** when it has a rational point, *i.e.,* a point (u, v) with u, v in the quotient field K of R, or when $b_1 \in K^{*2}$. (The latter corresponds to the rational point $(u', v') = (0, \sqrt{b_1})$ where $u' = 1/u$, $v' = v/u^2$; see *rational places* in Chapter 6.)

The torsor $(1')$ is obviously elliptic when b_2 is a square, say $b_2 = q^2$: it contains the rational point $(u, v) = (0, q)$. Applying Proposition 1.2.1 and replacing x by $x - a$ we obtain $\overline{E}$! The birational transformations are

$$x = \left[au^2 + 2q(v + q)\right]/u^2, \quad y = 2q\left[au^2 + 2q(v + q)\right]/u^3,$$

$$u = 2qx/y, \qquad v = qx\left[x^2 - \overline{b}\right]/y^2.$$

Using the final statement in Proposition 1.2.1, and remembering that we replaced x by $x - a$, we see that in this birational correspondence the point $(0, -q)$ on the torsor corresponds to the point $T = (0, 0)$ in $\overline{E}(K)$, while $(0, q)$ corresponds to O.

†Another term is *principal homogeneous space.* General definitions can only be given later.

In Chapter 6 we will prove that the torsor $(1')$ is elliptic iff it is birationally equivalent to $\overline{E}$; it would be awkward to prove this generalization right now.

Here is an application, where r denotes the rank of $E(K)$ (and of $\overline{E}(K)$ by the previous proposition):

$$\begin{aligned} r > 0 \quad &\Leftrightarrow \quad \overline{E}(K) \ \text{ is infinite} \\ &\Leftrightarrow \quad v^2 = bu^4 + au^2 + 1 \ \text{ has infinitely many points defined over } K. \end{aligned}$$

By **infinite descent** we will understand the technique of proving $r = 0$ by showing, using whatever means, that this particular torsor has only finitely many points. In a typical application when $R = \mathbf{Z}$, one shows that the only solution in non-negative integers N, M, e of $N^2 = bM^4 + aM^2e^2 + e^4$ with $e > 0$ and $\gcd(M, e) = 1$ is $M = 0$, $N = e = 1$. For example, Mordell [Mor67] shows this is so for $N^2 = pM^4 + e^4$ where p is any prime $\equiv 1 \bmod 8$ for which $x^4 \equiv 2 \bmod p$ has no solution ($p = 17, 41, 97, \ldots$). We prove this in a different way in example 4 below. (For yet another treatment see [Sil86, p.317].)

When we say, *e.g.*, that φ consists of certain elements $p, q, \ldots$ of K^*, we mean "modulo square factors", since the actual elements of φ are the cosets $pK^{*2}, qK^{*2}, \ldots$. Thus $p \in \varphi$ is short for $pK^{*2} \in \varphi$.

Proposition 3.6.4 † *Let*

$$E: \ y^2 = x(x^2 + ax + b)$$

be an elliptic curve defined over the UFD R, *where* $\operatorname{char} R \neq 2$. *Then* φ *consists of* 1 *together with those divisors* b_1 *of* b *for which*

$$N^2 = b_1M^4 + aM^2e^2 + b_2e^4, \quad \textit{where} \ \ b_1b_2 = b, \tag{1}$$

has a solution $N, M, e \in R$ *with* $e \neq 0$, *in other words, the torsor*

$$u^2 = b_1v^4 + av^2 + b_2. \tag{1'}$$

is elliptic. Then

$$\left(\frac{b_1M^2}{e^2}, \frac{b_1MN}{e^3}\right) \in E(K). \quad ‡$$

When (1) *has a solution then it has a solution satisfying*

$$e \neq 0, \quad \textit{and} \quad \gcd(M, e) = 1.$$

The same group φ *is obtained by taking those* b_1 *for which* (1) *has a solution satisfying*

$$e \neq 0, \quad \textit{and} \quad \gcd(N, M) = \gcd(N, e) = \gcd(M, e) = 1. \tag{2}$$

†We preserve the notation of Tate's 1961 lectures at Haverford College [Tat61].

‡Although this point is contributing to φ, and therefore is contributing in some sense to r, nevertheless it does not necessarily have infinite order: it may be in $\mathcal{T} - [2]\mathcal{T}$.

Remarks. If b_1 and b_1' are two divisors of b related in the manner $i^2 b_1 = j^2 b_1'$, then their complementary divisors are related by $j^2 b_2 = i^2 b_2'$, and a solution N, M, e of (1) gives the solution

$$(ijN)^2 = b_1'(jM)^4 + a(jM)^2(ie)^2 + b_2'(ie)^4.$$

Thus one need only check one representative divisor of b in $b_1 K^{*2}$. However this remark no longer applies if (2) is imposed, as will be explained in the proof.

Work can also be saved using the fact that φ is a group. For example, $b_1, b_1' \in \varphi \Rightarrow b_1 b_1' \in \varphi$, and since φ contains b (take $N = e = 1$, $M = 0$), therefore $b_1 \in \varphi \Leftrightarrow b_2 \in \varphi$. Thus there is no point in considering cases where b_2 is a square, hence in practice we never have $M = 0$.

If N, M, e satisfy (1) and (2), then also

$$\gcd(M, b_2) = 1, \quad \gcd(e, b_1) = 1.$$

Of course the same remarks apply to ϑ.

Proof. Suppose $(x, y) \in E(K)$ where $x \neq 0$, so that $x \in \varphi$. Since E is defined over the UFD R, by Proposition 2.2.2, we can write

$$x = \frac{m}{e^2}, \qquad y = \frac{n}{e^3},$$

where $\gcd(m, e) = \gcd(n, e) = 1$. (If $n = 0$ then e must be a unit.) The Weierstrass equation, cleared of denominators, is

$$n^2 = m(m^2 + ame^2 + be^4). \tag{3}$$

Let $b_1 = u \gcd(m, b)$, with u a unit to be chosen in a moment, and let $m = b_1 m_1$, $b = b_1 b_2$, so that $\gcd(m_1, b_2) = 1$. Equation (3) shows that $b_1^2 | n^2$, say $n = b_1 n_1$. Hence (3) becomes

$$n_1^2 = m_1(b_1 m_1^2 + a m_1 e^2 + b_2 e^4). \tag{4}$$

Since $\gcd(m_1, b_2) = \gcd(m_1, e) = 1$, the two factors on the right in (4) have $\gcd = 1$. It follows that each of these factors is a unit times a square, and now we choose u so that $m_1 = M^2$. Then M divides n_1, say $n_1 = MN$, and (4) becomes

$$N^2 = b_1 M^4 + aM^2 e^2 + b_2 e^4. \tag{1}$$

Thus $\phi(x) = \phi(b_1 M^2/e^2) = b_1 K^{*2}$.

Conversely, if $b_1 b_2 = b$ and (1) is satisfied by $N, M, e \in R$, with $e \neq 0$, then the Weierstrass equation is satisfied by $(x, y) = (b_1 M^2/e^2, b_1 MN/e^3)$.

If $\gcd(M, e) = d$, then $d^2 | N$ and (1) is satisfied by $N' = N/d^2$, $M' = M/d$ and $e' = e/d$, and $\gcd(M', e') = 1$. Thus, if (1) has a solution with $e \neq 0$, then it has one with $e \neq 0$ and $\gcd(M, e) = 1$. But it need not have a solution also satisfying $\gcd(N, M) = 1$, for example. Now we use the fact that φ is mod squares.

Let $\gcd(N, M) = g$, so that $g^2|b_2$, say $b_2 = g^2 b_2'$, $M = gM'$, $N = gN'$:

$$N'^2 = b_1' M'^4 + aM'^2 e^2 + b_2' e^4, \quad \text{where } \ b_1' = b_1 g^2.$$

Thus the net effect is to transfer a square factor from b_2 to b_1, which has no effect on the determination of whether or not b_1 belongs to φ. In this way we can assume that $\gcd(N, M) = 1$. Similarly by transfering a square factor from b_1 to b_2 we can also assume that $\gcd(N, e) = 1$. ■

3.6.3 Examples over Q

We illustrate the previous proposition with some E defined over the UFD $\mathbf{Z}$. Actually every subring of $\mathbf{Q}$ is a PID, but when E is defined over $\mathbf{Z}$, taking R larger than $\mathbf{Z}$ introduces more units, hence more divisors of b and $\bar{b}$, hence unnecessary work.

Here are two simple deductions from the previous proposition.

Corollary 3.6.5 *Let r denote the rank of $E(\mathbf{Q})$ where*

$$E : y^2 = x(x^2 + ax + b), \quad a, b \in \mathbf{Z}, \quad \bar{b} = a^2 - 4b, \quad b\bar{b} \neq 0.$$

(a) *If either $\bar{b} < 0$, or $a \leq 0$ and $b > 0$, then φ contains only positive divisors of b; similarly if either $b < 0$, or $a \geq 0$ and $\bar{b} > 0$, then ϑ contains only positive divisors of $\bar{b}$.*

(b) *Let $\omega(b)$ denote the number of distinct primes dividing b. Then*

$$r \leq \omega(b) + \omega(\bar{b}) - 1.$$

Proof. (a) Completing the square in (1) gives

$$N^2 = b_1 \left(M^2 + ae^2/2b_1\right)^2 - \bar{b}e^4/4b_1.$$

This is not possible when $b_1 < 0$, $\bar{b} < 0$ and $e \neq 0$ since the right side is negative. Similarly the right side is negative when $b_1 < 0$, $a \leq 0$ and $b_2 < 0$.

(b) The bound $r \leq \omega(b) + \omega(\bar{b})$ is obtained from $2^r = |\varphi||\vartheta|/4$ simply by allowing for all of the $2^{\omega(b)+1}$ possible positive and negative square-free divisors b_1 of b, and similarly for $\bar{b}$. Now b and $\bar{b}$ can't both be negative since $4b + \bar{b} = a^2$. It follows from (a) that negative divisors are disallowed in at least one of φ, ϑ. ■

Example 1

$$E : y^2 = x^3 + x, \quad \Delta = -2^6, \qquad \textbf{A64}$$

$$\overline{E} : y^2 = x^3 - 4x, \quad \Delta = 2^{12}. \qquad \textbf{B64}$$

Here $b = 1$, $\bar{b} = -2^2$, hence $r = 0$ by the corollary. Nagell-Lutz (Proposition 2.10.4) informs us that the point $(0, 0)$ is the only non-zero torsion point on **A64**, and therefore

$$\mathbf{A64}(\mathbf{Q}) = \{O, (0, 0)\} = C_2,$$

our first explicit determination of a Mordell-Weil group.

For this curve $\varphi = \{1\}$, and therefore $|\varphi||\vartheta|/4 = 1$ implies that $\vartheta = \{\pm 1, \pm 2\}$. Naturally one is curious about the rational points on the corresponding torsors:

$$N^2 = -M^4 + 4e^4 : \quad e = 1,\ M = 0,\ N = 2\,;$$

$$N^2 = \pm 2M^4 \mp 2e^4 : \quad e = M = 1,\ N = 0.$$

The Fermat curve $v^2 = u^4 + 1$ is transformed, by Proposition 1.2.1 via $u = 2x/y$, $v = -1 + 2x^3/y^2$, to **B64**. Let us complete the determination of $\mathbf{B64}(\mathbf{Q})$. To begin with,

$$\mathbf{B64}(\mathbf{Q})[2] = \{O, (0,0), (2,0), (-2,0)\}.$$

By Nagell-Lutz, any other point (x, y) would have $y = \pm 2^m$, $m \leq 4$. Then $x(x^2 - 4) = 2^{2m}$, hence $x > 2$ and $x = 2^k$ where $k > 1$. But then $x^2 - 4 = 4(x^{2(k-1)} - 1)$ contains a prime > 2 — a contradiction. Thus

$$\mathbf{B64}(\mathbf{Q}) = \{O, (0,0), (2,0), (-2,0)\} = C_2 \oplus C_2.$$

Changing to projective coordinates $x = X/Z, y = Y/Z$ and $u = U/W, v = V/W$, the birational transformation from **B64** to the Fermat curve is

$$(X, Y, Z) \longmapsto (U, V, W) = (2XY, Y^2 + 8XZ, Y^2) = (2YZ, X^2 + 4Z^2, X^2 - 4Z^2).$$

Two expressions for (U, V, W) are needed to cover all points on E — the first expression is not defined at $(X, Y, Z) = (0, 0, 1)$, while the second is not defined at $(0, 1, 0)$. That the two expressions agree when they are both defined can be checked using $Y^2Z = X^3 - 4XZ^2$. We note that both points $(\pm 2, 0, 1)$ map to the singular point $(0, 1, 0)$ on the Fermat curve. Thus the points on $V^2W^2 = U^4 + W^4$ in $\mathbf{P}^2(\mathbf{Q})$ are $(0, 1, \pm 1), (0, 1, 0)$. We can state this slightly differently:

Corollary 3.6.6 (Fermat) *The only rational points $(r, s, t) \in \mathbf{Q}^3$ on the surface*

$$r^2 = s^4 + t^4$$

are $\qquad (r, s, t) = (\pm q^2, q, 0) \quad$ *and* $\quad (\pm q^2, 0, q) \quad$ *for* $\ q \in \mathbf{Q}$.

Indeed each point (u, v) on $v^2 = u^4 + 1$, namely one of $(0, \pm 1)$, gives rise to the curve of rational points (q^2v, qu, q), $q \in \mathbf{Q}$, contained in the surface. The only rational points in the surface that are not in one of these curves are those of the form $(\pm q^2, q, 0)$, $q \in \mathbf{Q}^*$.

For the direct proof by 'Fermat descent', see [Har-Wr54, Thm.226].

Example 2 The complete list of examples for which the corollary gives $r = 0$ is as follows:

— with $b = 1$: $a = 0$ (**A64**), $a = 1$ (**A48**) and $a = -1$ (**A24**);

— $(a, b) = (\pm 6, 1)$ and $(\pm 3, 2)$ in effect repeat the example $(0, -1)$ given in the next group: **A32** appears, in one guise or another, as E or $\overline{E}$;

— with $b = -1$: any a such that $a^2 + 4 = p^n$ is a prime power, *viz.* $a = 0, \pm 1, \pm 2, \pm 3, \pm 5, \pm 7, \pm 11, \pm 13, \pm 15, \pm 17, \pm 27, \ldots$. The only cases of $n > 1$ that I know of are $0^2 + 4 = 2^2$, $2^2 + 4 = 2^3$ and $11^2 + 4 = 5^3$.

Example 3 The first curve in the catalog in [AntIV] or [Cre92] with a point of order 2 is

$$y^2 + xy + y = x^3 - x, \quad \Delta = -28. \qquad \textbf{A14}$$

In fact, Nagell-Lutz quickly confirms that the torsion subgroup is

$$\mathcal{T} = \{O, (-1, 0; 2), (0, -1; 3), (0, 0; 3), (1, -2; 6), (1, 0; 6)\} = C_6,$$

where the notation is $(x, y; n)$ for a point (x, y) of order n.

If we replace x, y by $(x - 4)/4, (y - x)/8$, the Weierstrass equation assumes the 'a-b' form we have been using:

$$E : y^2 = x(x^2 - 11x + 32), \quad a = -11,\ b = 32,$$

$$\overline{E} : y^2 = x(x^2 + 22x - 7), \quad \overline{a} = 22,\ \overline{b} = -7.$$

Since $32 \equiv 2$ mod squares, by part (a) of the corollary, $\varphi = \{1, 2\}$.

For ϑ the starting position is

$$\{1, -7\} \subseteq \vartheta \subseteq \{1, -7, -1, 7\},$$

and we must determine whether

$$N^2 = -M^4 + 22M^2e^2 + 7e^4 \qquad (*)$$

has a solution.

We don't notice any solutions with small values of the variables, so we attempt to prove that there is no solution. An obviously necessary condition for a solution to exist in that there be a solution in every field containing the quotient field of R. This time $\mathbf{Q} \hookrightarrow \mathbf{R}$ is of no help, so we consider $\mathbf{Q} \hookrightarrow \mathbf{Q}_p$ for appropriate p and seek solutions $N, M, e \in \mathbf{Z}_p$ with $e \neq 0$. As will be explained in the next section, the only p that can possibly give information in this example are 2 and 7. The latter does not help, *i.e.,* $(*)$ has a solution in $\mathbf{Z}_7$: choose $N = 0$ and $e = 1$ so that $(*)$ becomes the equation $f(M) = -M^4 + 22M^2 + 7 = 0$. Then $v_7(f(1)) = v_7(28) > 0$ and $v_7(f'(1)) = v_7(40) = 0$, so Hensel's lemma (Proposition 2.4.1) can proceed from $M = 1$.

However $(*)$ has no solution in $\mathbf{Z}_2$ with $e \neq 0$. For suppose N, M, e is such a solution; we can assume that $\gcd(M, e) = 1$, *i.e.,* M and e are not both multiples of 2. Then, by looking at $(*)$ mod 8, M and e must be odd, *i.e.,* $M = 2M_1 + 1$ and $e = 2e_1 + 1$ for some $M_1, e_1 \in \mathbf{Z}_2$, and then $N = 2N_1$. The equation takes the form $N_1^2 = 3 + 4i$ for some $i \in \mathbf{Z}_2$, which is impossible.

Conclusion: $|\vartheta| = 2$, $r = 0$ and therefore

$$\mathbf{A14}(\mathbf{Q}) = C_6.$$

Example 4 The first curve in the catalog with $r > 0$ and nontrivial torsion is

$$y^2 + xy = x^3 - x, \quad \Delta = 65. \qquad \mathbf{A65}$$

Nagell-Lutz informs us that

$$\mathcal{T} = \{O, (0,0)\} = C_2.$$

The substitutions $x, y \mapsto x/4, (y-x)/8$ yield

$$E : y^2 = x(x^2 + x - 16), \quad a = 1,\ b = -16,$$

$$\overline{E} : y^2 = x(x^2 - 2x + 65), \quad \overline{a} = -2,\ \overline{b} = 65.$$

Incidentally, substituting $x, y \mapsto 4x + 1, 4x + 8y$ in the equation for $\overline{E}$ gives

$$y^2 + xy = x^3 + 4x + 1, \quad \Delta = -65^2, \quad \mathcal{T} = \{O, (-1/4, 1/8)\}. \qquad \mathbf{B65}$$

$\{\pm 1\} \subseteq \varphi \subseteq \{\pm 1, \pm 2\}$ leads us to consider

$$N^2 = 2M^4 + M^2e^2 - 8e^4.$$

This has no solution with $e \neq 0$ in $\mathbf{Q}_5$ or in $\mathbf{Q}_{13}$ because the quadratic form

$$z^2 = 2x^2 + xy - 8y^2$$

has no non-trivial solution in either of these fields. It is time to remind ourselves how such statements concerning quadratic forms can be determined effortlessly.

We interrupt the sequence of examples to recall some basic facts about quadratic forms over global and local fields.

3.6.4 The Hilbert norm residue symbol

In this section K denotes a global field of characteristic $\neq 2$ and K_v its completion at $v \in \mathcal{M}(K)$. We recall the basic facts concerning the (quadratic) norm residue symbol $(s,t)_v$, defined for all $v \in \mathcal{M}(K)$ and $s, t \in K_v^*$. These facts are taken from the excellent sequence of exercises at the end of [Cas-Fr67].

- $(s,t)_v = \pm 1$.
- $(s,t)_v = 1$ except possibly when v is archimedean or when one of $v(2)$, $v(s)$, $v(t)$ is non-zero.

- $(s,t)_v = 1$ iff $z^2 = sx^2 + ty^2$ has a non-trivial solution $x, y, z \in K_v$; non-trivial means that at least one of x, y is non-zero.

- $(s,t)_v = (t,s)_v$.

- $(s,tt')_v = (s,t)_v(s,t')_v$, hence $(s,tt'^2)_v = (s,t)_v$, and therefore $(s,t)_v$ can be described by a finite table of values since K_v^*/K_v^{*2} is finite of order $4n$

$$\text{where} \qquad n = \begin{cases} 1 & \text{if } v(2) = 0, \\ 1/2 & \text{if } v \text{ is real archimedean}, \\ 1/4 & \text{if } v \text{ is complex archimedean}, \\ \#\,\text{residue field} & \text{otherwise.} \end{cases}$$

- $(s,-s)_v = 1$.

- $(s,t)_v = 1$ if $s + t \in K^{*2}$, in particular, $(s, 1-s)_v = 1$ for $s \neq 1$.

- If v is archimedean, then $(s,t)_v = 1$ except when v is real and s and t are both negative in the embedding $v : K \hookrightarrow K_v = \mathbf{R}$.

- *Product formula:* $\prod_v (s,t)_v = 1$; hence the number of v such that $(s,t)_v = -1$ is even. This formula embraces the law of quadratic reciprocity including the two supplementary laws.

- *Hasse principle for quadratic forms:* A non-degenerate quadratic form $f(x_1, \ldots, x_n)$ over K represents 0 over K, *i.e.,* there exist $c_1, \ldots, c_n \in K$ not all 0 such that $f(c_1, \ldots, c_n) = 0$ iff f represents 0 over K_v for all $v \in \mathcal{M}(K)$. In particular, if $s, t \in K^*$ then $sx^2 + ty^2 - z^2$ represents 0 over K iff $(s,t)_v = 1\, \forall v \in \mathcal{M}(K)$.

The values of the symbol in the case $K = \mathbf{Q}$ are as follows. We write $(s,t)_p$ in place of $(s,t)_{v_p}$.

$(s,t)_\infty = 1$ except when $s < 0$ and $t < 0$.

From the facts listed above, we can assume that s and t are square-free integers. When $p = 2$, two useful rules are

- when s is odd and $s \equiv s' \bmod 8$ then $(s,t)_2 = (s',t)_2$;

- (any $u \in \mathbf{Z}$) $(1 + 4u, t)_2 = (-1)^{uv_2(t)}$.

With these facts one can evaluate any $(s,t)_p$ for $s, t \in \mathbf{Q}^*$ using the following tables (*cf.* [Tat61]).

For an odd prime p, let $\overline{p} = (-1)^{(p-1)/2} =$ the Legendre symbol $(-1/p)$, and let q, q' (resp. g, g') denote any quadratic residues (resp. quadratic non-residues) $\bmod p$.

$s\backslash t$	q	g	qp	gp
q'	$+$	$+$	$+$	$+$
g'	$+$	$+$	$-$	$-$
$q'p$	$+$	$-$	$\overline{p}$	$-\overline{p}$
$g'p$	$+$	$-$	$-\overline{p}$	$\overline{p}$

For $p = 2$ the table is

$s\backslash t$	1	-3	3	-1	2	-6	6	-2
1	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
-3	$+$	$+$	$+$	$+$	$-$	$-$	$-$	$-$
3	$+$	$+$	$-$	$-$	$-$	$-$	$+$	$+$
-1	$+$	$+$	$-$	$-$	$+$	$+$	$-$	$-$
2	$+$	$-$	$-$	$+$	$+$	$-$	$-$	$+$
-6	$+$	$-$	$-$	$+$	$-$	$+$	$+$	$-$
6	$+$	$-$	$+$	$-$	$-$	$+$	$-$	$+$
-2	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$

Proposition 3.6.7 *Let K be a global field of characteristic $\neq 2$ and let $b_1, a, b_2 \in K$ with b_1, b_2 and $\overline{b} = a^2 - 4b_1b_2$ all non-zero. In order that*

$$N^2 = b_1M^4 + aM^2e^2 + b_2e^4 \tag{$\sharp$}$$

have a solution $N, M, e \in K$ with $e \neq 0$, it is necessary that for all $v \in \mathcal{M}(K)$ that are archimedean or for which at least one of $v(2)$, $v(b_1)$, $v(\overline{b})$ is non-zero, the norm residue symbol

$$(b_1, \overline{b})_v = 1.$$

The number of v for which $(b_1, \overline{b})_v = -1$ is even.[†]

In terms of presently used notation we have

$$b_1 \in \varphi \Longrightarrow (b_1, \overline{b})_v = 1,$$

and similarly

$$\overline{b_1} \in \vartheta \Longrightarrow (\overline{b_1}, b)_v = 1.$$

Remarks. The condition $(b_1, \overline{b})_\infty = 1$ in the case $K = \mathbf{Q}$, *i.e.*, $b_1 > 0$ when $\overline{b} < 0$, was already noted in Corollary 3.6.5.

For the v omitted in the proposition, namely those non-archimedean $v \in \mathcal{M}(K)$ for which $v(2) = v(b_1) = v(\overline{b}) = 0$, the norm residue symbols $(b_1, \overline{b})_v = 1$ automatically, as follows from properties listed above. More than that, ($\sharp$) has

[†]However it must not be construed that the number of v for which ($\sharp$) has no non-trivial solution in K_v is even. The equation ($*$) in Example 3 of the previous section fails to have a solution in $\mathbf{Q}_2$ only.

a solution in K_v for such v. This follows from the theorem of F.K. Schmidt that a curve of genus 1 over a finite field has a place of degree 1, hence is an elliptic curve, as we will explain in Chapter 6; but the proof of this stronger statement must wait until then.

Thus in practice there are three steps in attempting to determine φ (and analogously three steps for ϑ) as follows. Start with the group G_1 of divisors b_1 of b mod squares. Also, in practice, sufficient information about the norm residue symbols should be gathered so that the evaluation of the symbols is essentially a table look-up; see the tables for $\mathbf{Q}$ below.

1. A preliminary, easy sieving of G_1 by the norm residue test: $(b_1, \overline{b})_v = 1$ for the finitely many v specified in the proposition. The b_1 that survive form a subgroup G_2 since $(b_1, \overline{b})_v (b_1', \overline{b})_v = (b_1 b_1', \overline{b})_v$, a fact which can be exploited when G_1 is large.

2. For $b_1 \in G_2$, determine when there actually are solutions in K_v for all v (*i.e.,* for all the v in the proposition, presuming the theorem of F.K. Schmidt). This is do-able: if π denotes a uniformizer, then for successively higher n, consider all possible N, M, e not all 0 in $N^2 \equiv b_1 M^4 + \cdots \bmod \pi^n$ until either

(i) Hensel's lemma applied to one of the three variables gives a solution, or

(ii) there are no such solutions of the congruence, and therefore no solution in K_v.

And of course in practice there are shortcuts and efficiencies. It is a fact that the b_1 that survive this step form a subgroup G_3, known as the *Selmer group.* This will be a major topic in Chapter 10. However in the next section we can continue our examples without relying on unproved statements: we simply define G_3 to be the subset of G_1 for which there are solutions in K_v for all the v specified in the proposition.

3. There is no Hasse principle at work here: in general one only knows that $\varphi \subseteq G_3$ and there may very well be inequality. For $b_1 \in G_3$ one alternates between a search for a global solution, *i.e.,* a solution in K, and applying some special argument, perhaps using a finite extension of K, to prove that there is no global solution.

> *At present there is no effective algorithm known for step* 3. *This remains one of the major unsolved problems concerning elliptic curves over global fields.*

Proof. For ($\sharp$) to have a solution it is necessary that the quadratic form

$$z^2 = b_1 x^2 + axy + b_2 y^2 = b_1 (x + \frac{a}{2b_1} y)^2 - \frac{\overline{b}}{4b_1} y^2$$

have a solution $x, y \in K_v$ for each v with x and y not both 0. This implies that $(b_1, -\overline{b}/4b_1)_v = (b_1, \overline{b})_v (b_1, -b_1)_v = (b_1, \overline{b})_v = 1$. Applying this result to $\overline{b_1} \in \vartheta$, we have $(\overline{b_1}, \overline{\overline{b}})_v = (\overline{b_1}, 16b)_v = (\overline{b_1}, b)_v = 1$. ■

3.6.5 Continuation of examples over Q

In example 3 above, $(*)$ passes the norm residue test $(-1,32)_2 = (-1,2)_2 = 1$, yet $(*)$ has no solution in $\mathbf{Z}_2$. The proof was quite easy. In example 6 below there occurs a similar situation, but the proof that there is no solution in $\mathbf{Z}_p$ is considerably more subtle.

Example 4 (continued). $(2,65)_p$ is 1 when p is ∞ or 2, but is -1 when p is 5 or 13. Thus $\varphi = \{\pm 1\}$.

We must remember when we switch to ϑ, the norm residue criteria take the form $(\overline{b_1}, b)_v = 1$ where $\overline{b_1}|\overline{b}$. Since $b < 0$, all $\overline{b_1} < 0$ are eliminated, and since $(5,-1)_p = 1\,\forall p$, we are left to consider

$$\{1,65\} \subseteq \vartheta \subseteq \{1,5,13,65\}.$$

We spot the solution $N = 4$, $M = e = 1$ of

$$N^2 = 5M^4 - 2M^2e^2 + 13e^4,$$

and so $\vartheta = \{1,5,13,65\}$. This proves that $\mathbf{A65(Q)}$ and $\mathbf{B65(Q)}$ have rank 1. Moreover this solution tells us that

$$(\overline{b_1}M^2/e^2, \overline{b_1}MN/e^3) = (5,20) \in \overline{E}(K),$$

and applying $\beta : \overline{E}(\mathbf{Q}) \longrightarrow E(\mathbf{Q})$ (defined just before Lemma 3.6.2),

$$\beta(5,20) = (4,-4) \in E(\mathbf{Q}).$$

The transformation equations given earlier yield

$$(1,-1) \in \mathbf{A65}(\mathbf{Q}), \quad \text{and} \quad (1,2) \in \mathbf{B65}(\mathbf{Q}).$$

Calculations as done in §3.5.1 for $\mathbf{A37}$, still taking the formula for U on faith, show that in both cases these points represent the first (and only) successive minimum. Thus

$$\mathbf{A65}(\mathbf{Q}) = \{O,(0,0)\} \oplus \mathbf{Z}(1,-1), \tag{65a}$$

$$\mathbf{B65}(\mathbf{Q}) = \{O,(-1/4,1/8)\} \oplus \mathbf{Z}(1,2). \tag{65b}$$

Actually we can give direct, elementary proofs of both (65a) and (65b) without referring to heights at all; but first we must prove a lemma.

Some terminology: when the elliptic curve E is defined over a subfield K of $\mathbf{R}$ and $\Delta > 0$, let $\lambda < \lambda' < \lambda''$ denote the x-coordinates of the 2-division points, *i.e.,* the roots of $x^3 + (b_2/4)x^2 + (b_4/2)x + b_6/4$. Then $E^o(K) = \{(x,y) \in E(K) : x \geq \lambda''\} \cup \{O\}$ is a subgroup of $E(K)$ called the **even** or **neutral component**; $\{(x,y) \in E(K) : x \leq \lambda'\}$ is called the **odd component**. When the odd component is non-empty it is a coset of $E^o(K)$ and $E^o(K)$ has index 2 in $E(K)$.

Recall that according to our earlier definition, a Mordell-Weil basis $Q_1, \ldots, Q_r$ does not contain generators of the torsion subgroup $\mathcal{T}$. In general, $E(K)$ is (internally) the direct sum of $\mathcal{T}$ and $\langle Q_1, \ldots, Q_r \rangle$. In the case $r = 1$, we also refer to Q_1 as a **free generator**.

Lemma 3.6.8 *Let the elliptic curve E be defined over $\mathbf{Z}$ and satisfy the following four conditions.*

(i) *the rank of $E(\mathbf{Q})$ is 1;*

(ii) *$E(\mathbf{Q})$ contains a point Q of infinite order such that $Q + T$ is integral for all $T \in \mathcal{T}$, in particular, Q itself is integral;*

(iii) *$\Delta > 0$;*

(iv) *the odd component contains a rational point.*

Then a free generator R is to be found among the finitely many integral points in the odd component.

Proof. Let R be a free generator. Then $[n]R = Q + T$ for some $n \in \mathbf{Z}$ and $T \in \mathcal{T}$. By assumption (ii) and Proposition 2.10.1(a), R is integral. Since $R+T'$ is a free generator for $T' \in \mathcal{T}$, all $R+T'$ are integral, and we wish to show that some $R + T'$ is in the odd component. If not, then R is in the even component and therefore also all T' are in the even component. But then $E(\mathbf{Q}) = \langle R \rangle \oplus \mathcal{T}$ is contained in the even component, contrary to (iv).

There are only finitely many integral points (x, y) in the odd component since $\lambda \leq x \leq \lambda'$. ■

A65 satisfies the four conditions with $Q = (1, 0)$, $Q+T = (-1, 1)$, where $T = (0, 0)$ has order 2. Checking the integral x between $\lambda \approx -1.13$ and $\lambda' = 0$, we find that the only integral points in the odd component are $Q+T$, $-Q+T = (-1, 0)$ and T. Thus both $\pm Q + T$ (and therefore also both $\pm Q$) are free generators.

The lemma does not apply to **B65** directly since $\Delta = -65^2 < 0$, however (65a) implies (65b) as follows. Combining the isomorphisms $\mathbf{A65}(\mathbf{Q}) \approx E(\mathbf{Q})$ and $\mathbf{B65}(\mathbf{Q}) \approx \overline{E}(\mathbf{Q})$ with α gives a homomorphism $\alpha' : \mathbf{A65}(\mathbf{Q}) \longrightarrow \mathbf{B65}(\mathbf{Q})$; $\beta' : \mathbf{B65}(\mathbf{Q}) \longrightarrow \mathbf{A65}(\mathbf{Q})$ is defined similarly, and the composition of these two homomorphisms in either order is multiplication by 2. Now $\ker(\alpha)$ is the torsion subgroup $\mathcal{T}_A$ of $\mathbf{A65}(\mathbf{Q})$, and similarly $\ker(\beta') = \mathcal{T}_B$. In

$$\mathcal{T}_A \oplus \mathbf{Z} \xrightarrow{\alpha'} \mathcal{T}_B \oplus \mathbf{Z} \xrightarrow{\beta'} \mathcal{T}_A \oplus \mathbf{Z}$$

$\mathrm{im}(\alpha') = i\mathbf{Z} \subset \mathbf{Z}$ for some positive integer i and similarly $\mathrm{im}(\beta') = j\mathbf{Z} \subset \mathbf{Z}$. From $\beta'\alpha' = [2]$, we deduce $ij = 2$. Since

$$(-1, 1) \xmapsto{\alpha'} (0, 1) \xmapsto{\beta'} (4, -10) = [2](-1, 1),$$

we use Proposition 1.7.3 to test whether $(0,1)$ is a free generator or twice a free generator of $\mathbf{B65}(\mathbf{Q})$. We find

$$(0,1) = [-2](1,2),$$

and so $(1,2)$ is a free generator.

Corollary 3.6.9 *Suppose $E_{/\mathbf{Z}}$ satisfies* (i)–(iii) *of the lemma and that the odd component contains no integral points. Then the odd component contains no rational points.*

Rohrlich mentioned the example $y^2 = x(x^2+6x+2)$ in connection with Mazur's ideas on the (real) topology of rational points. Here $\mathcal{T} = \{O,(0,0)\}$, the rank of $E(Q)$ is 1 by simple 2-descent, $Q = (1,3)$ is a point of infinite order, and $Q + (0,0) = (2,-6)$. There are no integral (x,y) with $\lambda = -3-\sqrt{7} < x < \lambda' = -3+\sqrt{7}$, and so no rational points at all on the odd component. Replacing x by $x-2$ gives the curve 896D1 in Cremona's catalog [Cre92]. Another example in that catalog is 336E5: $y^2 = x^3 - x^2 - 12544x + 544960$. We will return to this subject in a later chapter.

We mention the following simple result; however it cannot be used effectively since we have no control over denominators to limit searches.

Lemma 3.6.10 *Let E be an elliptic curve defined over $\mathbf{Q}$ with $\Delta > 0$. Then the odd component contains either no rational points, or else it contains a Mordell-Weil basis.*

Proof. Let $\mathcal{C}$ denote the set of rational points on the odd component, suppose $\mathcal{C} \neq \emptyset$, and let $Q_1,\ldots,Q_r$ be a Mordell-Weil basis. If $\mathcal{C}$ contains a torsion point T, then for each $Q_i \notin \mathcal{C}$ replace Q_i by $Q_i + T$ to obtain a Mordell-Weil basis $\subset \mathcal{C}$. Thus suppose $\mathcal{T} \subset E^o(\mathbf{Q})$. Let $[n_1]Q_1 + \cdots + [n_r]Q_r + T \in \mathcal{C}$. Subtracting T and appropriate even multiples of the Q_i (all these points are in $E^o(\mathbf{Q})$), and renumbering the Q_i, we obtain $Q_1 + \cdots + Q_k \in \mathcal{C}$ where we can suppose that k is minimal. Then $k = 1$, for otherwise $Q_1 + \cdots + Q_{k-1}$ and Q_k are in $E^O(\mathbf{Q})$ and so is their sum. Thus $Q_1 \in \mathcal{C}$, and in the basis we can replace any $Q_i \notin \mathcal{C}$ with $Q_i + Q_1$ to obtain a basis contained in $\mathcal{C}$. ■

Example 5 Euler proved, in effect, that for $E : y^2 = x^3 + 1$, $E(\mathbf{Q})$ has rank 0; his proof by infinite descent is reproduced in [Dic52, vol.II, p.533]. Nagell [Nag25], also using infinite descent, generalized this as follows. Our proof uses the norm residue criteria (Proposition 3.6.7), and it is interesting to see by comparison how efficient that approach is.

Proposition 3.6.11 *Let D be a nonzero integer and let E be the elliptic curve defined over $\mathbf{Q}$ by the equation $y^2 = x^3 + D^3$. If*

$$p \text{ prime}, p|D \Longrightarrow p = 3 \text{ or } p \equiv 5 \bmod 12,$$

then the rank of $E(\mathbf{Q})$ is 0.

Remark. As a convenience, and not because of any particular difficulty, the determination of $\mathcal{T}$ is postponed to Corollary 7.3.4. The upshot is that

$$E(\mathbf{Q}) = \mathcal{T} = \{O, (-D, 0)\} = C_2$$

except (assuming D is square-free — see the begining of the following proof) when $D = 1$ and then

$$E(\mathbf{Q}) = \mathcal{T} = \{O, (-1, 0), (0, \pm 1), (2, \pm 3)\} = C_6.$$

Proof. We can assume that D is square-free; for if $D = d^2 D_1$, then replacing x, y with $d^2 x, d^3 y$ results in the equation $y^2 = x^3 + D_1^3$.

Replacing x with $x - D$, the equation becomes

$$E : y^2 = x(x^2 - 3Dx + 3D^2),$$

which is paired with

$$\overline{E} : y^2 = x(x^2 + 6Dx - 3D^2).$$

Thus

$$\{1, 3\} \subseteq \varphi, \qquad \{1, -3\} \subseteq \vartheta,$$

and we will exclude all other divisors of $b = 3D^2$ and $\overline{b} = -3D^2$ by the norm residue tests.

Suppose b_1 is a square-free divisor of $3D$ that is a member of φ. If b_1 is divisible by a prime $p \equiv 5 \bmod 12$, then $(b_1, \overline{b})_p = (b_1, -3)_p = -1$, a contradiction. Since $\overline{b} = -3D^2 < 0$, therefore $b_1 > 0$, hence $b_1 = 1$ or 3.

Similarly a square-free divisor $\overline{b_1}$ of $3D$ that is divisible by a prime $p \equiv 5 \bmod 12$ cannot be a member of ϑ since $(\overline{b_1}, b)_p = (\overline{b_1}, 3)_p = -1$. Finally, $-1 \notin \vartheta$ since $(-1, b)_3 = (-1, 3)_3 = -1$. ■

Example 6 An integer N is a **quartic residue** $\bmod p$ if it is a fourth power $\bmod p$, *i.e.,* if $x^4 \equiv N \bmod p$ has a solution $x \in \mathbf{Z}$. Some authors use the term *biquadratic*, but *quartic* is shorter. Recall that the quadratic fields $\mathbf{Q}(i)$ $(i = \sqrt{-1})$, $\mathbf{Q}(\sqrt{2})$ and $\mathbf{Q}(\sqrt{-2})$ have class number 1, *i.e.,* the rings of integers $\mathbf{Z}[i]$, $\mathbf{Z}[\sqrt{2}]$ and $\mathbf{Z}[\sqrt{-2}]$ are PID's. Also, the fundamental unit $1 + \sqrt{2}$ of $\mathbf{Z}[\sqrt{2}]$ has norm -1; thus a prime splits in one of these fields iff it is the norm of an integer in the corresponding PID.

Lemma 3.6.12 *Suppose p is a prime $\equiv 1 \bmod 8$, so that it splits in the three fields $\mathbf{Q}(i)$, $\mathbf{Q}(\sqrt{2})$ and $\mathbf{Q}(\sqrt{-2})$, say*

$$\begin{aligned} p &= (A + 4Bi)(A - 4Bi) = A^2 + 16B^2 \\ &= (C + D\sqrt{2})(C - D\sqrt{2}) = C^2 - 2D^2, \\ &= (F + G\sqrt{-2})(F - G\sqrt{-2}) = F^2 + 2G^2, \end{aligned}$$

where $A, B, \ldots, G \in \mathbf{Z}$. Also let i and $\sqrt{2}$ denote a root in $\mathbf{F}_p$ of $x^2 + 1$ and $x^2 - 2$ respectively.

(a) *The eight elements of* $\mathbf{F}_p$, $\pm 1 \pm i$ *and* $\pm 1 \pm \sqrt{2}$, *are all either quadratic residues or quadratic nonresidues.*

(b) *The following are equivalent:*

(i) 2 *is a quartic residue* $\bmod p$;

(i′) *when* $p \equiv 1 \bmod 16 : 1+i$ *is a quadratic residue* $\bmod p$;
when $p \equiv 9 \bmod 16 : 1+i$ *is a quadratic nonresidue* $\bmod p$;

(ii) B *is even*, i.e., p *has the form* $A^2 + 64B_1^2$;

(iii) C *is a quadratic residue* $\bmod p$;

(iii′) *there are nonzero integers* U, C', D', *where* C' *is a quadratic residue* $\bmod p$, *such that*

$$pU^2 = C'^2 - 2D'^2\,;$$

(iv) F *is a quadratic residue* $\bmod p$;

(iv′) *there are nonzero integers* V, F', G', *where* F' *is a quadratic residue* $\bmod p$, *such that*

$$pV^2 = F'^2 + 2G'^2\,.$$

Proof. (a) $(1+i)(1-i) = 2$ and $(1+\sqrt{2})(1-\sqrt{2}) = -1$ are squares in $\mathbf{F}_p$. The proof of this part is completed by noting the identity from [Str-Top94,p.1143]

$$(1+\sqrt{2})(1+i) = (1+\zeta_8)^2$$

where $\zeta_8 = (1+i)/\sqrt{2}$ is an 8th root of unity.

(b) The equivalence of (i) and (ii) was conjectured by Euler and proved by Gauss (*cf.* [Cox89, p.20]). For Dirichlet's beautiful proof see [Ire-Ro82, p.64, exercise 26] and [Sil86, p.318].

To prove the equivalence of (i) and (i′), we note that $p \equiv 1 \bmod 16$ iff $\mathbf{F}_p$ contains the 16th root of unity $\sqrt{(1+i)/\sqrt{2}}$, *i.e.,* $(1+i)/\sqrt{2}$ is a quadratic residue. Thus when $p \equiv 1 \bmod 16$, either $i+i$ is a quadratic residue and 2 is a quartic residue, *i.e.,* $\sqrt{2}$ is a quadratic residue (*e.g.* when $p = 113$), or $1+i$ and $\sqrt{2}$ are both quadratic non-residues (*e.g.* when $p = 17$). When $p \equiv 9 \bmod 16$, $\mathbf{F}_p$ contains only 8th roots of unity, hence $1+i$ and $\sqrt{2}$ cannot both be quadratic residues. In fact exactly one is. For suppose $\sqrt{1+i}$ and $\sqrt[4]{2}$ are both quadratic over $\mathbf{F}_p$. Since this field has a unique quadratic extension, $\sqrt{1+i} = a + b\sqrt[4]{2}$ for some $a, b \in \mathbf{F}_p$, $b \neq 0$. Squaring shows that $2ab\sqrt[4]{2} = 0$, *i.e.,* $a = 0$ and $(1+i)/\sqrt{2} = b^2$ in $\mathbf{F}_p$ which means that $p \equiv 1 \bmod 16$. (Examples: $1+i$ is a quadratic residue mod 41; 2 is a quartic residue mod 73.)

Assume (iii′). By replacing U, C', D' by $U/d, C'/d, D'/d$ where $d = \gcd(U, C', D')$, we can assume that $d = 1$, hence $p \nmid C'D'$. The Legendre symbol $(D'/p) = 1$ for any $p \equiv 1 \bmod 8$, in particular, $(D/p) = 1$; for $(-1/p) = (2/p) =$

1, and if q is an odd prime divisor of D' then $q \nmid U$ and $p \equiv C'^2/U^2 \bmod q$, hence $1 = (p/q) = (q/p)$. †Thus D'^2 is a quartic residue mod p and therefore $2 \equiv C'^2/D'^2 \bmod p$ is a quartic residue mod p iff C' is a quadratic residue mod p. This argument applies in particular to (iii).

The proof of (iv) $\Leftrightarrow$ (i) $\Leftrightarrow$ (iv′) is entirely similar: one finds that $(G'/p) = 1$, say $G' \equiv H^2 \bmod p$, and $\mathbf{F}_p$ contains a fourth root of -1, $\zeta := (1+i)/\sqrt{2}$, so $2 \equiv F'^2/\zeta^4 H^4 \bmod p$ is a quartic residue iff F' is a quadratic residue, in particular, iff F is a quadratic residue. ■

We put the results of the next example in

Proposition 3.6.13 *Let p be a prime number.*

(a) *The rank r of $E(\mathbf{Q})$, where $E : y^2 = x^3 + px$, satisfies*

$$\begin{array}{ll} r = 0 & \textit{if } \ p = 2, \\ r = 0 & \textit{if } \ p \equiv 7, 11 \bmod 16, \\ r \le 1 & \textit{if } \ p \equiv 3, 5, 13, 15 \bmod 16, \\ r \le 2 & \textit{if } \ p \equiv 1, 9 \bmod 16, \text{i.e., } p \equiv 1 \bmod 8. \end{array}$$

In the last case, $r = 0$ if 2 is not a quartic residue mod p.

(b) *The rank r of $E(\mathbf{Q})$, where $E : y^2 = x^3 - px$, satisfies*

$$\begin{array}{ll} r = 1 & \textit{if } \ p = 2, \\ r = 0 & \textit{if } \ p \equiv 3, 11, 13 \bmod 16, \\ r \le 1 & \textit{if } \ p \equiv 5, 7, 9, 15 \bmod 16, \\ r \le 2 & \textit{if } \ p \equiv 1 \bmod 16. \end{array}$$

Remarks. We do not pause to determine the torsion subgroup. Actually this subgroup is $O, (0,0)$ for all the $E(\mathbf{Q})$ and $\overline{E}(\mathbf{Q})$ involved in the proposition; a comprehensive result is given in Corollary 7.3.6.

For the proof by infinite descent of the $r = 0$ cases of part (a) of the proposition, see [Mor69, p.23].

There is a general conjecture (the **parity conjecture**: $(-1)^r = W_E$, the *root number* of E — to be defined later‡) that implies in the cases of the proposition where $r \le 1$, that actually $r = 1$; and in the cases where $r \le 2$, that $r = 0$ or 2. Mordell [Mor67, p.3] conjectured that when $p = A^2 + 64B_1^2$, then $r > 0$ for $y^2 = x^3 + px$ and so $r = 2$ by the parity conjecture; however computers have found this to be false, the first counter-example being $p = 257 = 1^2 + 64 \cdot 2^2$ for

†The idea of using quadratic reciprocity in this way seems to have originated with C. -E. Lind (*cf.* [Cas66, p.284]).

‡The term *parity conjecture* has been employed by others to mean something else. As will be explained later, the parity conjecture in the sense defined here is a consequence of the Taniyama conjecture coupled with the Birch, Swinnerton-Dyer conjecture. The E of the proposition satisfy the the Taniyama conjecture since they are CM — they have complex multiplication by i. It should also be mentioned that the parity conjecture for these E is included in an earlier conjecture of Selmer [Sel54a].

which $r = 0$. Recently Rose [Ros95] found that of the 625 primes of this form less than 50000, $r = 2$ in 367 cases, while $r = 0$ in the remaining 258 cases.

Heegner has suggested several similar examples, but I do not know the present-day status of his arguments; *cf.* [Mor67, p.4].

Bremner ([Mol89]) has carried out extensive numerical calculations for $y^2 = x^3 + px$ when $p \equiv 5 \bmod 8$.

Monsky [Mon92] has proved, among other results, that $r = 1$ for $y^2 = x^3+px$ when $p \equiv 5 \bmod 16$ by constructing a rational point using modular functions.

Proof. First consider

$$E : y^2 = x^3 + 2x, \qquad \overline{E} : y^2 = x^3 - 8x.$$

The norm residue test leaves

$$\varphi = \{1, 2\}, \qquad \{1, -2\} \subseteq \vartheta\{\pm 1, \pm 2\},$$

so we consider

$$N^2 = -M^4 + 8e^4, \quad \gcd(M, e) = 1.$$

Reduction mod 4 implies $M = 2M_1$ and $N = 4N_1$, hence $e = 2e_1$, contradicting $\gcd(M, e) = 1$. This eliminates -1 and 2 from ϑ, and $r = 0$.

In part (b) when $p = 2$, the point $P = (-1, 1)$ on $E(\mathbf{Q})$ has infinite order since $[2]P = (9/4, -21/8)$ is fractional, *cf.* Proposition 2.10.4. The initial estimate $|\varphi||\vartheta| \leq 8$ now confirms that $r = 1$.

Thus the remainder of the proof consists of finding upper bounds for r when p is an odd prime. We need to consider the following equations.

(α) $N^2 = -M^4 + pe^4$;

(β) $N^2 = -M^4 + 4pe^4$;

(γ) $N^2 = 2M^4 + 2upe^4$ ($u = 1$ or -1), or $2N_1^2 = M^4 + upe^4$ where $N = 2N_1$;

(δ) $N^2 = -2M^4 + 2pe^4$, or $2N_1^2 = -M^4 + pe^4$ where $N = 2N_1$.

In all four equations we can assume that $\gcd(M, e) = 1$. This implies that $p \nmid M$ and $p \nmid N$. Also e is odd (hence $e^4 \equiv 1 \bmod 16$) since e even $\Longrightarrow$ M odd and then none of the equations is possible mod 4.

We first gather information on (α)–(δ) by elementary congruence considerations involving the Legendre symbol (x/p) or the non-solvability of the equation in $\mathbf{Q}_2$.

(α) implies $(-1/p) = 1$, hence there is no solution when $p \equiv 3 \bmod 4$. Also there is no solution in $\mathbf{Q}_2$ when $p \equiv 13 \bmod 16$ since M and N have opposite parity and the possible values of $p \equiv N^2 + M^4 \bmod 16$ are 1,5, and 9.

(β) As in (α), there is no solution when $p \equiv 3 \bmod 4$.

(γ) implies $(2/p) = 1$, hence there is no solution when $p \equiv \pm 3 \bmod 8$. Now M must be odd and $up \equiv 2N_1^2 - 1 \equiv 1, 7$ or $15 \bmod 16$. Hence there is no solution in $\mathbf{Q}_2$ when $u = 1$, $p \equiv 9 \bmod 16$, or when $u = -1$, $p \equiv 7 \bmod 16$.

(δ) implies $(-2/p) = 1$, hence no solution when $p \equiv 5$ or $7 \bmod 8$. Also, there is no solution in $\mathbf{Q}_2$ when $p \equiv 11 \bmod 16$ since, as in (γ), $p \equiv 2N_1^2 + 1 \equiv 1, 3$, or $9 \bmod 16$.

Now let us go through the various cases of the proposition for odd p.

(a) Since $\overline{b} = -4p < 0$, therefore $-1 \notin \varphi$ and $\varphi = \{1, p\}$. We have

$$\{1, -p\} \subseteq \vartheta \subseteq \{\pm 1, \pm 2, \pm p, \pm 2p\}.$$

When $p \equiv 7$ or $11 \bmod 16$, the eliminations for (β), (γ) and (δ) leave $|\vartheta| = 2$, hence $r = 0$.

When $p \equiv 3, 5, 13, 15 \bmod 16$, the eliminations for, respectively, (β) and (γ), (γ) and (δ), (γ) and (δ), (β) and (δ) leave $|\vartheta| \leq 4$, hence $r \leq 1$ in these four cases.

That leaves $p \equiv 1$ or $9 \bmod 16$ with no eliminations, hence $|\vartheta| \leq 8$ and $r \leq 2$.

Part (b) is sorted out in a similar manner.

It remains to prove that when $p \equiv 1 \bmod 8$ if either (β) or (γ) with $u = -1$ has a solution in $\mathbf{Z}$ with $e \neq 0$ then 2 is a quartic residue $\bmod p$. It is not hard to check that both have solutions in all local fields $\mathbf{Q}_q$, so the proof must be of a different kind than we have used thus far.

In the case (γ), *i.e.*, $pe^4 = M^4 - 2N_1^2$, the conclusion is immediate from criterion (iii′) of the lemma.

To treat (β) we use the unique factorization in the PID $\mathbf{Z}[i]$. Taking (β) mod 4 shows that M and N are even, say $M = 2M_1$ and $N = 2N_1$:

(β') $N_1^2 = -4M_1^4 + pe^4$

Since e is odd, N_1 is odd, hence $N_1^2 \equiv 1 \bmod 8$, and $pe^4 \equiv p \equiv 1 \bmod 8$. Thus (β') implies $M_1 = 2M_2$, and our assumption is

$$\text{(ii}'\text{)} \qquad \exists \text{ nonzero integers } e, N_1, M_2 \text{ such that } pe^4 = N_1^2 + 64M_2^4.$$

We wish to prove (ii′)$\Longrightarrow$(ii) of the lemma.

We have $p = \pi\overline{\pi}$ where $\pi = A + 4Bi$ is irreducible. Assuming (ii′), we wish to prove that B is even. To see the idea of the proof, consider the simplest case $e = 1$. Then

$$p = \pi\overline{\pi} = \alpha\overline{\alpha} \quad \text{where} \quad \alpha = N_1 + 8M_2^2 i.$$

This implies $\pi = A + 4Bi = N_1 \pm 8M_2^2 i$, hence B is even.

In general let $e = \prod \sigma_j \overline{\sigma_j} \prod q_k$ where $\sigma_j = s_j + 2t_j i$ and $q_k \equiv 3 \bmod 4$ are irreducible; $\sigma_j = \pi$ is not excluded. The σ_j must occur in conjugate pairs since e is real. Then $N^2 \equiv -M^4 \bmod q_k$ where $q_k \nmid M$, which is not possible for $q_k \equiv 3 \bmod 4$, so in fact no q_k is present. Thus

$$\pi\overline{\pi} \prod \sigma_j^4 \overline{\sigma_j}^4 = \alpha\overline{\alpha}.$$

Since $\gcd(N_1, M_2) = 1$, $\alpha = N_1 + 8M_2^2 i$ cannot contain the prime $\sigma_j\overline{\sigma_j} = p_j$ as a factor. Thus we can choose notation so that

$$N_1 + 8M_2^2 i = \pi \prod \sigma_j^4 = (A + 4Bi) \prod (s_j + 2t_j i)^4$$

$$\Longrightarrow N_1 \equiv A + 4Bi \bmod 8 \Longrightarrow B \equiv 0 \bmod 2. \qquad \blacksquare$$

The statements of the proposition which leave the rank ambiguous can be made more explicit by bringing in relevant torsors. We give one example. When $p \equiv 5 \bmod 8$ we have

$$\{1, -p\} \subset \vartheta \subset \{\pm 1, \pm p\},$$

hence $r = 1$ iff $-1 \in \vartheta$ (equivalently $p \in \vartheta$). Re-writing the torsor $N^2 = -M^4 + 4pe^4$ yields the following.

> *If p is a prime $\equiv 5 \bmod 8$ and $E : y^2 = x^3 + px$ then the rank of $E(\mathbf{Q})$ is 1 iff*
>
> $$\exists u, v \in \mathbf{Q} \quad \textit{such that} \quad p = u^2 + 4v^4, \qquad \textit{and then}$$
>
> $$P = \left(\frac{u^2}{4v^2}, \frac{u^3}{8v^3} + uv\right) \qquad (\P)$$
>
> *is a point of infinite order.* [As remarked after the proposition, the parity conjecture implies that this is true for all $p \equiv 5 \bmod 8$, and Monsky has proved it is true for $p \equiv 5 \bmod 16$.]

Sometimes the representation $p = u^2 + 4u^4$ appears when p is written as the sum of two integer squares, *e.g.* $5 = 1^2 + 4 \cdot 1^4$; but usually one must look harder , *e.g.*

$$37 = \left(\frac{151}{25}\right)^2 + 4\left(\frac{3}{5}\right)^4.$$

The first expected example of $r = 2$ in part (a) is $p = 73 = 3^2 + 64$. Since $p = 3^2 + 4 \cdot 2^4$, (¶) gives the point

$$P = \left((3/4)^2, 3 \cdot 137/2^6\right) \in E(\mathbf{Q}).$$

However this turns out not to be one of the successive minima. A search for points, and calculations as indicated in §3.5.1, find that the successive minima are P_1, P_2 where

$$\begin{array}{ll} P_1 = (36, 222), & \hat{h}(P_1) = 3.699981, \\ P_2 = \left(73 \cdot (3/2)^2, 3 \cdot 7 \cdot 11 \cdot 73/8\right), & \hat{h}(P_2) = 4.366662, \\ P = -P_1 + P_2 + (0, 0), & \hat{h}(P) = 4.9486. \end{array}$$

Since $r < 5$, by Minkowski's result (Proposition 3.5.4(f)), P_1, P_2 is a Mordell-Weil basis.

Here is an example of rank 2 in case (b): the successive minima of $y^2 = x^3 - 17x$ over $\mathbf{Q}$ are

$$\begin{array}{ll} P_1 = (17, 68), & \hat{h}(P_1) = 1.172183, \\ P_2 = (-4, 2), & \hat{h}(P_2) = 1.755026. \end{array}$$

For the next few $p \equiv 1 \bmod 16$, those with $r = 2$ in case (b) are $p = 97, 241, 257, 337, 401, \ldots$, while those with $r = 0$ are $p = 113, 193, 353, \ldots$. I do not know any succinct criterion, such as we had in case (a), that guarantees $r = 0$.

The final proposition of this section treats an example considered by Stroeker and Top [Str-Top94]. The analysis is quite similar to that of the preceding proposition, but there are a few twists and turns, so a proof is included.

Proposition 3.6.14 *Let p be a prime number.*

(a) *The rank r of $E(\mathbf{Q})$, where $E : y^2 = x(x^2 - 2px + 2p^2)$, satisfies*

$$\begin{array}{ll} r = 0 & \textit{if } \ p = 2 \ \ \textit{or if } \ p \equiv \pm 3 \bmod 8, \\ r \le 1 & \textit{if } \ p \equiv -1 \bmod 8, \\ r \le 3 & \textit{if } \ p \equiv 1 \bmod 8. \end{array}$$

In the last case, $r \le 1$ unless $p \equiv 1 \bmod 16$ and 2 is a quartic residue $\bmod p$.

(b) *The rank r of $E(\mathbf{Q})$, where $E : y^2 = x(x^2 + 2px + 2p^2)$, satisfies*

$$\begin{array}{ll} r = 0 & \textit{if } \ p = 2, \\ r \le 1 & \textit{if } \ p \equiv \pm 3 \bmod 8, \\ r \le 2 & \textit{if } \ p \equiv \pm 1 \bmod 8, \end{array}$$

When $p \equiv 1 \bmod 8$ then $r = 0$ unless $1 + i$ is a quadratic residue $\bmod p$, *where i denotes (either) root of $u^2 = -1$ in $\mathbf{F}_p$.*

Remarks. Again the parity conjecture implies that when $r \le 1$, 2 or 3 then r is, respectively 1, 0 or 2, 1 or 3. *Cf.* [Str-Top94, 1.2 and 1.4].

$r = 3$ actually occurs in case (a); Stroeker and Top note the examples $p = 337$ and 1201. Examples of $r = 2$ in case (b) seem to be fairly common: $p = 31$, 41, 47 ....

Eventually one wants to see the Selmer groups more explicitly, as is done for the case (a) curves in [Str-Top94]; this will be done in Chapter 10.

Proof. The $p = 2$ cases are left to the reader.

Cases (a) and (b) can be treated together:

$$E : y^2 = x(x^2 - 2spx + 2p^2),$$

$$\overline{E} : y^2 = x(x^2 + 4spx - 4p^2)$$

where $s = 1$ for case (a) and $s = -1$ for case (b). For both cases,

$$\{1, 2\} \subset \varphi \subset \{1, 2, p, 2p\}, \quad \{1, -1\} \subset \vartheta \subset \{\pm 1, \pm 2, \pm p, \pm 2p\}.$$

Thus $r \le 3$ in all cases. The following facts complete the proof.

$\boxed{p \in \varphi} \Longrightarrow$

$$N^2 = pM^4 - 2spM^2e^2 + 2pe^4, \quad \textit{say } N = pN_1,$$

$$\Longrightarrow pN_1^2 = M^4 - 2sM^2e^2 + 2e^4.$$

Since we can assume that $\gcd(M, e) = 1$, this equation implies that $2 \nmid N_1$. Hence

$$p \equiv 1 - 2se^2 + 2e^4 \equiv \begin{cases} 1 \mod 8 & \text{in case (a)}, \\ 1 \mod 4 & \text{in case (b)}. \end{cases}$$

Also, $p \nmid e$ and $u^4 - 2su^2 + 2$ has a root in $\mathbf{F}_p$, *i.e.,* at least one of $s \pm i$ is a quadratic residue. When $p \equiv 1 \bmod 8$, by the lemma this is so iff $1 + i$ is a quadratic residue.

$\boxed{2 \in \vartheta} \Longrightarrow$

$$N^2 = 2M^4 + 4spM^2e^2 - 2p^2e^4, \quad \textit{say } N = 2N_1,$$

$$\Longrightarrow 2N_1^2 = M^4 + 2spM^2e^2 - p^2e^4.$$

This implies, since $\gcd(M, e) = 1$, that $2 \nmid Me$. Hence $M^4 \equiv p^2e^4 \equiv 1 \bmod 8$. From $2N_1^2 \equiv 2sp \bmod 8$ we deduce that $p \equiv 1$ (resp. 3) $\bmod 4$ in case (a) (resp. (b)).

In case (a) we can improve this to $p \equiv 1 \bmod 8$. For suppose $p \equiv 5 \bmod 8$. Then since the Legendre symbol $(2/p) = -1$, p must divide M and N_1, say $M = pM_1$, $N_1 = pN_2$, hence $2N_2^2 = p^2M_1^4 + 2pM_1^2e^2 - e^4$, where $p \nmid N_2e$. Thus $2N_2^2 \equiv -e^4 \bmod p$, contradicting $(-2/p) = -1$.

When $p \equiv 9 \bmod 16$ in case (a) we wish to prove that 2 is a quartic residue. We will derive a contradiction by assuming the contrary, which by the lemma means that $1 + i$ is a quadratic residue. Thus all of $\pm 1 \pm \sqrt{2}$ are quadratic residues. Following [Str-Top94,p.1145], we factor over $\mathbf{Z}[\sqrt{2}]$:

$$2N_1^2 = (M^2 + (1 + \sqrt{2})pe^2)(M^2 + (1 - \sqrt{2})pe^2).$$

The two factors on the right have greatest common divisor $\sqrt{2}$, hence

$$M^2 + (1 + \sqrt{2})pe^2 = \varepsilon\alpha^2\sqrt{2}$$

for $\varepsilon, \alpha \in \mathbf{Z}[\sqrt{2}]$ where ε is a unit, and so can be written in the form $(\pm 1 \pm \sqrt{2})^n$. Thus under the homomorphism $\mathbf{Z}[\sqrt{2}] \longrightarrow \mathbf{F}_p$ the equation takes the form $M^2 \equiv \beta^2\sqrt{2}$ where $\beta^2 \equiv \varepsilon\alpha^2$. This means that 2 is a quartic residue after all.

$\boxed{p \in \vartheta} \Longrightarrow N = pN_1$:

$$pN_1^2 = M^4 + 4sM^2e^2 - 4e^4$$

where $p \nmid Me$. Thus $u^4 + 4su^2 - 4$ has a root in $\mathbf{F}_p$. Since $u^2 = 2(-s \pm \sqrt{2}) \in \mathbf{F}_p$, therefore $p \equiv \pm 1 \bmod 8$ and at least one of $-s \pm \sqrt{2}$ is a quadratic residue. When $p \equiv 1 \bmod 8$, by the lemma the latter is true iff $1 + i$ is a quadratic residue.

In case (a) when $p \equiv 1 \bmod 8$ we can prove that in fact $p \equiv 1 \bmod 16$. Stroeker and Top again use the arithmetic of $\mathbf{Z}[\sqrt{2}]$; for variety we use a Lind-type argument. Write

$$pN_1^2 = 2M^4 - (2e^2 - M^2)^2, \quad \text{where } p \nmid Me,$$

and let q be an odd prime divisor of $2e^2 - M^2$. Clearly $q \neq p$ and $q \nmid Me$. Since $2e^2 \equiv M^2 \bmod q$, therefore $(2/q) = 1$. Hence from $pN_1^2 \equiv 2M^4 \bmod q$ we deduce $(p/q) = 1$, so by reciprocity $(q/p) = 1$. Thus $2e^2 - M^2 = 2^\alpha \beta$ where $\alpha \geq 0$, β is odd, and $(\beta/p) = 1$, say $\beta \equiv \gamma^2 \bmod p$. Hence $2M^4 \equiv 2^{2\alpha}\gamma^4 \bmod p$ which implies that 2 is a quartic residue $\bmod p$. Since we already have $\sqrt{1+i} \in \mathbf{F}_p$, this implies that $\mathbf{F}_p$ contains the 16th root of unity $\sqrt{(1+i)/\sqrt{2}}$. This means that $p \equiv 1 \bmod 16$.

$\boxed{2p \in \vartheta} \Longrightarrow N = 2pN_1$:

$$2pN_1^2 = M^4 + 2sM^2e^2 - e^4.$$

Since $\gcd(M, e) = 1$, therefore $\gcd(2p, Me) = 1$. Thus $u^4 + 2su^2 - 1$ has a root in $\mathbf{F}_p$. Since $u^2 = -s \pm \sqrt{2}$, this implies that $p \equiv \pm 1 \bmod 8$, and when $p \equiv 1 \bmod 8$ that $1 + i$ is a quadratic residue.

Since M and e are odd, $M^4 \equiv e^4 \equiv 1 \bmod 16$, hence $2pN_1^2 \equiv 2sM^2e^2 \bmod 16$, and $pN_1^2 \equiv sM^2e^2 \equiv s \bmod 8$. It follows that $p \equiv 1$ (resp. -1) $\bmod 8$ in case (a) (resp. (b)).

In case (a) a Lind-type argument applied to $2pN_1^2 = (M^2+e^2)^2 - 2e^4$, which is left to the reader, shows that 2 is a quartic residue $\bmod p$, hence $p \equiv 1 \bmod 16$. ■

Concerning $E(\mathbf{Q})$ with larger r,

> *it is conjectured that every integer $r \geq 0$ occurs; this remains an outstanding open problem.*

It is not even known if r is unbounded. Basing their work on a construction of Mestre, Nagao [Nag94] and Fermigier [Fer96] found $E_{/\mathbf{Q}}$ for which r is at least 21, resp. at least 22. We will return to this and related matters on several occasions in later chapters.

3.6.6 Second descent

For simplicity we restrict the discussion in this section to $E_{/Z}$.

In the notation we have been using, let

$$E : y^2 = x(x^2 + ax + b), \quad \overline{E} : y^2 = x(x^2 + \overline{a}x + \overline{b}), \quad a, b \in \mathbf{Z}$$

where $$\overline{a} = -2a, \quad \overline{b} = a^2 - 4b,$$

and suppose we have not been able to find a rational point on the locally solvable torsor

$$u^2 = b_1 v^4 + av^2 + b_2, \quad b_1 b_2 = b.$$

As usual we can assume that b_1 is square-free, $\neq 1$ and $\neq$ the square-free part of b. The following procedure, referred to as **second descent**, can sometimes discover a rational point (u, v) or prove that there is none. (By Proposition 3.6.4, a point (u, v) transforms to the point $(b_1 v^2, b_1 uv) \in E(\mathbf{Q})$.)

Since the quartic is locally solvable, by the Hasse principle the associated quadratic

$$u^2 = b_1 w^2 + aw + b_2$$

does have a rational point, say (u_0, w_0). We note that $w_0 = 0$ will not occur at this stage since then b_2 is a square and the quartic has an obvious rational point. ($b_1 = b/b_2 \equiv b$ mod squares would already be in φ. Of course the present discussion applies equally to ϑ.)

As explained in §1.4.3, all rational points on the quadratic are given parametrically by

$$w = \frac{w_0 t^2 - 2u_0 t + a + b_1 w_0}{t^2 - b_1} =: \frac{f(t)}{g(t)},$$

$$u = \frac{-u_0 t^2 + (a + 2b_1 w_0)t - u_0 b_1}{t^2 - b_1}.$$

Letting $\square$ denote (various) rational squares, we wish to find $w = \square$, *i.e.,* we seek $t \in \mathbf{Q}$ such that

$$f(t) = \delta * \square \quad \text{and} \quad g(t) = \delta * \square \tag{$*$}$$

for some nonzero squarefree $\delta \in \mathbf{Z}$.

Let k denote the denominator of w_0. Since $u_0^2 = b_1 w_0^2 + aw_0 + b_2$, therefore $ku_0 \in \mathbf{Z}$, hence $kf \in \mathbf{Z}[t]$. It follows that

$$R := \text{resultant}(kf, kg) = \lambda kf + \mu kg$$

for some linear $\lambda, \mu \in \mathbf{Z}[t]$,* and calculation shows that $R = k^4 \overline{b}$.

Suppose $t = t_1/t_2$, where $\gcd(t_1, t_2) = 1$, is a solution to $(*)$. Then

$$[t_2 \lambda(t)] \left[t_2^2 k f(t)\right] + [t_2 \mu(t)] \left[t_2^2 k g(t)\right] = t_2^3 k^4 \overline{b} \in \mathbf{Z}$$

where the factors in square brackets are integers. Thus $\delta n = t_2^3 k^4 \overline{b}$ for some $n \in \mathbf{Z}$. Now δ is square-free, also $t_2^2 g(t) = t_1^2 - b_1 t_2^2 = \delta * \square$ and $\gcd(t_1, t_2) = 1$.

*Explicit values can be calculated using the formula for R in [Con82, p.213]:

$$\begin{aligned} \lambda &= k^2(2u_0 t + a + 2b_1 w_0), \\ \mu &= k^2(-2u_0 w_0 t + 4u_0^2 - aw_0 - 2b_1 w_0^2). \end{aligned}$$

It follows that no prime dividing δ divides t_2, and therefore $\delta | k\bar{b}$. Thus we obtain a finite list of candidate δ.

Consider first the special case $\delta = 1$. The homogeneous form of the equation $t^2 - b_1 = s^2$ is $T^2 - b_1 U^2 = S^2$ which has the solution $(T, U, S) = (1, 0, 1)$. The general solution is

$$(\delta = 1) \quad t = \frac{\tau^2 + b_1}{2\tau}, \quad \tau \text{ a parameter.} \tag{1}$$

In general, since $g(t) = \delta * \square$ implies an equation of the form

$$\delta t^2 - \delta b_1 = s^2, \tag{$**$}$$

we have the norm residue conditions (*cf.* §3.6.4)

$$(\delta, \delta b_1)_p = 1 \quad \text{for } p = \infty \text{ and all prime divisors of } \delta b_1,$$

which may eliminate some $\delta \neq 1$.

Suppose that $\delta \neq 1$ and that δ survives the norm residue conditions so that there is a solution (s_0, t_0) of $(**)$. (The homogeneous equation $\delta T^2 - \delta b_1 U^2 = S^2$ does not have a solution with $U = 0$ since δ is not a square.) If it happens that $f(t_0) = \delta * \square$, then the procedure is done — we have found a rational point on the original quartic. Otherwise we look at the general solution

$$(\delta \neq 1) \quad t = \frac{t_0 \tau^2 - 2 s_0 \tau + \delta t_0}{\tau^2 - \delta}, \quad \tau \text{ a parameter.} \tag{2}$$

Since we want $f(t) = \delta * \square$, clearing denominators and substituting for t according to (1) or (2), we seek a value of τ such that, respectively,

$$4\tau^2 f(t) = \square \quad \text{or} \quad \delta(\tau^2 - \delta)^2 f(t) = \square.$$

Thus we obtain a list of one or more equations of the form

$$\text{quartic in } \tau = \square.$$

These equations are called the **descendant quartics**. The procedure now is to eliminate the descendant quartics that are not everywhere locally solvable, *i.e.,* not solvable in $\mathbf{R}$ or some $\mathbf{Q}_p$. (There is an algorithm for this involving Hensel's lemma; but we postpone the description to a later chapter when we have the theorem of F.K. Schmidt alluded to in §3.6.4.) If none survive then the original quartic has no rational points — the torsor is not elliptic — and the procedure is done. Finally, one searches for a rational point on one of the surviving descendant quartics.

Example (i)

$$E : y^2 = x(x^2 - 8x + 1), \quad \overline{E} : y^2 = x(x^2 + 16x + 60),$$

$$a = -8, \quad b = 1, \qquad \overline{a} = 16, \quad \overline{b} = 60,$$

and our usual starting position is

$$\varphi = \{1\}, \quad \{1, 15\} \subseteq \vartheta \subseteq \{\pm 1, \pm 2, \pm 3, \pm 5, \pm 6, \pm 10, \pm 15, \pm 30\},$$

hence for the common rank r of $E(\mathbf{Q})$ and $\overline{E}(\mathbf{Q})$we have the estimate

$$0 \leq r \leq 2.$$

But for $2^r = |\varphi||\vartheta|/4$ to make sense, ϑ must contain at least four elements. These are supplied by the other points of order 2 on $\overline{E}$, $(-6, 0)$ and $(-10, 0)$:

$$\{1, -6, -10, 15\} \subseteq \vartheta.$$

Apecs assures us that all the torsors corresponding to the elements in the upper bound $\{\pm 1, \ldots, \pm 30\}$ are elliptic. Second descent and the Hensel lemma algorithm alluded to above are of course fully implemented in apecs.

We apply second descent to $u^2 = -v^4 + 16v^2 - 60$, where $b_1 = -1$, $b_2 = -60$ and ϑ is taking the place of φ. Let us choose the point $(u_0, w_0) = (0, 6)$ on the corresponding quadratic $u^2 = -w^2 + 16w - 60$, so $k = 1$ and

$$w = \frac{6t^2 + 10}{t^2 + 1}.$$

Since $\overline{E}$ is playing the role of E, the correct value for the resultant is $R = k^4\overline{\overline{b}} = 16$ — we must avoid the mistake of taking the value $k^4 b = 1$. Since $b_1 < 0$, $(**)$ shows that we need only consider positive δ, hence the list of candidates for δ is 1,2.

$\underline{\delta = 1}$

$$t = \frac{\tau^2 - 1}{2\tau}$$

and the descendant quartic works out to

$$6\tau^4 + 28\tau^2 + 6 = \square.$$

This has no solution in $\mathbf{Q}_2$, hence none in $\mathbf{Q}$. For, by elementary reasoning, τ must be a 2-adic unit, and for such τ,

$$6\tau^4 + 28\tau^2 + 6 \equiv 8 \bmod 16,$$

which can never be a square in $\mathbf{Q}_2$.

$\underline{\delta = 2}$ $(**)$ has the solution $(s_0, t_0) = (2, 1)$, hence

$$t = \frac{\tau^2 - 4\tau + 2}{\tau^2 - 2}.$$

The descendant quartic divided by the square 16 is

$$2(\tau^4 - 3\tau^3 + 5\tau^2 - 6\tau + 4) = \square.$$

As in the case $\delta = 1$, this has no solution in $\mathbf{Q}_2$. For clearly τ must be a 2-adic integer and by taking $\tau = 1, \ldots, 15$ one finds that $2(\tau^4 - \cdots + 4)$ is never a square mod16.

We have eliminated $b_1 = -1$ and made progress:

$$0 \le r \le 1.$$

Next, second descent on $u^2 = 2v^4 + 16v^2 + 30$ appears to be a tad arduous (with $(u_0, w_0) = (4, -1)$ we have to deal with the four possibilities $\pm 1, \pm 2$ for δ), so we take $u^2 = -2v^4 + 16v^2 - 30$. On the quadratic we choose the point $(u_0, w_0) = (0, 3)$, so $k = 1$ and

$$f(t) = 3t^2 + 10, \quad g(t) = t^2 + 2.$$

Again $R = 16$, $(**)$ eliminates $\delta < 0$, and we are left with $\delta = 1$ or 2. We leave it to the reader to check that neither descendant quartic is solvable over $\mathbf{Q}_2$.

Since $-6 \in \vartheta$ and $-2 \notin \vartheta$ therefore $-6/-2 = 3 \notin \vartheta$; and by similar elementary reasoning we arrive at $\vartheta = \{1, -6, -10, 15\}$ and $r = 0$.

Example (ii) We say that a rational number n/d is *large* when $\max\{|n|, d\}$ is a large integer, and that a rational point (x, y) is *large* when x is large. Sometimes a relatively small rational point on one of the descendant quartics can produce a spectacularly large point on the torsor, one that could not have been found in a reasonable amount of time in a straightforward search.

Apecs supplied the following unexceptional example, chosen more or less at random. One of the descendant quartics produced by second descent on

$$E : y^2 = x(x^2 + 1001x + 1001)$$

is

$$-4\tau^4 + 242\tau^3 + 371\tau^2 - 1246\tau + 588 = \sigma^2$$

which was found to have the rational point

$$(\tau, \sigma) = (-12/137, 496518/137^2).$$

This transforms to the $E(\mathbf{Q})$ point

$$(1373811990322540873104/547262189^2,$$

$$562033981583503034679013611751 08/547262189^3).$$

3.6.7 A transcendental example

Let t be an indeterminate and

$$E : y^2 = x^3 - t(t-1)^2 x, \quad \Delta = 64t^3(t-1)^6, \quad j = 1728.$$

We will show that the rank r of $E(\mathbf{C}(t))$ is 2 by applying Proposition 3.6.4 with respect to the PID $\mathbf{C}[t]$. We note that besides the usual automorphism $P = (x, y) \mapsto -P = (x, -y)$, this group also has *complex multiplication by* $i = \sqrt{-1}$ (to be discussed in Chapter 9):

$$P = (x, y) \longmapsto \overline{P} = (-x, iy).$$

That this is a group automorphism, where of course we define $\overline{O} = O$, is immediate from the addition formulas (Proposition 1.7.1).

In Example 1 following Proposition 2.10.3 we determined the torsion subgroup $\mathcal{T}$ of $E(\mathbf{C}(t))$ to be of order 2: $\mathcal{T} = \{O, T = (0, 0)\}$; and in the course of applying Nagell-Lutz we uncovered the point $P_1 = (1 - t, (t - 1)^2)$ of infinite order.

Proposition 3.6.4 with $R = \mathbf{C}[t]$ and $K = \mathbf{C}(t)$ gives $\varphi = \{1, t, t-1, t(t-1)\}$; for $\mathbf{C}^{*2} = \mathbf{C}^*$, hence the set of divisors of $b = -t(t-1)^2$ mod squares is the indicated set, and $\phi_1(T) = -t(t-1)^2 K^{*2} = tK^{*2}$, and $\phi_1(P_1) = (1-t)K^{*2} = (t-1)K^{*2}$. Similarly for

$$\overline{E} : y^2 = x^3 + 4t(t-1)^2 x,$$

we find the points $(0, 0)$ and $(2i(t-1), 2(1+i)(t-1)^2)$, hence $\vartheta = \{1, t, t-1, t(t-1)\}$. Thus $2^r = 4 \times 4/4 = 4$, and $r = 2$.

We notice the point† $P_2 = (2it, (1-i)t(t+1)) \in E(\mathbf{C}(t))$, which has infinite order since $y^2 \nmid \Delta$, and the relations

$$\overline{P_1} = P_1 - P_2 + T, \quad \overline{P_2} = [2]P_1 - P_2.$$

An easy induction using the addition formulas shows that for all nonzero integers m, $[m]P_2 = (if_m, (1-i)g_m)$ where $f_m, g_m \in \mathbf{R}(t)$, hence $[m]P_2 \notin E(\mathbf{R}(t))$, whereas for all integers n and h, clearly

$$[n]P_1 + [h]T \in E(\mathbf{R}(t)).$$

Thus if $[m]P_2 + [n]P_1 + [h]T = O$, then $m = 0$, hence $[n]P_1 \in \mathcal{T}$ which implies $n = 0$. It follows that

$$\mathcal{T} \oplus \langle \{P_1, P_2\} \rangle$$

sits as a subgroup of finite index in $E(\mathbf{C}(t))$. Actually the index is 1 ([Cohn80]), but we lack the tools for a proof at this point in the notes. (*Exercise.* By Proposition 1.7.5(b), the index is odd.)

†discovered by searching for $P = (x, y)$ where $x|\Delta$ as suggested by Proposition 2.10.2(a).

3.7 Billing's upper bound for the rank

In this section we find an upper bound for the rank of $E(\mathbf{Q})$ where E is an elliptic curve defined over $\mathbf{Q}$ which does not have a point of order 2 defined over $\mathbf{Q}$, and so the method of simple 2-descent of the previous sections is not available.

Let E be defined over $\mathbf{Z}$:

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \quad a_i \in \mathbf{Z}.$$

By Mordell's theorem we can write

$$E(\mathbf{Q}) = \mathcal{T} \oplus \mathbf{Z}^r, \quad \mathcal{T} \text{ finite.}$$

We define the **2-minimal *b*-form** as

$$E' : y'^2 = x'^3 + 2^{2i-2}b_2x'^2 + 2^{4i-1}b_4x' + 2^{6i-2}b_6$$

where $i \in \mathbf{Z}$ is chosen minimal such that E' is defined over $\mathbf{Z}$. Then

$$x' = 2^{2i}x, \quad y' = 2^{3i}\eta = 2^{3i-1}(2y + a_1x + a_3)$$

define an isomorphism between $E(\mathbf{Q})$ and $E'(\mathbf{Q})$, and in particular, the two groups have the same rank r. We let $f = f(x') = x'^3 + ax'^2 + bx' + c$ denote the right side in the equation of E'. The polynomial discriminant of f is

$$D_f = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2 = 2^{12i-4}\Delta,$$

where Δ is the usual discriminant of E.

Now assume that $E(\mathbf{Q})[2] = O$, so f is irreducible over $\mathbf{Q}$. Let the roots of f in $\mathbf{C}$ be $\lambda, \lambda', \lambda''$. Then

$$D_f = \begin{vmatrix} 1 & \lambda & \lambda^2 \\ 1 & \lambda' & \lambda'^2 \\ 1 & \lambda'' & \lambda''^2 \end{vmatrix}^2 = (\lambda - \lambda')^2(\lambda' - \lambda'')^2(\lambda'' - \lambda)^2.$$

Let $\mathcal{O}$ be the ring of integers in the cubic field $L = \mathbf{Q}(\lambda)$, and let $1, \alpha, \beta$ be an integral basis of $\mathcal{O}$. For any $\gamma \in L$, we let γ', γ'' denote its conjugates. Then the field discriminant is

$$D_L = \begin{vmatrix} 1 & \alpha & \beta \\ 1 & \alpha' & \beta' \\ 1 & \alpha'' & \beta'' \end{vmatrix}^2.$$

Let $I = I(\lambda)$ denote the **index** of λ:

$$I = [\mathcal{O} : \mathbf{Z}[\lambda]] = \sqrt{D_f/D_L}.$$

I has the property that if $\gamma \in \mathcal{O}$, $\gamma \notin \mathbf{Z}[\lambda]$ and $p\gamma \in \mathbf{Z}[\lambda]$ for a prime p, then $p|I$. Of course D_L is uniquely determined by L: if γ, δ, ϵ is another integral basis then, using matrix notation, $(\gamma, \delta, \epsilon) = (1, \alpha, \beta)M$ where M is a 3×3 integral matrix with determinant ± 1. There is little room to choose D_f; even replacing λ by $d\lambda$ where d is a nonzero integer gives the curve

$$y'^2 = x'^3 + adx'^2 + bd^2x' + cd^3$$

which is obtained from the equation for E' by replacing x', y' with $x'/d, y'/d\sqrt{d}$. Unless d is a square, this twist of E can very well have an r different from that of E.

We make a definition:

> An elliptic curve E defined over $\mathbf{Z}$ is **quasi-supersingular at 2** if it satisfies the four conditions

(QSS1) $E(\mathbf{Q})[2] = O$, so with the above notation, $L = \mathbf{Q}(\lambda)$ is a cubic field;

(QSS2) 2 is ramified in L; let v denote the valuation on L with $v(2) \geq 2$;

(QSS3) $v(\lambda) \geq 2$;

(QSS4) there exists an odd integer N such that for all $(x, y) \in [N]E'(\mathbf{Q})$, where E' is the 2-minimal b-form of E, we have $v(x) \leq 0$.

Lemma 3.7.1 *An E defined over $\mathbf{Z}$ with supersingular reduction at 2* [†] *is quasi-supersingular at 2; furthermore, 2 is totally ramified in L and $\pi = \lambda/2$ is a uniformizer for the valuation v.*

Proof. Since a_1 is even and a_3 is odd, by the formulas in §1.1, we can write $b_2 = 4\beta_2$, $b_4 = 2\beta_4$ where $\beta_2, \beta_4 \in \mathbf{Z}$, and b_6 is odd. Thus

$$\pi^3 + 2\beta_2\pi^2 + 2\beta_4\pi + 2b_6 = 0$$

which is Eisenstein at 2. This proves (QSS1)–(QSS3).

In (QSS4) we can take $N = 15$. For by Proposition 2.10.4(c), if $\widetilde{E}$ denotes the reduction of $E \bmod 2$, $|\widetilde{E}(\mathbf{F}_2)| = 1$, 3 or 5. Thus $[15]E(\mathbf{Q}) \subseteq E_1(\mathbf{Q})$ where $E_1(\mathbf{Q})$ is the kernel of the reduction homomorphism, *i.e.*, O and all nonzero points of the form $P = (m/(2e)^2, t/(2e)^3) \in E(\mathbf{Q})$ where $m, e, t \in \mathbf{Z}$ and m and t are *odd*. The 2-minimal b-form is $y'^2 = x'^3 + b_2x'^2 + 8b_4x' + 16b_6$, and the isomorphism $\theta : E \longrightarrow E'$ is defined by $\theta(x, y) = (x', y') = (4x, 8y + 4a_1x + 4a_3)$. Thus the points $\theta(P) \in [15]E'(\mathbf{Q})$ have $x(\theta(P)) = m/e^2$ where m is odd. ■

An element of L is defined to be **positive** if its image is positive under the embedding $L \longrightarrow \mathbf{R}$ defined by $\lambda \mapsto$ the smallest of the 1 or 3 real roots of f.

[†] For the moment we interpret this assumption to mean that a_1 is even and a_3 is odd; *cf.* Proposition 2.10.4(c). After we clarify the notion of *isomorphism* between elliptic curves in the next chapter, we can understand this to mean that E is $\mathbf{Q}$-isomorphic to an elliptic curve defined over $\mathbf{Z}$ with a_1 even, a_3 odd.

Proposition 3.7.2 *Let E be an elliptic curve defined over $\mathbf{Z}$ and suppose $E(\mathbf{Q})[2] = O$; let the 2-minimal b-form be*

$$E' : y'^2 = x'^3 + ax'^2 + bx' + c = f(x'),$$

where $f(x')$ is irreducible over $\mathbf{Q}$, let λ be a root of $f(x')$, let $L = \mathbf{Q}(\lambda)$, let L have discriminant D_L and ring of integers $\mathcal{O}$, let I denote the index of $\mathbf{Z}[\lambda]$ in $\mathcal{O}$, and let $\mathrm{Cl}(\mathcal{O})$ denote the ideal class group of $\mathcal{O}$.

(a) [Bil38] *The rank r of $E(\mathbf{Q})$ satisfies*

$$r \le n_L + 2n_p + n_q + n_h,$$

where

$$\begin{aligned} n_L &= \begin{cases} 1 & \text{if } D_L < 0, \\ 2 & \text{if } D_L > 0, \end{cases} \\ n_p &= \text{the number of primes } p \text{ dividing } I \text{ which factor into three} \\ &\quad \text{distinct prime ideals in } \mathcal{O}, \\ n_q &= \text{the number of primes } q \text{ dividing } I \text{ whose factorization into} \\ &\quad \text{prime ideals in } \mathcal{O} \text{ has the form } PP' \text{ or } P^2P',\ P \ne P', \\ n_h &= \dim_{\mathbf{F}_2} \mathrm{Cl}(\mathcal{O})[2]. \end{aligned}$$

(b) *This upper bound for r can be reduced by 1 in either of the following situations.*

(b1) *The following two conditions are satisfied:*

(i) *The polynomial f is irreducible* $\bmod 2$ *(hence $\mathcal{O}/[2] \approx \mathbf{F}_8$, and $T = \{c_0 + c_1\lambda + c_2\lambda^2 : c_i \in \{0, \pm 1, 2\}$ is a set of representatives for the elements of $\mathcal{O}/[4]$);*

(ii) *$\mathcal{O}$ contains a positive unit ϵ such that for all $\alpha \in T$, $\epsilon\alpha^2$ is not congruent* $\bmod 4$ *to 1 or any one of $c_0 - \lambda$ where $c_0 \in \{0, \pm 1, 2\}$.*

(b2) *The following two conditions are satisfied:*

(i) *E is quasi-supersingular at 2; let v be the unique valuation on L with $v(2) > 1$, and let V be the ring of v and π a uniformizer;*

(ii) *$\mathcal{O}$ contains a unit $\epsilon \equiv 1 + \pi \bmod \pi^2 V$.*

Proof.

(a) For ideals A, B in $\mathcal{O}$, define the ideal greatest common divisor by

$$\mathrm{igcd}(A, B) = \prod P^{\min\{v_P(A), v_P(B)\}}$$

where the product is over the prime ideals P in $\mathcal{O}$ and $v_P(A)$ denotes the exponent of P in the prime ideal factorization of A. For $\gamma \in \mathcal{O}$, let $[\gamma]$ denote the principal ideal $\gamma\mathcal{O}$. The norm from L to $\mathbf{Q}$, both for elements and ideals, is denoted N.

A nonzero point in $E'(\mathbf{Q})$ has the form $(m/e^2, n/e^3)$ where $m, e, n \in \mathbf{Z}$, $e > 0$ and $\gcd(m, e) = 1$, equivalently $\gcd(n, e) = 1$. Let

$$\varphi = \phi_1\left(E'(\mathbf{Q})\right) = \left\{(m - \lambda e^2)L^{*2} : (m/e^2, n/e^3) \in E'(\mathbf{Q})\right\} \cup \{1L^{*2}\}.$$

By Proposition 3.2.1, $\varphi \approx E'(\mathbf{Q})/[2]E'(\mathbf{Q})$. Since $|\mathcal{T}|$ is odd, therefore $[2]\mathcal{T} = \mathcal{T}$, hence $|\varphi| = 2^r$. Thus we wish to obtain an upper bound on the dimension of the $\mathbf{F}_2$-vector space φ.

From

$$n^2 = (m - \lambda e^2)(m - \lambda' e^2)(m - \lambda'' e^2)$$

we see that any prime ideal occuring to an odd power in $[m - \lambda e^2]$ must also divide $[\gamma]$ where $\gamma = (m - \lambda' e^2)(m - \lambda'' e^2)$. Now $\operatorname{igcd}([m - \lambda e^2], [\gamma])$ divides the different

$$(\lambda - \lambda')(\lambda - \lambda'') = f'(\lambda) = 3\lambda^2 + 2a\lambda + b;$$

for if P is any prime ideal such that $v_P(m - \lambda e^2) > 0$, then $v_P(e) = 0$ and the Taylor expansion of f about the point $x = \lambda$, $f(x) = (x - \lambda)f'(\lambda) + \cdots$, when evaluated at $x = m/e^2$ and multiplied by e^6, shows that $v_P(f'(\lambda)) \geq \min\{v_P(m - \lambda e^2), v_P(\gamma)\}$. It follows that $\operatorname{igcd}([m - \lambda e^2], [\gamma])$ also divides $\mathrm{N}(f'(\lambda)) = -D_f$. Thus we can write

$$[m - \lambda e^2] = AB^2 \tag{$*$}$$

where A is a square-free ideal dividing D_f. There are only finitely many possibilities for A. (In fact we are in the process of proving weak Mordell-Weil for the class of E at hand.) This equation implies

$$\mathrm{N}(m - \lambda e^2) = n^2 = \mathrm{N}(A)\mathrm{N}(B)^2,$$

hence $\mathrm{N}(A)$ is a square.

For $\nu = 1, 2, 3$ and a prime p, let $P_\nu, P'_\nu, \ldots$ denote distinct prime ideals with $\mathrm{N}(P_\nu) = p^\nu$ (when such exist). If p is a prime divisor of D_f there are five possible factorizations of $[p]$:

1. $[p] = P_3$;
2. $[p] = P_1^3$;
3. $[p] = P_1^2 P'_1$;
4. $[p] = P_1 P_2$;
5. $[p] = P_1 P'_1 P''_1$.

In cases 1 and 2, $\operatorname{igcd}(A, [p]) = [1]$ since otherwise we would have $v_p(\mathrm{N}(A)) = 3$ or 1, respectively, contrary to the fact that $\mathrm{N}(A)$ is a square.

In case 3, either $\mathrm{igcd}(A,[p]) = [1]$ or $P_1P_1'|A$ and $v_p(\mathrm{N}(A)) = 2$.

Similarly in case 4, either $\mathrm{igcd}(A,[p]) = [1]$ or $P_1 \nmid A$, $P_2|A$ and $v_p(\mathrm{N}(A)) = 2$.

In case 5, besides $\mathrm{igcd}(A,[p]) = [1]$ there are three possibilities: A is divisible by precisely two of P_1, P_1', P_1'', and again $v_p(\mathrm{N}(A)) = 2$.

We can now prove

$$p\,|\mathrm{N}(A) \Longrightarrow p|\,I.$$

In case 3 we have $P_1P_1'|\,[m-\lambda e^2]$, so

$$[p]\,\big|P_1^2P_1'^2 \quad \text{and} \quad P_1^2P_1'^2\big|\,[m-\lambda e^2]^2,$$

hence
$$\gamma := \frac{m^2 - 2\lambda m e^2 + \lambda^2 e^4}{p} \in \mathcal{O} \ \text{ and } \ p\gamma \in \mathbf{Z}[\lambda].$$

Since $p \nmid e$, therefore $\gamma \notin \mathbf{Z}[\lambda]$, which implies $p|I$.

In cases 4 and 5 we must have $p \nmid D_L$ since any prime dividing D_L is ramified. Thus $p\,|D_f$ and $p \nmid D_L$, hence $p|I$.

It follows that the number of possible A is at most $2^{2n_p+n_q}$, where n_p (resp. n_q) is the number of prime divisors of I of type 5 (resp. of type 3 or 4).

For each A that occurs we choose a particular equation $(*)$ and write it as $[\mu_A] = AB_A^2$. In the list of μ_A we include $\mu_{[1]} = 1$, which may not actually arise from an equation $(*)$, but corresponds to the point O. Then any equation $(*)$ can be written

$$[m-\lambda e^2] = [\mu_A](BB_A^{-1})^2.$$

Thus the ideal class of BB_A^{-1} belongs to $\mathrm{Cl}(\mathcal{O})[2]$. If $J_1, \ldots, J_{n_h}$ are ideals whose classes form a $\mathbf{F}_2$-basis of the latter group, and $J_i^2 = [\xi_i]$ where $\xi_i \in \mathcal{O}$, then

$$(BB_A^{-1})^2 = [\alpha^2 \xi_1^{\nu_1} \xi_2^{\nu_2} \cdots]$$

for some $\alpha \in L^*$ and $\nu_i \in \{0,1\}$. Thus

$$m - \lambda e^2 = u\mu_A \alpha^2 \prod_{i=1}^{n_h} \xi_i^{\nu_i} \tag{$**$}$$

where u is a unit in $\mathcal{O}$.

We are at liberty now to choose λ as a real root of f, and in the case $D_f > 0$, as the smallest of the three real roots. This implies $m - \lambda e^2 > 0$, and in particular, all $\mu_A > 0$. Replacing ξ_i by $-\xi_i$ as necessary, we can choose all $\xi_i > 0$, and then $u > 0$. The group of positive units in $\mathcal{O}$ is free of rank n_L. Thus the number of distinct $m - \lambda e^2 \bmod L^{*2}$, *i.e.*, $|\varphi|$, is at most 2^n where $n = n_L + 2n_p + n_q + n_h$.

(b) The upper bound for $\dim_{\mathbf{F}_2}\varphi$ just stated was obtained by allowing in the right side of $(**)$ all positive units u, all $2^{2n_p+n_q}$ values of μ_A including $\mu_A = 1$,

and all possible combinations of $\nu_i \in \{0, 1\}$. Thus we can reduce the estimate by 1 when we can show that one of these values does not actually occur.

(b1) Let $P = (m/e^2, n/e^3) \in E'(\mathbf{Q})$ and define $z = m - \lambda e^2$. If e is even then m and n are odd and the Weierstrass equation implies that $n^2 \equiv m^3 \bmod 4$, hence $z \equiv m \equiv 1 \bmod 4$. If e is odd then $z \equiv c_0 - \lambda \bmod 4$ where $c_0 \in \{0, \pm 1, 2\}$.

Thus assumption (ii) implies that ϵ, which was counted as a possible right side in $(**)$ in our initial estimate, does not in fact occur.

(b2) Assume the two conditions in (b2) and let N denote the odd integer specified in (QSS4). Because of the isomorphism

$$[N]E'(\mathbf{Q})/[2N]E'(\mathbf{Q}) \approx E'(\mathbf{Q})/[2]E'(\mathbf{Q}),$$

we have

$$\varphi = \{\phi_1(P) : P = (m/e^2, n/e^3) \in [N]E'(\mathbf{Q})\} \cup \{1L^{*2}\},$$

where $m, n, e \in \mathbf{Z}$ with m, n odd. Since 2 and λ are 0 $\bmod \pi^2 V$, therefore $m - \lambda e^2 \equiv 1 \bmod \pi^2 V$. Each $\alpha \in \mathcal{O}$ is congruent $\bmod \pi^2 V$ to one of

$$0,\ 1,\ \pi,\ 1 + \pi,$$

hence α^2 is congruent $\bmod \pi^2 V$ to either 0 or 1. Thus $\epsilon\mu_1\alpha^2 \equiv 1 + \pi$ or $0 \bmod \pi^2 V$ does not occur on the right side of $(**)$. ■

3.7.1 Examples

The first example is due to Washington [Was87, Th.1]; the L that arise were named *the simplest cubic fields* by Shanks. (For a generalization see [Kaw-Na92].)

Corollary 3.7.3 *With the notation of the proposition, let E be given by*

$$y^2 = f(x) = x^3 + Mx^2 - (M + 3)x + 1$$

where M is a positive integer such that $M^2 + 3M + 9$ is square-free. Then E' coincides with E, L is a cyclic cubic extension of $\mathbf{Q}$ and

$$r \leq 1 + n_h.$$

Proof. f is irreducible over $\mathbf{Q}$ since it is irreducible mod 2, and L is cyclic over $\mathbf{Q}$ since $D_f = (M^2 + 3M + 9)^2$. In fact

$$\lambda' = 2 - (M + 1)\lambda - \lambda^2, \quad \lambda'' = -(2 + M) + M\lambda + \lambda^2.$$

These three roots are units since their norm is -1. Choosing the notation so that $\lambda < \lambda' < \lambda''$ in a real embedding, one finds that λ is negative and λ' and λ'' are positive.

In the Corollary to his Proposition 1, Washington proves that $1, \lambda, \lambda^2$ form an integral basis. We take this fact as 'background' number theory; it implies that $I = 1$, hence $n_p = n_q = 0$. Thus part (a) of the proposition yields $r \leq 2 + n_h$, and it remains to check that (b1) applies.

In fact computer calculation shows that for all nonzero $\alpha \in T$, $\lambda''\alpha^2 \equiv c_0 + c_1\lambda + c_2\lambda^2 \bmod 4$ where $c_2 \not\equiv 0 \bmod 4$. ■ †

In the following examples, reference will be made to the tables of cubic fields in Appendix B of [Coh93], and for 'pure' cubic fields $\mathbf{Q}(\sqrt[3]{m})$, to Table 1 in [Cas50]. The next example will be the key to proving in chapter 8 that no $E_{/\mathbf{Q}}$ has a point of order of order 11 defined over $\mathbf{Q}$.

Corollary 3.7.4 *For*

$$E : y^2 + y = x^3 - x^2, \quad \Delta = -11, \qquad \textbf{A11}$$

the rank of $E(\mathbf{Q})$ *is* 0. *More precisely,*

$$E(\mathbf{Q}) = \{O, (0,0), (1,0), (1,-1), (0,-1)\} = C_5.$$

Proof. That the torsion subgroup $\mathcal{T} = C_5$ as indicated is an easy application of Nagell-Lutz and is left to the reader.

Condition (b2)(i) of the proposition is met since **A11** has supersingular reduction at 2; $f = x_1^3 - 4x_1^2 + 16$, so $D_f = -2^8 11$ and $\pi^3 - 2\pi^2 + 2 = 0$ is the Eisenstein equation. The discriminant of the order $\mathbf{Z}[\pi]$ is -44, and so L must be the unique cubic field of discriminant -44. In fact calculation shows that, using the notation of Table B.3 of [Coh93, p.509], $\alpha = \epsilon = 1 - \pi$ and $h = 1$. Thus condition (ii) is met. Since $\mathcal{O} = \mathbf{Z}[\alpha] = \mathbf{Z}[\pi]$, the index $I = 2^3$, hence $n_p = n_q = 0$, and $r = 0$ follows. ■

The preceding corollary was first proved by Billing and Mahler [Bil-Ma40]; they used part (a) of the proposition, but an *ad hoc* argument in place of (b2) that applied only to **A11**. We will find (b2) useful also for other curves.

Corollary 3.7.5 *For the curve*

$$E : y^2 + y = x^3 - 7, \quad \Delta = -3^9, \qquad \textbf{B27}$$

$E(\mathbf{Q})$ *has rank* 0 *and* $\mathcal{T} = \{O, (3,4), (3,-5)\}$. *The* 2*-minimal b-form of this curve is*

$$y'^2 = x'^3 - 432,$$

hence, by Corollary 1.4.3*, we have Euler's result that Fermat's last theorem is true for exponent* 3.

† λ' does not work: $\lambda'\alpha^2 \equiv 1 - \lambda$ for appropriate α.

Proof. E has supersingular reduction at 2, and $\mathcal{T}$ is as stated by Nagell-Lutz. $L = \mathbf{Q}(\pi)$ where $\pi = \sqrt[3]{2}$ with integral basis $1, \pi, \pi^2$, $D_L = -108$, $I = 6^3$ and $h = 1$. Thus $n_L = 1$, $n_p = n_q = n_h = 0$. Since $\epsilon = 1 + \pi + \pi^2$, (b2) gives $r = 0$. ■

Corollary 3.7.6 *For*

$$E : y^2 + y = x^3 - x, \quad \Delta = 37, \qquad \textbf{A37}$$

$E(\mathbf{Q}) = \langle (0,0) \rangle \approx \mathbf{Z}$. *The integral points on* $E(\mathbf{Q})$ *are precisely the ten listed in the table in example* 4, §2.10.1.

Proof. This is another example that enjoys the reduction afforded by part (b2) of the proposition. Again the reduction at 2 is supersingular. In this case, $D_f = 2^8 37$, $\pi^3 - 4\pi + 2 = 0$, $\mathcal{O} = \mathbf{Z}[\pi]$, and L is the unique totally real cubic field of discriminant 148. Using table B.4 [*ibid.*] one finds, after some minor calculation, that $\alpha = 1 + \pi + \pi^2$, $\epsilon_1 = \alpha$, and $\epsilon_2 = -3 + \pi^2$. Condition (b)(ii) is satisfied by $\epsilon = \epsilon_1$. We have $I = 8$ hence $n_p = n_q = 0$, and $h = 1$ hence $n_h = 0$. This time $n_L = 2$ so $r \le 1$. Since $(0,0)$ has infinite order, $r = 1$.

The discriminant $\Delta = 37$ is positive; the real points of order 2 have x-coordinates approximately -1.1, .27, and .84, hence the integral points in the odd component are precisely the four points $Q = (0,0)$, $-Q = (0,-1)$, $[3]Q = (-1,-1)$, and $[-3]Q = (-1,0)$. By Lemma 3.6.8, Q is a free generator. Noting that

$$[8]Q = (21/5^2, 69/5^3), \quad [12]Q = (23 \cdot 59/29^2, 2^3 \cdot 23 \cdot 157/29^3)$$

are fractional, the determination of all integral points follows from an application of part (b) of the

Lemma 3.7.7 (a) *Let E be an elliptic curve defined over* $\mathbf{Z}$ *and let* $Q \in E(\mathbf{Q})$. *Then there is an integer* $a \ge 0$ *such that* $[2^i]Q$ *is integral (resp. non-integral) for* $0 \le i < a$ *(resp.* $i \ge a$*).*

(b) *Let* $E_{/\mathbf{Z}}$ *satisfy conditions* (i)–(iv) *of* Lemma 3.6.8 *and*

(v) *the torsion subgroup* $\mathcal{T}$ *of* $E(\mathbf{Q})$ *has odd order, for example* $\mathcal{T} = O$.

Then all integral points can be determined by the following algorithm:

— *Let* $R_1, \ldots, R_n$ *be the integral points in the odd component;*

— *for each* R_i *let* a_i *be the smallest positive integer such that* $[2^{a_i}]R_i$ *is non-integral.*

Then the integral points in $E(\mathbf{Q})$ *are precisely*

$$[2^j]R_i : 1 \le i \le n,\ 0 \le j < a_i.$$

Proof. (a) We first note that $E(\mathbf{R})[3] = C_3$. Geometrically this is clear: a point $S \ne O$ is of order 3 iff $[2]S = -S$ iff S is a flex, *i.e.,* the tangent at S cuts through the curve. There are no flexes on the odd component (algebraically one

says that $S = [2]([2]S)$ is on the even component) and there are precisely two flexes in the affine part of the plane on $E^o(\mathbf{R})$ for the following reason. Think of tracking the slope of the tangent on the upper branch of the curve starting from the leftmost point on $E^o(\mathbf{R})$ and letting x increase. The tangent at the leftmost point is vertical, then the slope becomes finite and at first decreases, then attains a minimum — at the upper flex — and then increases monotonically to ∞ since $y \sim x^{3/2}$.

Algebraically this can be proved by calculating the Sturm sequence for the 3-division polynomial ψ_3. One finds that for all $E_{/\mathbf{R}}$, ψ_3 has two real roots and a pair of conjugate complex roots. For one of the real roots the values of y are real, while for the other they are complex. (This involves some work, but it is a good refresher course in Sturm theory.)

As a temporary notation, let ξ denote the x-coordinate of the two real flexes, define the **left half** to be $\{(x, y) \in E^o(\mathbf{Q}) : x \leq \xi\}$ and the **right half** to be the complement of the left half in $E^o(\mathbf{Q})$.

Now let $Q \in E(\mathbf{Q})$. Since $[nn']Q$ integral implies that $[n]Q$ is integral (Proposition 2.10.1(a)), the point is to prove that a is *finite*. Since O is non-integral by definition, the statement is clear for $Q \in \mathcal{T}$ by Proposition 2.10.4(a). Thus we assume that Q is a point of infinite order in the neutral component such that all $[2^i]Q$ are integral, and wish to derive a contradiction.

Let m_i denote the absolute value of the slope at $(x_i, y_i) := [2^i]Q$ and M the slope at the upper flex. If $[2^i]Q$ is in the left half then, since $m_i > M$, $[2^{i+1}]Q$ is in the right half. If $[2^i]Q$ is in the right half then $x_{i+1} \leq x_i - 1$, and therefore for some $j > 0$, $[2^{i+j}]Q$ is in the left half. Thus we would accumulate infinitely many distinct integral points in the left half; but this is not possible since the range of x in the left half is finite.

(b) Since $|\mathcal{T}|$ is odd, every $T \in \mathcal{T}$ can be written uniquely as $[2]T'$, and $\mathcal{T}$ is contained in the neutral component. Let R be a free generator in the odd component (the existence was proved in Lemma 3.6.8), so every point in $E(\mathbf{Q})$ is uniquely expressible in the form $[n]R + T$. Suppose $[n]R + T$ is integral, where $n = 2^j m$, m odd. Then $[n]R + T = [2]([2^{j-1}m]R + T')$ implies that $[2^{j-1}m]R + T'$ is integral. Continuing in this way, we find that $[m]R + T''$ is integral for some $T'' \in \mathcal{T}$. Also $[m]R+T''$ is in the odd component since $[m]R$ is in the odd component and T'' is in the even component. Hence $[m]R+T'' = R_i$ for some i, and $[n]R + T = [2^j]R_i$. ■

We state the remaining examples more concisely; as before, α, h, ϵ, refer to the tables in [Coh93].

$$y^2 + y = x^3 - x^2 - 736x - 18020, \quad \Delta = -11 \cdot 47^6 \qquad \mathbf{A11} * (-47)$$

is the quadratic twist of **A11** by -47 — to be explained in the next chapter. For this curve, $\mathcal{T} = O$, $f = x^3 - 4x^2 - 2^9 23x - 2^4 7^2 1471$, $D_f = -2^8 11 \cdot 47^6$, $L = \mathbf{Q}(\pi)$ where $\pi^3 - 2\pi^2 - 2^7 23\pi - 2 \cdot 7^2 1471 = 0$, $D_L = -44$, $\epsilon = \alpha = (15 + \pi)/47 \equiv 1 + \pi \bmod 2$, $I = 2^3 47^3$, $[47] = P_1 P_1' P_1''$, $n_p = 1$, $n_q = 0$,

$h = 1$ hence $n_h = 0$, (b2) applies and $r \leq 2$. In fact $r = 2$, two independent points being $Q_1 = (110, 1104)$ and $Q_2 = (205/4, 2205/8)$ with height pairing determinant

$$\det(\langle Q_i, Q_j \rangle) = \begin{vmatrix} 1.0073 & .3145 \\ .3145 & 1.7887 \end{vmatrix} = 1.703 \ .$$

For a connection between this curve and the Galois theory of $\mathbf{Q}$, see [Ser92, p.53]. Also, the significance of $r > 0$ for this curve will come out in §7.3.1: $m = -47$ is an 'extraordinary' value.

Here is an example where the reduction in (b) of the proposition does not apply.

$$E : y^2 = x^3 - x + 1 = f, \quad \Delta = -2^4 23 \qquad \textbf{C92}$$

has $\mathcal{T} = O$. A root of f is $\lambda = -1/\alpha = -\epsilon$ where $\alpha^3 + \alpha^2 - 1 = 0$ and $L = \mathbf{Q}(\lambda) = \mathbf{Q}(\alpha)$; $D_f = D_L = -23$, $I = 1$, $h = 1$. Hence $r \leq n_L = 1$. In fact $E(\mathbf{Q}) = \langle (1, 1) \rangle$.

Here is an example that brings n_h into play. The cubic field with smallest discriminant and $h > 1$ is $L = \mathbf{Q}(\alpha)$ where $a^3 + 4\alpha - 1$ for which $D_L = -283$ and $h = 2$, so $n_h = 1$. We take

$$E : y^2 = x^3 + 4 - 1$$

so by construction, $I = 1$. Thus $r \leq 2$ and in fact $E(\mathbf{Q}) = \langle Q_1, Q_2 \rangle$ where $Q_1 = (1, 2)$ and $Q_2 = (5, 12)$ are independent.

Of course the above examples were carefully chosen; "usually" r is less than the bound given by the proposition. Then to determine r another approach is needed. In [Bir-Sw63], Birch and Swinnerton-Dyer introduce the method of classifying torsors of E defined over $\mathbf{Q}$. In [Cre92], Cremona describes the method in detail and uses it to determine the ranks of the elliptic curves in his catalog; but for this, one must refer to the second (1997) edition for the latest refinements including his remarkable criterion to test the equivalence of two quartics. This affords a great simplification to the method (and is incorporated in the apecs procedure Crem.)

3.7.2 An example of Selmer

We now fulfill a promise made in Chapter 1.

Corollary 3.7.8 *The Selmer curve*

$$3U^3 + 4V^3 + 5W^3 = 0$$

has a point in $\mathbf{P}^2(\mathbf{R})$ *and in each* $\mathbf{P}^2(\mathbf{Q}_p)$*, but no point in* $\mathbf{P}^2(\mathbf{Q})$*.*

Proof. We first prove that the set of rational points is empty. As explained in §1.4.1, this will follow if we prove that the only point in $\mathbf{P}^2(\mathbf{Q})$ on

$$X^3 + Y^3 + 60Z^3 = 0$$

is $(X, Y, Z) = (1, -1, 0)$. By Proposition 1.4.1, this curve is birationally equivalent with the Weierstrass equation $y^2 = x^3 - 432 \cdot 60^2$. Replacing x, y with $4x, 8y$ and dividing the equation by 64, this becomes

$$E : y^2 = x^3 - 2^2 3^5 5^2.$$

By Nagell-Lutz, $\mathcal{T} = O$, and we wish to prove that $r = 0$.

This will follow from the proposition using part (b2), but this time the curve is only quasi-supersingular and not supersingular at 2.

We take $\pi = \sqrt[3]{30}$ and calculate $D_f = -2^4 3^{13} 5^4$, $\lambda = 3\pi^2$, $L = \mathbf{Q}(\lambda) = \mathbf{Q}(\pi)$. From [Cas50, Table 1], an integral basis is $1, \pi, \pi^2$, hence $D_L = -2^2 3^5 5^2 (= a_6!)$ and $I = 2 \cdot 3^4 5$, a positive fundamental unit is $\epsilon = 1 + 9\pi - 3\pi^2$ and the class number $h = 3$.

These data give $n_L = 1$, $n_p = n_q = n_h = 0$, hence part (a) of the proposition gives $r \leq 1$, and it remains to show that E is quasi-supersingular at 2. (QSS1)–(QSS3) are clear from the preceding. We now verify (QSS4) with $N = 1$.

Let $(x, y) \in E(\mathbf{Q})$. Clearly $x \neq 0$. Suppose $x = 2^a x_1$ where $a > 0$ and $v(x_1) = 0$. Then $y = 2y_1$ where $v(y_1) = 0$ and

$$y_1^2 = 2^{3a-2} x_1^3 - 3^5 5^2.$$

If $a \geq 2$ we deduce the contradiction $1 \equiv -3 \bmod 8$; if $a = 1$ we deduce $1 \equiv 2x_1 - 3 \bmod 8$, hence the contradiction $x_1 \equiv 2 \bmod 4$. We conclude that $v(x) \leq 0$. This completes the proof that $E(\mathbf{Q}) = O$.

It remains to prove that the Selmer curve has points defined over every local field K.

$K = \mathbf{R}$: no problem.

$K = \mathbf{Q}_2$: take $V = 0$, $W = 1$ and apply Hensel to $f(U) = 3U^3 + 5$ starting at $U = 1$.

$K = \mathbf{Q}_3$: take $U = 0$, $W = -1$ and apply Hensel to $f(V) = 4V^3 - 5$ starting at $V = 2$.

$K = \mathbf{Q}_5$: take $V = -1$, $W = 0$ and apply Hensel to $f(U) = 3U^3 - 4$ starting at $U = 2$.

$K = \mathbf{Q}_p$, $p > 5$: use Hensel's lemma starting with a point (U, V, W) defined over $\mathbf{F}_p$ with not all three of $U, V, W \equiv 0 \bmod p$, whose existence is guaranteed by the following lemma.

Lemma 3.7.9 *Let q be a prime power and $a, b, c \in \mathbf{F}_q$. Then $\mathbf{P}^2(\mathbf{F}_q)$ contains a point on*

$$aU^3 + bV^3 + cW^3 = 0. \qquad (\sharp)$$

Remark. This can be regarded as another example of the theorem of Schmidt mentioned in the remarks after Proposition 3.6.7 (together with degenerate cases where $3abc = 0$). However we have the following direct and elegant proof by M. Hall, Jr., as reported in [Sel54, p.218].

Proof. If $abc = 0$, say $a = 0$, we can take $U = 1, V = W = 0$. Thus suppose $abc \neq 0$. If $q \equiv 0$ or $2 \bmod 3$ then each element of $\mathbf{F}_q$ has a unique cube root, and therefore there is no shortage of solutions. Thus suppose $q \equiv 1 \bmod 3$. Then the set of nonzero cubes form a subgroup K_0 of ${\mathbf{F}_q}^*$ of order $\overline{q} = (q-1)/3$; we denote the cosets by K_1, K_2.

If two of a, b, c are in the same coset, say a and b, then $a = bv^3$ for some v and we can take $(U, V, W) = (1, -v, 0)$. The hard case is when the three coefficients lie in the three cosets, say $a \in K_0$, $b \in K_1$ and $c \in K_2$. We want to prove there are $x_i \in K_i$ such that $x_0 + x_1 + x_2 = 0$, for there are U, V, W such that $x_0 = aU^3$, $x_1 = bV^3$ and $x_2 = cW^3$.

We define a 3×3 matrix of integers by

$$\alpha_{ij} = \#\{x \in K_i : x + 1 \in K_j\}.$$

We have

$$\alpha_{00} + \alpha_{01} + \alpha_{02} = \overline{q} - 1$$

since the one member -1 of K_0 is 'lost' due to $-1 + 1 = 0$. Otherwise the $\overline{q}$ members of K_i when augmented by 1 are distributed without loss among the three K_j:

$$\alpha_{10} + \alpha_{11} + \alpha_{12} = \overline{q}, \quad \alpha_{20} + \alpha_{21} + \alpha_{22} = \overline{q}.$$

Thus

$$\alpha_{10} + \alpha_{11} + \alpha_{12} = \alpha_{00} + \alpha_{01} + \alpha_{02} + 1. \tag{$\flat$}$$

Since $x \mapsto -x$ effects a permutation on each of the sets K_i, and

$$x_j = x_i + 1 \Longrightarrow -x_i = -x_j + 1,$$

where x_i denotes a typical member of K_i, therefore the matrix is symmetric:

$$\alpha_{ji} = \alpha_{ij}.$$

The map $x \mapsto x^{-1}$ defines a pemutation on K_0 and interchanges K_1 and K_2. We interpret subscripts $\bmod 3$; thus $x_i^{-1} \in K_{-i} = K_{3-i}$, and $x_i^{-1}x_j \in K_{j-i}$. Since

$$x_j = x_i + 1 \Longrightarrow x_i^{-1}x_j = 1 + x_i^{-1},$$

we have $\alpha_{ij} = \alpha_{-i,j-i}$, which implies in particular that $\alpha_{11} = \alpha_{20} = \alpha_{02}$. Thus ($\flat$) becomes

$$\alpha_{12} = \alpha_{00} + 1 \geq 1.$$

This means there are x_1, x_2 such that $x_2 = x_1 + 1$, or $x_0 + x_1 + (-x_2) = 0$ where $x_0 = 1$. ■

For an outline of a different approach to the corollary, see [Cas66, p.206].

In [Sel51], Selmer gives extensive theory and tables of results concerning diophantine equations of the type $aU^3 + bV^3 + cW^3 = 0$ defined over $\mathbf{Q}$. The examples are not always as 'easy' as the one we have just worked out. *E.g.*, he shows that

$$U^3 + 10V^3 + 33W^3 = 0$$

has no point in $\mathbf{P}^2(\mathbf{Q})$, yet has points defined over every local field as in the example in our corollary, but this time $X^3+Y^3+330Z^3 = 0$ has infinitely many points. (In Proposition 1.4.4 we have a rational map going in one direction only.) Nagell's algorithm converts the latter cubic to the Weierstrass form $y^2 = x^3 - 3^5 5^2 11^2$ which has the point $(91, 136)$ of infinite order. Thus the method of proof of the preceding corollary does not apply to this example.

Chapter 4

Twists

Taking *elliptic curve* to mean a Weierstrass equation, we define the notion of isomorphism betwen elliptic curves and describe simplified representatives of the isomorphism classes. When elliptic curves 'become' isomorphic over some extension field they are called *twists* of one another; we will see that this is so iff they have the same j-invariant. The last section considers elliptic curves over finite fields, studiously postponing (to Chapter 6) the proof of the Riemann Hypothesis for $E_{/\mathbf{F}_q}$ to emphasize the elementary nature of the results, in particular Corollary 4.7.11 which was previously thought to be rather more difficult to obtain. For the same reason, the introduction of Galois cohomology is postponed to a later chapter.

4.1 Isomorphisms of elliptic curves

Let E be an elliptic curve defined over the field K by the Weierstrass equation $F = y^2 + a_1xy + \cdots - a_6 = 0$ with function field $L = K(x, y)$ and point O at ∞. Recall from §2.2.2 that associated to O is a valuation w_∞ in $\mathrm{gam}_K(L)$ characterized by the property that x has negative value, and in fact $w_\infty(x) = -2$, $w_\infty(y) = -3$. If E' is another elliptic curve over K with data $L' = K(x', y')$, a'_i, ${w_\infty}'$, we wish to study K-isomorphisms, that is, K-algebra isomorphisms, $L \longrightarrow L'$ that preserve O in the sense that the image of x has negative ${w_\infty}'$-value. (Later, once we have defined morphisms between algebraic varieties, we will properly describe this as an isomorphism $f : E' \longrightarrow E$ preserving O.) To simplify notation we take $L = L'$ and seek $x', y' \in L$ with appropriate properties.

For later purposes it is convenient to include in the discussion the invariant differential encountered in Chapter 2:

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{dx}{F_y(x, y)} = \frac{-dy}{F_x(x, y)} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}.$$

Proposition 4.1.1 *With E, L, w_∞ as above and $x', y' \in L$, the following are equivalent:*

(i) x', y' satisfy a Weierstrass equation $y'^2 + a_1'x'y' + \cdots - a_6' = 0$, $w_\infty(x') < 0$ and $K(x', y') = L$;

(ii) there are $r, s, t \in K$ and $u \in K^*$ such that

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

equivalently,

$$\begin{aligned} u^2x' &= x - r, \\ u^3y' &= y - sx + sr - t. \end{aligned}$$

Then

$$u^3(2y' + a_1'x' + a_3') = 2y + a_1x + a_3, \qquad u^{-1}\omega' = \omega,$$

and

$$\begin{aligned} ua_1' &= a_1 + 2s \\ u^2a_2' &= a_2 - sa_1 + 3r - s^2 \\ u^3a_3' &= a_3 + ra_1 + 2t \\ u^4a_4' &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st \\ u^6a_6' &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1 \\ u^8b_8' &= b_8 + 3rb_6 + 3r^2b_4 + r^3b_2 + 3r^4 \\ u^6b_6' &= b_6 + 2rb_4 + r^2b_2 + 4r^3 \\ u^4b_4' &= b_4 + rb_2 + 6r^2 \\ u^2b_2' &= b_2 + 12r \end{aligned}$$

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

Proof. If x' and y' are defined as in (ii) then obviously (i) is true. Conversely assume that x', y' satisfy the conditions in (i). Proposition 2.2.9(a) applied to the Weierstrass equation satisfied by x', y' implies that w_∞ is the only valuation in $\mathrm{gam}_K L$ for which x' has negative value, in fact the values of x', y' are $-2, -3$ respectively. Therefore for all $w \neq w_\infty$, $w(x') \geq 0$, $w(y') \geq 0$ hence by Proposition 2.2.10

$$x', y' \in \bigcap_{w \neq w_\infty} V(w) = \bigcap A_P = A = K[x, y].$$

Thus $x' = \alpha_1 y + \alpha_2$, $y' = \alpha_3 y + \alpha_4$ for some $\alpha_i \in K[x]$. Now for $\alpha_i \neq 0$, $w_\infty(\alpha_i) = -2\deg\alpha_i$ is even and $w_\infty(\alpha_i y) = -2\deg\alpha_i - 3$ is odd. Therefore $w_\infty(x') = \min\{w_\infty(\alpha_1 y), w_\infty(\alpha_2)\} = -2$ forces $\alpha_1 = 0$, $\deg\alpha_2 = 1$. Similarly we see that $\deg\alpha_3 = 0$ and $\deg\alpha_4 \leq 1$ or $\alpha_4 = 0$. Substituting into the

Weierstrass equation for x', y' shows that the leading coefficients of α_2, α_3 can be written as u^{-2}, u^{-3} respectively for some $u \in K^*$.

The remaining formulas are best checked on the computer. ■

Let $\tau = [r, s, t, u]$ denote the transformation described in the proposition. The set of these transformations comprise a group G:

$$[r', s', t', u'][r, s, t, u] = [r + u^2 r', s + us', t + u^2 sr' + u^3 t', uu'],$$

$$[0, 0, 0, 1] \qquad \text{acts as } 1,$$

and
$$\tau^{-1} = [-u^{-2} r, -u^{-1} s, u^{-3}(rs - t), u^{-1}].$$

The group G acts on the set $\mathcal{W}$ of Weierstrass equations, which for the present discussion can be identified with the set $K^5 = \{(a_1, a_2, a_3, a_4, a_6)\}$, according to the formulas of the proposition:

$$[r, s, t, u](a_1, a_2, a_3, a_4, a_6) = (u^{-1}(a_1 + 2s), \ldots) .$$

The set $\mathcal{W}$ is partitioned as $\mathcal{E} \cup \mathcal{S}$, where $\mathcal{E}$ consists of the non-singular Weierstrass equations — those for which $\Delta \neq 0$ — and $\mathcal{S}$ consists of the singular equations. G acts separately on $\mathcal{E}$ and $\mathcal{S}$.

For example $\tau = [0, 1, 4, 2]$ transforms

$$E = (0, 1, 0, 8, 16), \quad \textit{i.e.,} \quad y^2 = x^3 + x^2 + 8x + 16 \quad \text{with } \Delta = -2^{13} * 13$$

to
$$\tau E = (1, 0, 1, 0, 0), \quad \textit{i.e.,} \quad y^2 + xy + y = x^3 \quad \text{with } \Delta = -26.$$

However we will not continue to use the $(a_1, \ldots)$ notation, but rather use expressions such as 'let E be the equation $y^2 = x^3 + \cdots$'.

At this point we take as the definition of **elliptic curve** (**defined over K**) a Weierstrass equation $E \in \mathcal{E}$ (whose coefficients $a_1, \ldots, a_6$ lie in K), and we define the transformations τ described in the proposition as **isomorphisms** from one elliptic curve to another. Thus two elliptic curves are **isomorphic over K**, or are **K-isomorphic**, when they fall in the same G-orbit of $\mathcal{E}$. Typical notation is $\tau E = E'$. When the identity of K is clear, we say simply that E and E' are isomorphic. The orbit is an **abstract elliptic curve**, and the individual members of the orbit are referred to as **models** of the abstract elliptic curve; these terms are qualified by *defined over* K when necessary.

These isomorphisms have a concrete meaning in terms of points on the curves.

Proposition 4.1.2 *Let E and $[r, s, t, u]E = E'$ be isomorphic elliptic curves defined over the field K. Then*

$$(a, b) \mapsto (a', b') = (u^{-2}(a - r), u^{-3}(b - sa + sr - t))$$

augmented by $O_E \mapsto O_{E'}$ defines a group isomorphism

$$E'(K) \longrightarrow E(K).$$

Proof. At some level all of this is 'obvious'; but let us attempt to give a reasoned account.

Let X, Y be independent transcendentals over K and let F be the polynomial $Y^2+a_1XY+a_3Y-X^3-a_2X^2-a_4X-a_6$, so that the affine coordinate ring of E is $A = K[X,Y]/(F)$ where the denominator (F) is the principal ideal generated by F. The statement that $(a,b) \in E(K)$ is equivalent to the statement that the ideal $(X-a, Y-b)$ generated by the two polynomials $X-a$ and $Y-b$ contains (F) (so that the surjective K-algebra homomorphism $K[X,Y] \longrightarrow K$ determined by $X \mapsto a$, $Y \mapsto b$ factors through A).

Define the polynomials

$$X' = u^{-2}(X-r), \qquad Y' = u^{-3}(Y - sX + sr - t);$$

then

$$X = u^2X' + r, \qquad Y = u^3Y' + su^2X' + t,$$

and $K[X',Y'] = K[X,Y]$. When we substitute these expressions for X, Y into F, divide by u^6 and collect terms, we obtain a polynomial F' in X', Y' that is the Weierstrass polynomial for E'. Since u is a unit, $(F') = (F)$. Also

$$(X'-a', Y'-b') = (u^{-2}(X-a), u^{-3}(Y-b-s(X-a))) = (X-a, Y-b).$$

Thus $(a,b) \in E(K) \Leftrightarrow (a',b') \in E'(K)$, and so we at least have a map $E(K) \longrightarrow E'(K)$, which we denote $\overline{\tau}$. Since $O_{E'} \mapsto O_E$ and $(a',b') \mapsto (u^2a'+r, u^3b' + su^2a' + t)$ define an inverse, $\overline{\tau}$ is a bijection.

Since the transformations between the X, Y and X', Y' coordinates are linear, colinearity of points is preserved, hence $\overline{\tau}$ is a homomorphism. In detail, we pass to projective coordinates: $X = U/W$, $Y = V/W$, and $X' = U'/W'$, $Y' = V'/W'$, so the transformations are $U = u^2U'+rW$, $V = u^3V'+su^2U'+tW'$ and $W = W'$. Substitution shows that if (a_i, b_i, c_i), $i = 1,2,3$ are the points of intersection of the line $\alpha U + \beta V + \gamma W = 0$ with E, then their images (a_i', b_i', c_i') are the points of intersection of $\alpha'U' + \beta'V' + \gamma'W' = 0$ with E', where $\alpha' = u^2(\alpha + s\beta)$, $\beta' = u^3\beta$ and $\gamma' = r\alpha + t\beta + \gamma$ are not all zero. ■

The following obvious corollary is worth stating explicitly.

Corollary 4.1.3 *Isomorphic $E_{/K}$ have isomorphic m-torsion subgroups $E(K)[m]$ for every positive integer m.* ■

The subgroup $\{\tau : \tau E = E\}$ of G that stabilizes E is the **automorphism group** of E (over K) and is denoted $\mathrm{aut}_K E$. The automorphism groups of isomorphic E are conjugate subgroups of G. The group $\mathrm{aut}_K E$ always contains at least the two elements $[1]$ and $[-1] = [-1]_E = [0, -a_1, -a_3, -1]$. [‡] Note that as elements of G, $[-1] \neq [1]$, (even when $\mathrm{char}\, K = 2$ because then $\Delta \neq 0$ implies that at least one of $a_1, a_3 \neq 0$), but their action on E is identical.

[‡]Later we will see that $\mathrm{aut}_K E$ is the group of invertible elements in a ring $\mathrm{end}_K E$; for the time being we can think of $[-1]$ as the endomorphism of the abelian group $E(L)$ given by $[-1](x,y) = (x, -y - a_1x - a_3)$.

4.2 Simplified Weierstrass equations

For general K there is no canonical way to pick models of elliptic curves, that is, particular representatives of the G-orbits, but at least each orbit contains a subset of simplified forms as we now describe. We also describe the stabilizer of this subset in terms of the following subgroups of G:

$$\begin{aligned} R &= \{[r,0,0,u]\}, \\ S &= \{[0,s,0,1]\}, \\ T &= \{[s^2,s,t,u]\}, \text{ (characteristic 2 only)}, \\ U &= \{[0,0,0,u]\}. \end{aligned}$$

where r, s, t run through K and u through K^*.

We have the semidirect product decomposition

$$1 \longrightarrow G_1 \longrightarrow G \underset{s}{\overset{p}{\rightleftarrows}} K^* \longrightarrow 1$$

where $G_1 = \{[r,s,t,1]\}$ and the projection $p[r,s,t,u] = u$ is split by the section $s(u) = [0,0,0,u]$; *cf.* [Con82], p.401. U is isomorphic (via s) with the multiplicative group K^* and S with the additive group K^+, while R and T are noncommutative being describable as semidirect products.

Lemma 4.2.1 *Let* $\operatorname{char} K \neq 2$ *or* 3 *and let* E *be defined over* K *with covariants* c_4, c_6. *Then* E *is* K*-isomorphic to its c-form*

$$E_1\colon\ y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864}.$$

Hence if E' *is also defined over* K, *with covariants* c_4', c_6', *then* E *and* E' *are* K*-isomorphic iff* $\exists u \in K^*$ *such that*

$$c_4 = u^4 c_4', \qquad c_6 = u^6 c_6'.$$

Remarks. The assumption $1728 = 2^6 3^3 \neq 0$ in K and the relations

$$c_4^3 - c_6^2 = 1728\Delta, \qquad {c_4'}^3 - {c_6'}^2 = 1728\Delta'$$

mean that any two of the equations

$$c_4 = u^4 c_4', \quad c_6 = u^6 c_6', \quad \Delta = u^{12}\Delta'$$

implies the third.

For alternative criteria that E and E' be K-isomorphic, see Proposition 4.4.1.

Proof. First complete the square in y and then complete the cube in x:

$$[-b_2/12, 0, 0, 1][0, -a_1/2, -a_3/2, 1]E = E_1.$$

This is the change to (ξ, η) coordinates mentioned at the beginning of chapter 1.

If E and E' are K-isomorphic, say $[r, s, t, u]E = E'$, then the relations between their covariants follow from the transformation equations of the previous proposition.

Conversely assume these relations. Then E is K-isomorphic to

$$\begin{aligned} [0,0,0,u]E_1\colon\ y^2 &= x^3 - \frac{u^4 c_4}{48}x - \frac{u^6 c_6}{864} \\ &= x^3 - \frac{c_4'}{48}x - \frac{c_6'}{864} \end{aligned}$$

which in turn is K-isomorphic to E'. ■

Proposition 4.2.2 *For each $E \in \mathcal{E}$ there exists a $\tau \in G$ such that the equation for $E' = \tau E$ has the following simplified form according to the case indicated in the boxed assumptions. A particular τ (which is not unique) is given in each case.*

The subgroup stabilizing the class of simplified forms is denoted stab *in each case; for classes* (a), (b) and (d), *for each allowed value j_0 of j, the group* stab *is also the subgroup stabilizing the subset of that class with $j = j_0$.* †

(a) $\boxed{\operatorname{char} K \neq 2, 3}$

$$y^2 = x^3 + a_4'x + a_6', \qquad c_4' = c_4 = -48a_4', \qquad c_6' = c_6 = -864a_6',$$

$$\text{hence} \qquad \Delta' = -16(4{a_4'}^3 + 27{a_6'}^2), \qquad j = -48^3{a_4'}^3/\Delta';$$

$\tau = [-b_2/12, 0, 0, 1][0, -a_1/2, -a_3/2, 1]$; stab $= U$;

(b) $\boxed{\operatorname{char} K = 3 \textit{ and } j \neq 0, \text{ i.e., } E \textit{ ordinary}}$

$$y^2 = x^3 + a_2'x^2 + a_6', \qquad c_4' = c_4 = {a_2'}^2, \qquad c_6' = c_6 = -{a_2'}^3,$$

$$\text{hence} \qquad \Delta' = -{a_2'}^3 a_6', \qquad j = -{a_2'}^3/a_6';$$

$\tau = [-b_4/b_2, 0, 0, 1][0, a_1, a_3, 1] = [-b_4/b_2, a_1, a_3 - a_1b_4/b_2, 1]$; stab $= U$;

(c) $\boxed{\operatorname{char} K = 3 \textit{ and } j = 0, \text{ i.e., } E \textit{ supersingular}}$

$$y^2 = x^3 + a_4'x + a_6', \qquad c_4' = c_4 = c_6' = c_6 = j = 0, \qquad \Delta' = -{a_4'}^3;$$

†For example, the subset of class (a) with $j = 1728$ consists of the equations $y^2 = x^3 + a_4x$, $a_4 \in K^*$ and the stabilizing subgroup of this subset is again U. This observation is used in the proof of Proposition 4.7.6 where we count E defined over a finite field with given j. We note also that all these curves have a point of order 2, namely $(0, 0)$. In fact, an elliptic curve of invariant 1728 has a point of order 2 either always, sometimes, or never according as $\operatorname{char} K$ is, respectively, not 2 or 3, 3, or 2.

this simplified form exists with $a_6' = 0$ *iff* $E(K)$ *contains a point of order* 2;[‡] $\tau = [0, a_1, a_3, 1]$*;* stab $= R$*;*

(d) $\boxed{\text{char}\, K = 2 \textit{ and } j \neq 0, \text{ i.e., } E \textit{ ordinary}}$

$$y^2 + xy = x^3 + a_2'x^2 + a_6', \quad c_4' = c_6' = 1, \quad \Delta' = a_6', \quad j = 1/a_6';$$

$\tau = [a_3/a_1, 0, a_4/a_1 + a_3^2/a_1^3, a_1]$*;* stab $= S$*;*

(e) $\boxed{\text{char}\, K = 2 \textit{ and } j = 0, \text{ i.e., } E \textit{ supersingular}}$

$$y^2 + a_3'y = x^3 + a_4'x + a_6', \qquad c_4' = c_6' = j = 0, \qquad \Delta' = {a_3'}^4.$$

$\tau = [a_2, 0, 0, 1]$*;* stab $= T$*;*

Moreover, for each of the five classes, stab *has the following strong stabilizing property: if* $\sigma \in G$ *is such that* $\sigma E = E'$ *for any particular pair* E, E' *in that class then* $\sigma \in$ stab.

Proof. (a) The existence of the simplified form is given by the lemma. Clearly any element $\sigma \in U$ stabilizes this class and also stabilizes any subclass with j fixed since E and τE have the same j-invariant by the previous proposition. If E and E' are two such simplified forms and $\sigma = [r, s, t, u]$ is an isomorphism from E to E', then, since char $K \neq 2, 3$, the transformation equations give in succession $2s = 0$, hence $s = 0$, then $3r = 0$, hence $r = 0$, then $2t = 0$, hence $t = 0$. This proves $\sigma \in U$ and the strong stabilizing property. The determination of stab and the verification of the strong stabilizing property in the remaining cases is equally straightforward, so the details will be omitted. Similarly $\tau E = E'$ for the stated τ in each case is left to the reader.

(b) and (c) Since char $= 3$ we can at least complete the square in y and so assume $a_1' = a_3' = 0$. Then $c_4' = {a_2'}^2$ hence $j := {c_4'}^3/\Delta$ is $0 \Leftrightarrow a_2' = 0$. If $a_2 \neq 0$, $[a_4/a_2, 0, 0, 1]$ yields $a_4' = 0$.

Assuming the form (c), if $a_6' = 0$ then $(0, 0)$ is a point of order 2. Conversely if $(r, 0)$ is a point of order 2 then, since char $K = 3$, $[-r, 0, 0, 1]$ gives a new equation of the same type but with $a_6' = 0$.

(d) and (e) Since char $= 2$ we have $c_4 = a_1^4$, hence $j = 0 \Leftrightarrow a_1 = 0$. All that remains is to verify that $\tau E = E'$ in the two cases. ■

Since we will have occasion to refer to these special Weierstrass equations, we make a formal definition. A Weierstrass equation of one of the types listed in the foregoing proposition is a **simplified** Weierstrass equation of **class** (a), (b), (c), (d) or (e).

[‡]By Corollary 4.1.3, the existence of a point of order 2 does not depend on the equation chosen in the isomorphism class of E.

4.3 Twists, quadratic and otherwise

An elliptic curve defined over K can be regarded as being defined over any extension field F of K.

Proposition 4.3.1 *For elliptic curves E, E' defined over K with respective invariants j, j', the following are equivalent*

- *E and E' are F-isomorphic for some extension F of K;*
- *E and E' are $\overline{K}$-isomorphic where $\overline{K}$ denotes the algebraic closure of K;*
- *$j = j'$.*

Remark. It will be seen in the proof that usually F can be taken to be a quadratic extension, and in any case a finite extension of K. This will be made explicit in Proposition 4.4.1.

Proof. If $\tau E = E'$ with τ defined over F then $j = j'$ by Proposition 4.1.1 (with F written in for K) and, incidentally, $\tau \in \mathrm{stab}$ as classified in the previous proposition, by the strong stabilizing property.

Conversely let $j = j'$. We will construct a $\overline{K}$-isomorphism $\tau = [r, s, t, u]$ from E to E'. (This construction will be examined more closely in Proposition 4.5.1; at the moment we only wish to prove existence.) We can assume that E and E' are both in simplified form, and the assumption $j = j'$ assures that they are in the same class. We treat the five classes separately. For each class we make a choice of $\tau = [r, s, t, u] \in \mathrm{stab}$ that satisfies the five transformation equations $ua'_1 = a_1 + 2s$ through $u^6 a'_6 = a_6 + \text{etc.}$

(a) By the formulas for j and $j - 1728$ given in § 1.2,

$$j = j' \implies c_4^3/\Delta = {c'_4}^3/\Delta', \quad j - 1728 = j' - 1728 \implies c_6^2/\Delta = {c'_6}^2/\Delta',$$

$$\text{hence} \qquad {c'_4}^3 c_6^2 = c_4^3 {c'_6}^2.$$

We can choose $u \in \overline{K}$ so that

$$c_4 = u^4 c'_4, \qquad c_6 = u^6 c'_6;$$

this is clear if either $c_4 = c'_4 = 0$ or $c_6 = c'_6 = 0$ and otherwise take

$$u = \sqrt{\frac{c'_4 c_6}{c_4 c'_6}}.$$

That $\tau = [0, 0, 0, u]$ is an isomorphism from E to E' was already observed in the proof of Lemma 4.2.1

(b) Since E and E' are in simplified form we have $a_i = a'_i = 0$ for $i = 1, 3, 4$, and all of a_2, a'_2, a_6, a'_6 are nonzero since the discriminants are nonzero. Also

$$j = j' \implies -a_3^2/a_6 = -{a'_3}^2/a'_6.$$

For $\tau = [0,0,0,u] \in \text{stab} = U$ the transformation equations reduce to

$$u^2 a_2' = a_2, \qquad u^6 a_6' = a_6$$

which have the solution $u = \sqrt{a_2/a_2'}$.

(c) This time, using $3 = 0$ in K, the equations reduce to

$$u^4 a_4' = a_4 \ (\neq 0) \quad \text{and} \quad u^6 a_6' = a_6 + r a_4 + r^3$$

which can be solved successively for $u, r \in \overline{K}$ to give $\tau = [r,0,0,u] \in R$.

(d) Since $j = j' \Rightarrow a_6 = a_6'$, there is only one equation that is not automatically satisfied: $a_2' = a_2 + s + s^2$. Taking either root $s \in \overline{K}$ gives $\tau = [0,s,0,1] \in S$.

(e) We have $a_i = a_i' = 0$ for $i = 1,2$, and $2 = 0$, and we want $\tau = [s^2, s, t, u] \in T$. The equations to be satisfied are

$$\begin{aligned} u^3 a_3' &= a_3 \ (\neq 0), \\ u^4 a_4' &= a_4 + s a_3 + s^4 \\ u^6 a_6' &= a_6 + s^2 a_4 + s^6 + t a_3 + t^2. \end{aligned}$$

These can be solved successively for $u, s, t \in \overline{K}$. ■

Two elliptic curves E, E' defined over K satisfying the conditions of the above proposition are **twists** (of each other).

Example. Let K be a function field with constant field K_0, and let E be defined over K. Then E is said to be **constant** when $j \in K_0$, and in that case for some finite extension F of K, E is F-isomorphic to an E' defined over K_0. An instance of this occured in Example 1 following Proposition 2.10.3; there $F = \mathbf{C}(t,t')$ was of degree 4 over $K = \mathbf{C}(t)$.

Given E with invariant j, the smallest field over which there is defined a twist is $K_0(j)$ where K_0 denotes the prime subfield of K. For j is a rational function in the coefficients $a_1, \ldots$, and so the field must contain $K_0(j)$. Conversely if $j \neq 1728, 0$ the "generic-j" curve

$$y^2 + xy = x^3 - \frac{36}{j - 1728} x - \frac{1}{j - 1728}$$

is defined over this field and has j-invariant $= j$. When $j = 0$, 1728 take $y^2 = x^3 + 1$, $y^2 + y = x^3 + x$ respectively (the latter when $\operatorname{char} K = 2$ or 3).

If $\tau = [r,s,t,u]$ we use the convenient abbreviation $K(\tau) := K(r,s,t,u)$. Clearly

$$K(\tau^{-1}) = K(\alpha\tau) = K(\tau\alpha) = K(\tau) \quad \forall \alpha \text{ defined over } K. \qquad ¶$$

When $E' = \tau E$, where E' and E are defined over K, the transformation equations imply that $K(\tau)$ is a finite extension of K: $u^{12} = \Delta/\Delta' \in K^*$, and then,

at least when $\operatorname{char} K \neq 2, 3$, the equations for a_1, a_2, a_3 show in succession that $s, r, t \in K(u)$. When $\operatorname{char} K = 2$ or 3 one needs to use the other equations; a careful analysis of the extension $K(\tau)/K$ is made in Proposition 4.5.1 below.†

When E and E' are twists we call the minimal $[K(\tau)\colon K]$ for an isomorphism $\tau E = E'$ the **degree** of the twist. Thus there are twists of degree 1, in which case E and E' are K-isomorphic, twists of degree 2, and, as we will see in the next proposition, possibly higher degree twists only when $j = 0$ or 1728. We define **quadratic twist** to mean a twist of degree ≤ 2: it is convenient to allow the possibility that E and E' are already isomorphic over K. Most quadratic twists can be constructed as follows.

If $\boxed{\operatorname{char} K \neq 2}$, we can take a b-form for E in the notation

$$y^2 = x^3 + a_2x^2 + a_4x + a_6.$$

For $d \in K^*$, the **quadratic twist by *d*** of E, denoted by any one of $d * E$, E^d, $E * d$ is given by

$$y^2 = x^3 + da_2x^2 + d^2a_4x + d^3a_6\ ; \tag{s}$$

replacing x, y with dx, d^2y gives the equivalent quasi-Weierstrass form

$$dy^2 = x^3 + a_2x^2 + a_4x + a_6.$$

The latter form makes it obvious that E and E^d are isomorphic over $K(\sqrt{d})$. For a full-blown Weierstrass equation, the first version works out to (see the proof of the next proposition for details)

$$y^2 + a_1xy + a_3y = \\ x^3 + (a_2d + a_1^2(d-1)/4)x^2 + (a_4d^2 + a_1a_3(d^2-1)/2)x + \\ a_6d^3 + a_3^2(d^3-1)/4.$$

If $\boxed{\operatorname{char} K = 2}$ the definition one should adopt, now for any $d \in K$, is

$$y^2 + a_1xy + a_3y = x^3 + (a_2 + a_1^2d)x^2 + a_4x + a_6 + a_3^2d.$$

This assumes a simpler appearance when E is in the simplified form of class (d) or (e):

If $\qquad E : y^2 + xy = x^3 + a_2x^2 + a_6 \quad$ or $\quad y^2 + a_3y = x^3 + a_4x + a_6,$

then

$$E^d : y^2 + xy = x^3 + (a_2 + d)x^2 + a_6 \quad \text{or} \quad y^2 + a_3y = x^3 + a_4x + a_6 + a_3^2d. \tag{s$'$}$$

†The set of isomorphisms τ from E to E' defined over any extension field of K is the (reducible) algebraic variety in affine 4-space consisisting of the solutions $[r, s, t, u]$ of the set of 5 transformation equations for $a_1, \ldots, a_6$. Since $a_1, \ldots, a_1', \ldots$ are all in K, these solutions are all algebraic over K, so this variety is 0-dimensional and consists of finitely many points.

Proposition 4.3.2 *Let E be an elliptic curve over the field K and $d \in K$, with $d \neq 0$ when* $\operatorname{char} K \neq 2$. *Let ϑ be a root of*

$$\vartheta^2 = 1/d \quad \textit{if } \operatorname{char} K \neq 2, \qquad \vartheta^2 + \vartheta = d \quad \textit{if } \operatorname{char} K = 2,$$

so $K(\vartheta)$ is a separable extension of K of degree ≤ 2. Define

$$\tau = \begin{cases} [0, a_1(\vartheta - 1)/2, a_3(\vartheta^3 - 1)/2, \vartheta] & \textit{if char } K \neq 2, \\ [0, a_1\vartheta, a_3\vartheta, 1] & \textit{if char } K = 2. \end{cases}$$

(a) $\tau E = E^d$, *hence E^d is a quadratic twist of E.*

(b) *If* $\operatorname{char} K \neq 2, 3$, *and $j = 1728$, then E has an exceptional quadratic twist*† *E' : their simplified forms are*

$$E\colon \quad y^2 = x^3 + ax, \qquad E'\colon \quad y^2 = x^3 - 4ax$$

with isomorphism $[0, 0, 0, 1 + \sqrt{-1}]E' = E$.

(c) *All twists are quadratic except when $j = 1728$ or 0, and then there may be twists of higher degree. Except when* $\operatorname{char} K = 2$ *or* 3 *and $j = 0$, every quadratic twist of E is K-isomorphic to E^d for some $d \in K$, or to the exceptional twist E' in paragraph* (b).

(d) *If* $\operatorname{char} K \neq 2$ *(resp.* $\operatorname{char} K = 2$*), then $d \mapsto \tau$ defines a group homomorphism from the multiplicative group K^* (resp. the additive group K^+) to the group G. Thus*

$$\textit{if } \operatorname{char} K \neq 2, \ (E * d_1) * d_2 = E * (d_1 d_2), \quad (E * d) * (d^{-1}) = E, \quad E * 1 = E;$$

$$\textit{if } \operatorname{char} K = 2, \ (E * d_1) * d_2 = E * (d_1 + d_2), \quad (E * d) * (-d) = E, \quad E * 0 = E.$$

(e) *If E_1 and E_2 are K-isomorphic, then so are their twists E_1^d and E_2^d : if $\sigma E_1 = E_2$ where $\sigma = [r, s, t, u]$ and τ_i denotes τ as defined for E_i, $i = 1, 2$, then $\sigma^d E_1^d = E_2^d$ where*

$$\sigma^d = \tau_2 \sigma \tau_1^{-1} = \begin{cases} [rd, s, t + a_1 r(1 - d)/2, u] & \textit{if } \operatorname{char} K \neq 2, \\ \sigma \quad (!) & \textit{if } \operatorname{char} K = 2. \end{cases}$$

(f) *Suppose E has invariant $j \neq 1728$ or 0. Then $E * d_1$ and $E * d_2$ are K-isomorphic iff*

$$d_1 = \begin{cases} d_2 u^2 & \textit{when } \operatorname{char} K \neq 2 \ \textit{ for some } \ u \in K^*, \\ d_2 + s + s^2 & \textit{when } \operatorname{char} K = 2 \ \textit{ for some } \ s \in K. \end{cases}$$

---

†In terms to be explained later, this is a sort of "rational manifestation" of complex multiplication by i, the simplest example being the 2-isogeny

$$\mathbf{A32}\colon y^2 = x^3 - x \longrightarrow \mathbf{B32}\colon y'^2 = x'^3 + 4x', \quad \text{where}$$

$$x' = \frac{x^2 - 1}{x}, \qquad y' = \frac{x^2 + 1}{x^2} y.$$

Remarks. When $\operatorname{char} K \neq 2$ or 3, an E with $j = 1728$ is K-isomorphic with its c-form $E_{\mathrm{c}} : y^2 = x^3 + bx$, and a twist, in other words any E' with $j = 1728$, is K-isomorphic with an equation of the form $E'_{\mathrm{c}} : y^2 = x^3 + b'x$. Since $E'_{\mathrm{c}} = [0, 0, 0, \sqrt[4]{b/b'}]E_{\mathrm{c}}$, E and E' are isomorphic over the field $K(\sqrt[4]{b/b'})$. Hence the twist can be quadratic or of degree 4.

We note also that $E * (-1)$ is K-isomorphic with E since $E_{\mathrm{c}} * (-1) : y^2 = x^3 + (-1)^2 bx$ is unchanged. Thus paragraph (f) cannot be extended to include $j = 1728$.

The previous remarks also apply when $\operatorname{char} K = 3$ to E with $j = 0$ and a point of order 2; *cf.* Proposition 4.2.2.

Similarly, twists with $j = 0$ (still with $\operatorname{char} K \neq 2, 3$), which can be taken in the form $y^2 = x^3 + a_6$, are isomorphic over $K(\sqrt[6]{a_6/a'_6})$, and the twist can be quadratic, or of degree 3 or 6.

In general there are other twists of degree 2 when $\operatorname{char} K = 2$ or 3 and $j = 0$. These are the classes (c) and (e) of Proposition 4.2.2 where stab is a 2-parameter subgroup of G, and it is hard (unnatural?) to sort out which twists are quadratic.

Proof. (a) and (b): We will require the following notation. In the group G of K-isomorphisms, define

$$\lambda = \lambda_E = \begin{cases} [0, -a_1/2, -a_3/2, 1] & \text{if } \operatorname{char} K \neq 2, \\ [a_3/a_1, 0, (a_1^2 a_4 + a_3^2)/a_1^3, a_1] & \text{if } \operatorname{char} K = 2 \text{ and } j \neq 0, \\ [a_2, 0, 0, 1] & \text{if } \operatorname{char} K = 2 \text{ and } j = 0, \end{cases}$$

so that λE is in b-form or simplified form of type (d) or (e). Secondly, define

$$\mu = \begin{cases} [0, 0, 0, \vartheta] & \text{if } \operatorname{char} K \neq 2 \\ [0, \vartheta, 0, 1] & \text{if } \operatorname{char} K = 2; \end{cases}$$

Then calculation shows that $\tau = \lambda^{-1}\mu\lambda$. Thus λ brings the Weierstrass equation into b-form or simplified form, then μ effects the twist in the simplified form (s) or (s′), and finally λ^{-1} restores the twisted equation to a-form. Of course if E is in b-form or simplified form, then $\lambda = 1$ and $\tau = \mu$. In any case, $K(\tau) = K(\vartheta)$ is of degree ≤ 2 over K.

(c) Let E and $E' = \sigma E$ be twists.

Case 1: $j \neq 1728$ or 0 and $\operatorname{char} K \neq 2$. Applying the K-isomorphisms λ_E and $\lambda_{E'}$, we can assume that E and E' are already in b-form. The stabilizer R of the set of b-forms enjoys the "strong stabilizing" property: $\sigma E = E' \Longrightarrow \sigma \in R$, say $\sigma = [r, 0, 0, u]$ for r, u in some extension field. Since $j \neq 0$ or 1728, none of c_4, c_6, c'_4, c'_6 is 0, hence $u^4 c'_4 = c_4$ and $u^6 c'_6 = c_6$ imply that $u^2 \in K^*$. To complete the proof that the twist is quadratic in this case, we prove that $r \in K$. If $\operatorname{char} K \neq 3$ this follows from $u^2 b'_2 = b_2 + 12r$; if $\operatorname{char} K = 3$ the transformation equation for a_4, with $s = t = 3 = 0$, reduces to $u^4 a'_4 = a'_4 + 2ra_2$. Since $j \neq 0$ and $a_1 = 0$, therefore $a_2 \neq 0$, hence $r \in K$.

Writing $d = 1/u^2 \in K^*$, we have

$$\sigma = [ru^{-2}, 0, 0, 1][0, 0, 0, u] = [rd, 0, 0, 1]\mu,$$

hence $E' = [rd, 0, 0, 1]E^d$ is K-isomorphic withE^d.

Case 2: $j = 1728$ or 0 and $\operatorname{char} K \neq 2$ or 3. When $j = 1728$, resp. $j = 0$, by a K-isomorphism we can take E in c-form:

$$y^2 = x^3 + a_4 x,\ a_4 \neq 0, \quad \text{resp.} \quad y^2 = x^3 + a_6,\ a_6 \neq 0.$$

We can also take $E' = \tau E$ in c-form, and then the transformation equations imply that $\tau = [0, 0, 0, u]$ where $u^4 = a_4/a_4'$, resp. $u^6 = a_6/a_6'$. Assuming the twist is quadratic, we must ascertain when the polynomial $U^4 - a$, resp. $U^6 - a$, has a quadratic factor. There are the obvious cases

$$U^4 - d^2 = (U^2 - d)(U^2 + d), \quad U^6 - d^3 = (U^2 - d)(U^4 + dU^2 + d^2)$$

which correspond to twists of the form E^d, and just the one "nonobvious" case

$$U^4 + 4b^4 = (U^2 - 2bU + 2b^2)(U^2 + 2bU + 2b^2)$$

which gives rise to the exceptional quadratic twist.

Case 3: $j \neq 0$ and $\operatorname{char} K = 2$. Applying the K-isomorphisms $\lambda_{E'}$ and λ, we can assume that E and E' are in class (d) of Proposition 4.2.2. The transformation equations for $\sigma \in \operatorname{stab} = S$ reduce to $a_2' = a_2 + s + s^2,\ a_6' = a_6$, so again the twist is (separable) quadratic, and $E' = E * (a_2 + a_2')$.

(d) is clear from the formula $\tau = \lambda^{-1}\mu\lambda$ and the fact that $\mu = 1$ when $d = 1$, resp. $d = 0$, in the two cases.

(e) again is by direct calculation.

(f) First, let $\operatorname{char} K \neq 2$. Let c_4, c_6 denote the covariants of E and γ_4, γ_6, resp. γ_4', γ_6', those of $E * d_1$ and $E * d_2$. Then $\gamma_4 = c_4 d_1^2$, $\gamma_6 = c_6 d_1^3$, and similarly for γ_4', γ_6'.

If $E * d_2 = [r, s, t, u]E * d_1$ then $\gamma_4 = u^4\gamma_4'$ and $\gamma_6 = u^6\gamma_6'$. Since $j \neq 1728, 0$, neither of these quantities is 0, hence $d_1^2 = d_2^2 u^4$ and $d_1^3 = d_2^3 u^6$, and we conclude that $d_1 = d_2 u^2$.

Conversely for any $u, d \in K^*$, $[0, 0, 0, 1/u](E * d) = E * (du^2)$.

Second, let $\operatorname{char} K = 2$. Since $j \neq 0$, therefore $a_1 \neq 0$.

If $E * d_2 = [r, s, t, u](E * d_1)$, then the transformation equations

$$\begin{aligned} ua_1 &= a_1, \\ u^3 a_3 &= a_3 + ra_1, \\ u^2(a_2 + a_1^2 d_2) &= a_2 + a_1^2 d_1 + sa_1 + r + s^2 \end{aligned}$$

imply in succession that $u = 1$, $r = 0$ and $d_2 = d_1 + S + S^2$ where $S = s/a_1$.

Conversely for any $s, d \in K$, $[0, a_1 s, 0, 1](E * d) = E * (d + s + s^2)$. ■

4.4 The isomorphism algorithm

The two propositions in this section collect some facts in a convenient form that are implicit in the previous section, and do not contain any essentially new results.

We have just seen that it is very easy to decide when two elliptic curves E and E' defined over K with invariants j and j' are twists: iff $j = j'$. Obviously some additional condition is needed to insure that they are K-isomorphic.

Proposition 4.4.1 *Let E, E' be two elliptic curves defined over K; let unprimed data such as j refer to E, and primed data j' etc. to E'. Then E is K-isomorphic to E' iff $j = j'$ and the following condition is satisfied.*

(i) $\boxed{\operatorname{char} K \neq 2,\ j \neq 0 \text{ or } 1728}$

$$\sqrt{c_6/c_6'} \in K;$$

(ii) $\boxed{\operatorname{char} K \neq 2 \text{ or } 3,\ j = 1728}$

$$\sqrt[4]{c_4/c_4'} \in K;$$

(iii) $\boxed{\operatorname{char} K \neq 2 \text{ or } 3,\ j = 0}$

$$\sqrt[6]{c_6/c_6'} \in K;$$

(iv) $\boxed{\operatorname{char} K = 3,\ j = 0}$ *With E, and similarly E' in simplified form $y^2 = x^3 + a_4x + a_6$,*

$$K \text{ contains } u := \sqrt[4]{a_4/a_4'} \text{ and, for some choice of } u,$$

$$\text{a root of } \quad r^3 + a_4r + a_6 - u^6a_6' = 0;$$

(v) $\boxed{\operatorname{char} K = 2,\ j \neq 0}$ *With E, and similarly E' in simplified form* $y^2 + xy = x^3 + a_2x^2 + a_6$,

$$K \text{ contains a root } s \text{ of } s^2 + s + a_2 + a_2' = 0;$$

(vi) $\boxed{\operatorname{char} K = 2,\ j = 0}$ *With E and similarly E' in simplified form* $y^2 + a_3y = x^3 + a_4x + a_6$,

$$K \text{ contains } u := \sqrt[3]{a_3/a_3'} \text{ and, for some choice of } u, \text{ roots } s, t \text{ of}$$

$$\begin{cases} s^4 + a_3s + a_4 + u^4a_4', \\ t^2 + a_3t + s^6 + a_4s^2 + a_6 + u^6a_6'. \end{cases}$$

Proof. If there is an isomorphism $\tau = [r, s, t, u]\colon E \longrightarrow E'$ then by Proposition 4.1.1,

$$j = j', \quad c_4 = c_4'u^4, \quad c_6 = c_6'u^6.$$

We recall that

$$j = 0 \iff c_4 = 0 \text{ since } j = c_4^3/\Delta \text{ and}$$

$$j = 1728 \iff c_6 = 0 \text{ since } j - 1728 = c_6^2/\Delta.$$

(i) Since $j \neq 0, 1728$, none of c_4, c_4', c_6, c_6' is 0. If τ exists then $\sqrt{c_6/c_6'} = u^3 \in K$. Conversely suppose $j = j'$ and $\sqrt{c_6/c_6'} = a^2$, $a \in K^*$. Then

$$j = \frac{c_4^3}{\Delta} = \frac{{c_4'}^3}{\Delta'}, \quad j - 1728 = \frac{c_6^2}{\Delta} = \frac{{c_6'}^2}{\Delta'},$$

hence if we set $b = c_4'c_6/c_4c_6'$ then $c_4 = c_4'b^2$, $c_6 = c_6'b^3$. It follows that $b = u^2$ where $u = a/b$. When $\operatorname{char} K \neq 3$ (resp. $= 3$) we take E and E' in the simplified form (a) (resp. (b)) of Proposition 4.2.2. Then (in both situations) $[0, 0, 0, u]$ is the required isomorphism.

(ii) If τ exists we have $c_4 = c_4'u^4 \neq 0$, so $\sqrt[4]{c_4/c_4'} \in K$. Conversely suppose $j = j'$ and $\sqrt[4]{c_4/c_4'} = u$, $u \in K$. Taking E and E' in the simplified form (a), with $a_6 = a_6' = 0$, $[0, 0, 0, u]$ is the required isomorphism.

The proof of (iii) is similar to that of (ii).

In cases (iv), (v) and (vi) we can assume that E and E' are in the simplified forms of (c), (d) and (e) respectively of Proposition 4.2.2. Then the conditions are just the obviously necessary and sufficient conditions for the existence of τ given by the transformation equations of Proposition 4.1.1 in these special cases. (The polynomials required to have roots in these cases will be examined more carefully in Proposition 4.5.1 below.) ■

As a simple example, an application of (i) (resp. (v)) recovers the case $\operatorname{char} K \neq 2$ (resp. $\operatorname{char} K = 2$) of Proposition 4.3.2(f).

Proposition 4.4.2 *When* $\operatorname{char} K \neq 2$ *or* 3*, the following algorithm decides whether* E *and* E' *are* K*-isomorphic (automatically taking care of cases where* $j = 0$ *or* 1728*); all exits from the procedure are via the* `return` *command.*

```
if j ≠ j' then return("no")
if Δ/Δ' ∉ K*^12 then return("no")
if √-1 ∈ K then return("yes")
choose any u ∈ K* such that Δ = Δ'u^12
if c_6 = c'_6 u^6 then return("yes")
return("no")
```

Proof. By Proposition 4.1.1 it is necessary that $j = j'$, $\Delta = \Delta' u^{12}$ and $c_6 = c'_6 u^6$ for some $u \in K^*$. This verifies the `"no"` returns. We now explain the `"yes"`'s.

The relations

$$j = c_4^3/\Delta = {c'}_4^3/\Delta', \quad \Delta = \Delta' u^{12}$$

imply $c_4^3 = {c'}_4^3 u^{12}$, hence $c_4 = {c'}_4 u^4 \rho$ where $\rho^3 = 1$, $\rho \in K$. Replacing u by $u\rho^2$ we can assume that $c_4 = c'_4 u^4$. Next, the relation $1728\Delta = c_4^3 - c_6^2$ and the analogous one for E' imply that $c_6^2 = c_6'^2 u^{12}$, hence $c_6 = \pm c'_6 u^6$. If the $+$ sign obtains or if we can replace u by $u\sqrt{-1}$ to change the sign to $+$, then we have the required τ. We have now exhausted the possibilities for u and the response `"yes"`. ■

We quote a set of 8 examples from [Com-Na87] which we will refer to again later to illustrate various phenomena. The curves are $E_1, \ldots, E_8$ and are defined over various quadratic fields $\mathbf{Q}(\sqrt{d})$. In each case ϵ stands for a fundamental unit: $\epsilon = 8 + 3\sqrt{7},\ 32 + 5\sqrt{41},\ 8 + \sqrt{65}$ for $d = 7, 41, 65$ respectively; and σ stands for the nontrivial automorphism of the field. Thus $\epsilon^\sigma = \epsilon^{-1}$, $-\epsilon^{-1}$, $-\epsilon^{-1}$ respectively.[‡] For all the E_i, $a_1 = 1$ and $a_3 = a_6 = 0$.

	d	a_2	a_4	Δ	j
E_1	7	-8ϵ	ϵ^3	ϵ^6	255^3
$E_1^\sigma = E_2$	7	$-8\epsilon^{-1}$	ϵ^{-3}	ϵ^{-6}	255^3
E_3	41	0	$-\epsilon$	ϵ^4	$(\epsilon - 16)^3/\epsilon$
$E_3^\sigma = E_4$	41	0	ϵ^{-1}	ϵ^{-4}	$\epsilon(\epsilon^{-1} + 16)^3$
E_5	65	8ϵ	ϵ^3	ϵ^6	257^3
E_6	65	$40\epsilon + 1$	$25\epsilon^3$	$(5\epsilon)^6$	257^3
E_7	65	2ϵ	ϵ^2	ϵ^6	17^3
E_8	65	$10\epsilon + 1$	$25\epsilon^2$	$(5\epsilon)^6$	17^3

[‡]We use exponential notation for Galois action; thus Galois modules are right modules.

In general for any automorphism σ of the field K and E with Weierstrass coefficients $a_1, \ldots$, we let E^σ denote the curve with Weierstrass coefficients $a_1^\sigma, \ldots$. Since Δ and j are rational functions in $a_1, \ldots, a_6$, E^σ has discriminant Δ^σ and j-invariant j^σ. Thus it is immediate that E_3 and E_4 are not isomorphic: their invariants are unequal.

Next let us check that E_1 and E_2 are not isomorphic over $K = \mathbf{Q}(\sqrt{7})$. The discriminant of E_2 is $\sigma(\epsilon^6) = \epsilon^{-6}$, so in the algorithm we can take $u = \pm\epsilon$. We find that $c_6'\epsilon^6/c_6 = -1$ (to decide whether this ratio is 1 or -1 of course requires only a rough calculation) and since $\sqrt{-1} \notin K$ we conclude that they are not isomorphic.

The algorithm applied to the curves E_5, E_5^σ proceeds quite similarly except this time the ratio is $+1$, and therefore they are isomorphic over $\mathbf{Q}(\sqrt{65})$. Similarly E_i is isomorphic to E_i^σ for $i = 6, 7, 8$.

4.5 Automorphisms and fields of definition

Let E, E' be two elliptic curves defined over the field K, and for any field F containing K, let $\mathrm{Isom}_F(E, E')$ denote the set of isomorphisms from E to E' defined over F. Clearly

$$\mathrm{Isom}_F(E', E) = \{\tau^{-1} : \tau \in \mathrm{Isom}_F(E, E')\}$$

and
$$\mathrm{Isom}_F(E, E) = \mathrm{aut}_F E.$$

Let $\overline{K}$ be an algebraic closure K, and K^{sep} the separable closure of K in $\overline{K}$.

If $\tau \in \mathrm{Isom}_{\overline{K}}(E, E')$ and $\sigma \in \mathrm{aut}_{\overline{K}} E$, then $\tau \circ \sigma$ is another isomorphism, and every isomorphism τ' from E to E' is of this form: take $\sigma = \tau^{-1} \circ \tau'$.

In the next proposition we see that $\mathrm{aut}_{\overline{K}} E$ is finite and that most of the time $K(\tau)$ is a quadratic extension of K and in any case is a separable extension.

Proposition 4.5.1 (a) *The group* $\mathrm{aut}_{\overline{K}} E$ *is finite. Its order* n *is* 2 *with the following exceptions:*

— *when* $\mathrm{char}\, K \neq 2$ *or* 3 : *if* $j = 0$ *then* $n = 6$, *while if* $j = 1728$ *then* $n = 4$;

— *if* $\mathrm{char}\, K = 3$ *and* $j = 0$ *then* $n = 12$;

— *if* $\mathrm{char}\, K = 2$ *and* $j = 0$ *then* $n = 24$.

(b) *If* τ *is an isomorphism from* E *to* E' *then* $K(\tau)$ *is a finite separable*† *extension of* K; *the degree* $[K(\tau) : K]$ *divides* n *with the following exceptions: when* $j = 0$ *and* $\mathrm{char}\, K$ *is, respectively, not* 2 *or* 3, 3, 2, *then the degree can be, respectively,* 4, 8, *one of* $9, 16, 18$.

(c) *The subgroup index* $i = [\mathrm{aut}_{\overline{K}} E : \mathrm{aut}_K E]$ *is* 1 *with the following exceptions:*

†We do not assume that K is perfect since we will want to consider examples such as $y^2 = x^3 + tx^2 - tx$ where $K = \mathbf{F}_p(t)$.

— *if* $j = 1728$, $\operatorname{char} K \neq 2$ *or* 3, *and* $\sqrt{-1} \notin K$ *then* $i = 2$;
— *if* $j = 0$, $\operatorname{char} K \neq 2$ *or* 3, *and* $\sqrt{-3} \notin K$ *then* $i = 3$;
— *if* $j = 0$, $\operatorname{char} K = 3$ *and* E *is in simplified form we present the values of* i *in a table:*

$r^3 + a_4 r - a_6$ *has a root in* K *and* $\sqrt{-1} \in K$	$\sqrt{-a_4} \in K$	i
yes	*yes*	1
no	*yes*	2
yes	*no*	3
no	*no*	6

— *if* $j = 0$, $\operatorname{char} K = 2$ *and* E *is in simplified form, then* i *is* 24 *divided by the number of solutions* (u, s, t) *in* K^3 *of the system*

$$\begin{aligned} u^3 &= 1, \\ s^4 + a_3 s + a_4(u+1) &= 0, \\ t^2 + a_3 t + a_3 s^3 + a_4 s^2 &= 0. \end{aligned}$$

Proof. By Corollary 4.3.1, E and E' have the same invariant j. In the proof we deal tandemly with the "isomorphism case" ($\tau\colon E \longrightarrow E'$) and the "automorphism case" ($\tau \in \operatorname{aut}_{\overline{K}} E$ and all $a'_i = a_i$.)

For each class we simplify the Weierstrass forms of E and E' as in Proposition 4.2.2 by applying appropriate isomorphisms defined over K — this does not change j or $K(\tau)$ and replaces $\operatorname{aut}_{\overline{K}} E$ by an isomorphic group (a conjugate subgroup in G as explained in Section 2).

Case 1: $\operatorname{char} K \neq 2$ or 3. Let E, E' be in simplified form. As in Proposition 4.4.1, cases (i)–(iii), $j = j'$ implies $c_4'^3 c_6^2 = c_4^3 c_6'^2$ which allows us to choose $u \in K^{\mathrm{sep}}$ such that

$$u^4 c'_4 = c_4, \qquad u^6 c'_6 = c_6.$$

Then $\tau = [0, 0, 0, u]$ is an isomorphism from E to E'.

Conversely any isomorphism is in stab $= U$ and we have $u^4 a'_4 = a_4$, $u^6 a'_6 = a_6$. If $j \neq 0$ or 1728 we deduce $u^2 \in K$ so $[K_1 : K] \leq 2$ and for automorphisms $u^2 = 1$. Thus $\operatorname{aut}_{\overline{K}} = \{[1], [-1]\}$ and all is clear in these cases.

Next for the cases $j = 0$, resp. $j = 1728$ (still with $\operatorname{char} \neq 2, 3$) we have $u^6 \in K$, resp. $u^4 \in K$ for an isomorphism, and the possibilities $u^6 = 1$, resp. $u^4 = 1$ for automorphisms. Now $u^6 \in K \Longrightarrow [K(u)\colon K] = 1$, 2, 3, 4 or 6 † and $u^4 \in K \Longrightarrow [K(u)\colon K] = 1$, 2 or 4. Again all is clear. Incidentally for the cases so far

$$\operatorname{aut}_{\overline{K}} E = \{[0, 0, 0, u] : u^n = 1]\}$$

†Here is an example of degree 4: $K = \mathbf{Q}(t)$, t transcendental, $E\colon y^2 = x^3 + 1$, $E'\colon y^2 = x^3 + t^3$ so $\operatorname{Isom}_{\overline{K}}(E, E') = \{[0, 0, 0, u]\colon u \in \{\pm t^{-1/2}, \pm t^{-1/2}(1 \pm \sqrt{-3})/2\}\}$. If $\tau = [0, 0, 0, t^{-1/2}(1 + \sqrt{-3})/2]$ then $[K(\tau)\colon K] = 4$.

is cyclic.

Case 2: $\operatorname{char} K = 3$. Consider an isomorphism $[r, 0, 0, u] \in \operatorname{stab}$ mapping E to E', both in simplified form. If $j \neq 0$ then $a_2 \neq 0$, $u^2 a_2' = a_2$ and we are in the situation with $n = 2$ as before. If $j = 0$ then $a_2 = 0$, $a_4 \neq 0$ (in order that $\Delta \neq 0$), and

$$u^4 a_4' = a_4, \qquad u^6 a_6' = a_6 + r a_4 + r^3.$$

For an isomorphism we have $u^4 \in K$, hence $K(u)/K$ is a separable extension of degree dividing 4 and, since the derivative of the last equation for r is $3r + a_4 = a_4 \neq 0$, $K(u, r)/K(u)$ is separable of degree 1, 2 or 3.[‡] For automorphisms, $u^4 = 1$ gives 4 possibilities for u, and for each u the other equation gives 3 possibilities for r which are distinct since the derivative is nonzero. Thus aut has order 12 in this case; in fact it is the unique noncommutative semidirect product of C_4 by C_3, as one can easily verify. We must check that $[K_{\mathrm{aut}} : K] \mid 12$. When $u = \pm 1$ the equation for r is $r(r^2 + a_4) = 0$, thus K_{aut} contains $\sqrt{-a_4}$. When $u = \pm\sqrt{-1}$ the equation for r is $r^3 + a_4 r - a_6 = 0$. The result follows from the fact that the discriminant of this cubic, in characteristic 3, is $-a_4^3$ which is a square once $\sqrt{-a_4}$ is adjoined. (If $r = \theta$ is one root the others are $\theta \pm \sqrt{-a_4}$.)

Conversely suppose $j = j'$. If $j \neq 0$, we choose $r = 0$ and $u \in K^{\mathrm{sep}}$ to satisfy $u^2 a_2' = a_2$, while if $j = 0$ we choose $u, r \in K^{\mathrm{sep}}$ to satisfy $u^4 a_4' = a_4$, $u^6 a_6' = a_6 + r a_4 + r^3$. In both cases $[r, 0, 0, u]$ is an isomorphism in stab.

Case 3: $\operatorname{char} K = 2$, $j \neq 0$. If $j = j'$, *i.e.*, $a_6 = a_6'$ then there are two isomorphisms $[0, s, 0, 1]$ given by the two roots of the separable equation $s^2 + s + a_2 = a_2'$, and once again we have the simple $n = 2$ situation.

Case 4: $\operatorname{char} K = 2$, $j = 0$. The a_3-equation in the simplified form is $u^3 a_3' = a_3$, hence 3 values for u and $[K(u) \colon K] = 1$, 2 or 3. The a_2-equation gives $r = s^2$ which when substituted into the a_4-equation gives the separable quartic

$$s^4 + a_3 s + a_4 = u^4 a_4',$$

hence 4 distinct roots for each value of u, and $[K(u, s) \colon K(u)] = 1$, 2, 3 or 4. The a_6-equation is the separable quadratic for t

$$t^2 + a_3 t + s^6 + s^2 a_4 + a_6 = u^6 a_6',$$

thus 2 distinct roots for each pair s, u, and $[K(u, s, t) \colon K] = 1$ or 2. (It is not hard to construct examples of $[K(\tau) \colon K] = 9$, 16 and 18 .) From this we see that $n = 24$.

Now let us show that $[K_{\mathrm{aut}} : K] | 24$. Taking $a_i = a_i'$, the equations become

[‡]Here is an example of $[K(\tau) \colon K] = 8$: $K = \mathbf{F}_3(s, t)$ where s and t are independent transcendentals, $E \colon y^2 = x^3 - sx$, $E' \colon y^2 = x^3 - stx$ so $\operatorname{Isom}_{\overline{K}}(E, E') = \{[r, 0, 0, u] \colon r \in \{0, \pm\sqrt{s}\}, u \in \{\pm t^{-1/4}, \pm\sqrt{-1} t^{-1/4}\}\}$, whereas $\operatorname{aut}_{\overline{K}} E = \{[r, 0, 0, u] \colon r \in \{0, \pm\sqrt{s}\}, u \in \{\pm 1, \pm\sqrt{-1}\}\}$. — 12 in all. If $\tau = [\sqrt{s}, 0, 0, \sqrt{-1} t^{-1/4}]$ then $[K(u, r) \colon K] = 2$ and $[K(\tau) \colon K] = 8$.

$$\begin{aligned} u^3 &= 1, \\ s^4 + a_3 s + a_4(u+1) &= 0, \\ t^2 + a_3 t + a_3 s^3 + a_4 s^2 &= 0. \end{aligned}$$

Let $1, \zeta, \zeta^2 = 1 + \zeta$ denote the values of u. When $u = 1$ the values of s are $0,\ \alpha = \sqrt[3]{a_3},\ \zeta\alpha,\ \zeta^2\alpha$. If $s = \theta$ is a root when $u = \zeta^2$ then the other roots† are $\theta+\alpha,\ \theta+\zeta\alpha,\ \theta+\zeta^2\alpha$ and the roots when $u = \zeta$ are $\zeta\theta,\ \zeta\theta+\alpha,\ \zeta\theta+\zeta\alpha,\ \zeta\theta+\zeta^2\alpha$. If t_1 denotes one value of t then the other value is $t_1 + a_3$. When $s = 0$, then $t_1 = 0$; when $s = \alpha$ then

$$t_1 = \zeta a_3 + \zeta\theta^2\alpha + \zeta^2\theta\alpha^2$$

with similar expressions when $s = \zeta\alpha,\ \zeta^2\alpha$; when $u = \zeta^2$ (resp. ζ) then $t_1 = \zeta s^3$ (resp. $\zeta^2 s^3$).Thus $K(\zeta, \alpha, \theta)$ is a normal extension of K of degree dividing 24.

Part (c) now follow easily; we leave the details to the reader. In the case $j = 0$, $\operatorname{char} K = 2$ we did not state the results more explicitly in tabular form as we did when $\operatorname{char} K = 3$ since the situation is rather complicated with many subcases. ■

The most exotic aut that occurs in the above proposition is a group of order 24. As will be seen in a later chapter, it is the semidirect product of the cyclic group of order 3 by the quaternion group of order 8; a concrete realization is $\mathbf{SL}_2(\mathbf{F}_3)$. Once we have defined and identified $\operatorname{end}_{\overline{K}} E$ as the maximal order of Hurwitz quaternions in the skew field of rational quaternions, then a more natural definition of aut is the group of units $\pm 1,\ \pm i,\ \pm j,\ \pm k,\ (\pm 1 \pm i \pm j \pm k)/2$.

The group of order 12 that occurs in characteristic 3 can be described as the semidirect product of the cyclic group of order 4 by the cyclic group of order 3.

4.6 Legendre and Deuring forms

A **Legendre form** is an elliptic curve with an equation

$$y^2 = x(x-1)(x-\lambda);$$

for this equation

$$c_4 = 16(\lambda^2 - \lambda + 1), \qquad \Delta = 16\lambda^2(\lambda-1)^2, \qquad j = \frac{2^8(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2}.$$

In order that $\Delta \neq 0$, it is necessary that $\operatorname{char} K \neq 2$ and that $\lambda \neq 0$ or 1. The points $(0,0),\ (1,0),\ (\lambda,0)$ have order 2.

†I am indebted to Y. Zhang for supplying some of the details.

Proposition 4.6.1 *Let K be a field of characteristic not 2 and let E be an elliptic curve defined over K. Then*

(a) *E is K-isomorphic to a Legendre form defined over K iff the following two conditions are met:*

(i) $|E(K)[2]| = 4$, i.e., *all points of order 2 are defined over K, say*

$$y^2 = (x-a)(x-b)(x-c), \qquad (*)$$

(we may take E in b-form since $\operatorname{char} K \neq 2$)

(ii) *and at least one of $\pm(a-b)$, $\pm(b-c)$, $\pm(c-a)$ is a square in K^*.*

Suppose this is the case, say E is K-isomorphic to the Legendre form with parameter λ. Then the Legendre forms with parameter λ' that are K-isomorphic with E are as follows.

- *if $\lambda \in K^{*2}$ then $\lambda' = 1/\lambda$;*
- *if $-\lambda \in K^{*2}$ then $\lambda' = (\lambda-1)/\lambda$;*
- *if $\lambda - 1 \in K^{*2}$ then $\lambda' = 1/(1-\lambda)$;*
- *if $1-\lambda \in K^{*2}$ then $\lambda' = \lambda/(\lambda-1)$;*
- *if $-1 \in K^{*2}$ then $\lambda' = 1-\lambda$.*

(b) *In general, E acquires a twist in Legendre form over a separable extension $K(\lambda)$ of degree at most 6, and E is isomorphic to a Legendre form over a separable extension of degree at most 12.*

Proof. (a) A Legendre form obviously satisfies (i) and (ii). Conversely, suppose $b-a = u^2$. Then $[a,0,0,u]$ transforms $(*)$ to Legendre form with $\lambda = (c-a)/(b-a)$.

The Weierstrass coefficients of the Legendre form are $a_1 = a_3 = a_6 = 0$, $a_2 = -(\lambda+1)$, $a_4 = \lambda$. Suppose $[r,s,t,u]$ transforms this Legendre form to another with parameter λ'. The transformation equation $ua_1' = a_1 + 2s$ shows that $s = 0$, similarly that for a_3 shows $t = 0$, and now the transformation equation for a_6 is $0 = r\lambda - r^2(\lambda+1) + r^3$, whence $r = 0$, 1 or λ. The equations for a_2 and a_4 are

$$\begin{aligned} -u^2(\lambda'+1) &= -(\lambda+1) + 3r, \\ u^4\lambda' &= \lambda - 2r(\lambda+1) + 3r^2. \end{aligned}$$

One now deals with the three values of r separately, using simple calculations. For example when $r = 0$ one deduces $\lambda(1-u^{-2}) = u^2 - 1$, hence either $u = \pm 1$ (no change in the Legendre form), or $\lambda = u^2$ and then $\lambda' = 1/\lambda$. The two other cases are only slightly more involved.

(b) For the first statement it is only necessary to choose λ as any root of

$$256(\lambda^2 - \lambda + 1)^3 - \lambda^2(\lambda - 1)^2 j(E) = 0.$$

We check the separability when $\operatorname{char} K = 3$. If $j = 0$ then $\lambda = -1$ so $K(\lambda) = K$ is a separable extension. Otherwise it follows by the derivative test: the above polynomial for λ and its derivative have no common root.

For the last statement we can take E in the form $y^2 = x^3 + ax^2 + bx + c$. The root field of the cubic is a separable extension since the polynomial discriminant $\frac{1}{16}\Delta \neq 0$. Then adjoining a square root if necessary as explained in (a) allows one to write the equation in Legendre form. ■

When $K = \mathbf{R}$ and $\Delta > 0$, one of $\pm(a - b)$ in condition (a)(ii) is positive, and so we can state the classical case $E_{/\mathbf{R}}$ as follows.

Corollary 4.6.2 *Let E be defined over the real field $\mathbf{R}$. Then E is $\mathbf{R}$-isomorphic with a Legendre form $y^2 = x(x-1)(x-\lambda)$, $\lambda \in \mathbf{R}$, iff $\Delta > 0$; and then the other possible choices for the parameter are as follows.*

$$\begin{aligned} \lambda < 0 \quad &: \quad \lambda' = (\lambda - 1)/\lambda \text{ or } \lambda/(\lambda - 1); \\ 0 < \lambda < 1 \quad &: \quad \lambda' = \lambda/(\lambda - 1)/ \text{ or } 1/\lambda; \\ 1 < \lambda \quad &: \quad \lambda' = 1/\lambda \text{ or } 1/(1 - \lambda). \quad \blacksquare \end{aligned}$$

The **Deuring form** is

$$y^2 + \delta xy + y = x^3;$$

for this Weierstrass equation

$$c_4 = \delta(\delta^3 - 24), \qquad \Delta = \delta^3 - 27, \qquad j = \delta^3(\delta^3 - 24)^3/(\delta^3 - 27).$$

The point $(0, 0)$ has order 3.

Proposition 4.6.3 *Let the elliptic curve E be defined over the field K. Then, except when $\operatorname{char} K = 3$ and $j = 0$, there exists an extension $K(\delta)$ of degree at most 12 over which E acquires a twist in Deuring form. When $\operatorname{char} K \neq 3$, the extension $K(\delta)/K$ is separable for every choice of δ.*

Proof. Choose δ to be any root of

$$\delta^3(\delta^3 - 24)^3 - (\delta^3 - 27)j(E) = 0,$$

except that when $\operatorname{char} K = 3$, when the equation is $\delta^{12} - \delta^3 j = 0$, we require $\delta \neq 0$. The corresponding Deuring normal form is defined over $K(\delta)$ and has the same j as E. The formula for j makes it clear that we have considered all possibilities for δ.

When $\operatorname{char} K \neq 3$, let $\epsilon = \delta^3$. Then it is easy to prove (we omit the details) that $K(\epsilon)/K$ and $K(\delta)/K(\epsilon)$ are separable. ■

4.7 Finite fields

We conclude this chapter with examples of $E * d$ defined over finite fields $\mathbf{F}_q$ where $q = p^n$ is a prime power.

For a positive integer m and field K, we let $\boldsymbol{\mu}_m(K)$ denote the group of m-th roots of 1 in K. We regard K as a subfield of an algebraic closure $\overline{K}$, so that

$$\boldsymbol{\mu}_m(\overline{K}) \cap K = \boldsymbol{\mu}_m(K).$$

The following elementary lemma will be used on several occasions.

Lemma 4.7.1 *Let p be a prime, let n be a positive integer, let $q = p^n$, and let m be a positive divisor of $q - 1$.*

An element $c \in \mathbf{F}_q{}^*$ is an m-th power in $\mathbf{F}_q{}^*$ iff $c^{(q-1)/m} = 1$, and then c has m distinct m-th roots in $\mathbf{F}_q$. In particular, taking $c = 1$, $\mathbf{F}_q$ contains all m m-th roots of $1 : \boldsymbol{\mu}_m(\mathbf{F}_q) = \boldsymbol{\mu}_m(\overline{\mathbf{F}_q})$. In any case, $c^{(q-1)/m} \in \boldsymbol{\mu}_m(\mathbf{F}_q)$.

Thus the following are equivalent:

- *m is a divisor of $p - 1$;*
- *$c^{(q-1)/m} \in \mathbf{F}_p$ for all $c \in \mathbf{F}_q$;*
- *$\mathbf{F}_p$ contains all the m-th roots of 1.*

Proof. This is all immediate from the fact that $\mathbf{F}_q{}^*$ is cyclic: if g is a generator then the m-th powers in $\mathbf{F}_q{}^*$ are the $(q-1)/m$ elements of the form g^{jm}, $j \bmod (q-1)/m$, and these are the roots of the polynomial $x^{(q-1)/m} - 1$, *i.e.*, the $(q-1)/m$-th roots of 1. Replacing m in this by its complementary divisor $(q-1)/m$, the m-th roots of 1 are $g^{j(q-1)/m}$, $j \bmod m$. In general, for any M, the smallest overfield of $\mathbf{F}_q$ containing $\boldsymbol{\mu}_M(\overline{\mathbf{F}_q})$ is $\mathbf{F}_{q^k}$ where k is minimal such that $M|(q^k - 1)$.

$g^e \in \mathbf{F}_p$ iff e is a multiple of $(q-1)/(p-1)$. In particular, $g^{j(q-1)/m} \in \mathbf{F}_p{}^*$ $\forall j$ when m divides $p - 1$. ■

4.7.1 The trace of Frobenius: preliminaries

Let $E(a)$ denote the Weierstrass equation for E in which the value a has been substituted for the variable x, and define the symbol $(E(a)/q)$ be 1 less than the number of solutions $y \in \mathbf{F}_q$ of the equation $E(a)$. For example when E is $y^2 + xy = x^3 + 1$ and $K = \mathbf{F}_2$ then $E(0)$ is the equation $y^2 = 1$ which has the single root $y = 1$, hence $(E(0)/2) = 0$; $E(1)$ is the equation $y^2 + y = 0$ has two roots hence $(E(1)/2) = 1$. Similarly when $K = \mathbf{F}_3$ we find $(E(x)/3) = 1, 0, 1$ when $x = 0, 1, 2$ respectively.

From the definition of $(E(x)/q)$ it follows that the number of points on E defined over $\mathbf{F}_q$ (remember to add 1 for the point at ∞) is

$$|E(\mathbf{F}_q)| = 1 + \sum_{x \in \mathbf{F}_q} \left(1 + \left(\frac{E(x)}{q}\right)\right) = q + 1 + \sum_{x \in \mathbf{F}_q} \left(\frac{E(x)}{q}\right).$$

We write this as $q+1-t$ and call t the **trace of Frobenius**:

$$t = t(E,q) = -\sum_{x\in\mathbf{F}_q}\left(\frac{E(x)}{q}\right).$$

The following is immediate from Proposition 4.1.2.

Lemma 4.7.2 *The trace of Frobenius is an $\mathbf{F}_q$-isomorphism invariant: if E and $E' = \tau E$ are elliptic curves over $\mathbf{F}_q$ where τ is an isomorphism defined over $\mathbf{F}_q$, then*

$$t(E',q) = t(E,q) \quad \blacksquare$$

When p is odd and E is in Weierstrass b-form $y^2 = f(x) = x^3 + \cdots$ then $(E(a)/q)$ is the (generalized) **Legendre symbol** (d/q) where $d = f(a)$, whose values are given by

$$\left(\frac{d}{q}\right) = \begin{cases} 1 & \text{when } d \text{ is a quadratic residue, } i.e.\ d \in {\mathbf{F}_q}^{*2}, \\ 0 & \text{when } d = 0, \\ -1 & \text{when } d \text{ is a quadratic nonresidue, } i.e. \text{ neither of the above.} \end{cases}$$

When q is odd, any E is isomorphic to its b-form, and so by the lemma we can assume that E is the b-form in order to calculate $t(E,q)$. Then, multiplying the right side of the b-form by 4, which does not affect the value of the Legendre symbol, we have

Proposition 4.7.3 *For q odd*

$$t(E,q) = -\sum_{x\in\mathbf{F}_q}\left(\frac{4x^3+b_2x^2+2b_4x+b_6}{q}\right) \quad \blacksquare$$

Recall Euler's relation: when q is odd and $d \in \mathbf{F}_q$ then $d^{(q-1)/2}$ is in the prime subfield $\mathbf{F}_p$ and

$$\left(\frac{d}{q}\right) \equiv d^{(q-1)/2} \bmod p.$$

This is the case $m = 2$ of Lemma 4.7.1.

In order to deal with cases $q = 2^n$ of characteristic 2 we define the symbol

$$\left(\frac{d}{2^n}\right) = \begin{cases} 1 & \text{when } s^2+s = d \text{ has a solution } s \in \mathbf{F}_{2^n}, \\ -1 & \text{when it does not.} \end{cases}$$

For example $(0/2) = 1$ and $(1/2) = -1$.

Since the terminology is available for use in the characteristic 2 case, we define d to be a **quadratic residue** in $\mathbf{F}_{2^n}$ when $(d/2^n) = 1$, and a **quadratic**

nonresidue when $(d/2^n) = -1$. Here are some simple facts whose proofs we leave to the reader:

(i) $\left(\dfrac{d_1 + d_2}{2^n}\right) = \left(\dfrac{d_1}{2^n}\right)\left(\dfrac{d_2}{2^n}\right)$;

(ii) exactly half the elements of $\mathbf{F}_{2^n}$ are quadratic residues — they form an additive subgroup of $\mathbf{F}_{2^n}$ of index 2.

(iii) $\left(\dfrac{d^2}{2^n}\right) = \left(\dfrac{d}{2^n}\right)$, hence for $k \geq 0$, $\left(\dfrac{d^{2^k} + d}{2^n}\right) = 1$.

Again by the lemma, for purposes of calculating $t(E, 2^n)$, we can assume that E is in one of the simplified forms (d) or (e) of Proposition 4.2.2.

Proposition 4.7.4 *Let E be an elliptic curve defined over $\mathbf{F}_{2^n}$.*

(i) *When $j \neq 0$ we can write E in the simplified form*

$$y^2 + xy = x^3 + a_2x^2 + a_6, \quad \Delta = a_6, \quad j = 1/a_6,$$

and then

$$t(E, 2^n) = -\sum_{x \in \mathbf{F}_{2^n}^*} \left(\frac{(x^3 + a_2x^2 + a_6)/x^2}{2^n}\right).$$

(ii) *When $j = 0$ we can write E in the simplified form*

$$y^2 + a_3y = x^3 + a_4x + a_6, \quad \Delta = a_3^4,$$

and then

$$t(E, 2^n) = -\sum_{x \in \mathbf{F}_{2^n}} \left(\frac{(x^3 + a_4x + a_6)/a_3^2}{2^n}\right).$$

Proof. (i) Since $\mathbf{F}_{2^n}$ is perfect, $(E(0)/2^n) = 0$. When $x \neq 0$ and we substitute $y = xs$, the equation can be written as $s^2 + s = (x^3 + a_2x^2 + a_6)/x^2$, and the formula for t follows.

(ii) In this case we can substitute $y = a_3s$, and the equation can be written $s^2 + s = (x^3 + a_4x + a_6)/a_3^2$. ■

Proposition 4.7.5 *Let E be an elliptic curve defined over the finite field $\mathbf{F}_q$ and let $d \in \mathbf{F}_q$. When q is odd assume that $d \neq 0$. Then*

$$t(E^d, q) = \left(\frac{d}{q}\right) t(E, q).$$

Hence if d is a quadratic nonresidue, $E \mapsto E^d$ sets up a bijection between the set of E which have $q + 1 - t$ points and the set of those which have $q + 1 + t$.

Proof. First let q be odd. Then $t(E^d, q)$ is given by the formula in Proposition 4.7.3 with b_2, b_4, b_6 replaced by db_2, d^2b_4, d^3b_6 respectively. Replacing x by dx in the sum we have

$$t(E^d, q) = -\sum_{x \in \mathbf{F}_q} \left(\frac{d^3}{q}\right) \left(\frac{4x^3 + b_2x^2 + 2b_4x + b_6}{q}\right) = \left(\frac{d}{q}\right) t(E, q).$$

Second let $q = 2^n$. If $j \neq 0$ then $t(E^d, 2^n)$ is given by the first formula of the previous proposition with a_2 replaced by $a_2 + d$:

$$t(E^d, 2^n) = -\sum_{x \in \mathbf{F}_q{}^*} \left(\frac{(x^3 + a_2x^2 + a_6)/x^2 \ + \ d}{2^n}\right) = \left(\frac{d}{q}\right) t(E, 2^n),$$

using the basic rule $((d' + d)/2^n) = (d'/2^n)(d/2^n)$. The proof when $j = 0$ is similar: in the second formula of the previous proposition a_6 is replaced by $a_6 + a_3^2 d$. ■

4.7.2 An application of Burnside's formula

Proposition 4.7.6 *Let $\nu = \nu(q, j)$ denote the number of $\mathbf{F}_q$-isomorphism classes of elliptic curves with given j. Then $\nu = 2$ except for certain cases when $j = 0$ or 1728 as detailed in the following table.*

q	j	ν
1 or 5 mod 12	1728	4
1 mod 6	0	6
$3^n, n$ odd	0	4
$3^n, n$ even	0	6
$2^n, n$ odd	0	3
$2^n, n$ even	0	7

Remark. In passing from one field $\mathbf{F}_q$ to an overfield $\mathbf{F}_{q^k}$, in general some non-isomorphic E over $\mathbf{F}_q$ become isomorphic over the larger field; but the proposition assures that except in some cases where $j = 0$ or 1728, new isomorphism types appear to compensate for this loss exactly.

Proof. The value $\nu = 2$ when $j \neq 0, 1728$ is immediate from earlier results: since the quadratic residues form a subgroup of index 2 in $\mathbf{F}_q{}^*$ (resp. in $\mathbf{F}_q{}^+$ when q is even), Proposition 4.3.2(f) implies that for each $j \in \mathbf{F}_q \backslash \{0, 1728\}$, there are precisely two $\mathbf{F}_q$-isomorphism classes of E with invariant j; they are represented by the generic-j curve and a twist of it by a quadratic nonresidue.

However it is interesting to see a quite different method of proof, a method which will uniformly determine ν in all cases. The method is Burnside's theorem:

Let the finite group Γ act on the finite set X. Then the number of orbits is the average size of $\mathrm{Fix}(g) = \{x \in X : gx = x\}$:

$$\#orbits = \frac{1}{|\Gamma|} \sum_{g \in \Gamma} |\mathrm{Fix}(g)|.$$

We use Burnside's theorem to determine ν in the simplest case ($p > 3$, $j \neq 0, 1728$) and the hardest case (the bottom line of the table), leaving the intermediate cases to the reader.

Case $p > 3$, $j \neq 0, 1728$: As explained in Proposition 4.2.2, the isomorphism classes that we wish to enumerate are the orbits of the group U acting on the set X of Weierstrass equations

$$E\colon\ y^2 = x^3 - (c_4/48)x - (c_6/864)$$

whose invariant is the given j. For such E we have $c_4 c_6 \neq 0$ and

$$\Delta = \frac{c_4^3}{j} = \frac{c_6^2}{j - 1728}.$$

Thus $|X|$ is the number of pairs $c_4, c_6 \in \mathbf{F}_q{}^*$ satisfying the equation $(j - 1728)c_4^3/j = c_6^2$. There are $(q-1)/2$ choices of c_4 to make $c_4^3(j - 1728)/j$ a quadratic residue, and then two choices for c_6. Hence $|X| = q - 1$.

Consider any $\tau = [0, 0, 0, u] \in U$ with $\mathrm{Fix}(\tau) \neq \emptyset$, say $E \in \mathrm{Fix}(\tau)$. The transformation equations $u^4 a_4 = a_4$, $u^6 a_6 = a_6$ imply that $u^2 = 1$, *i.e.*, $u = \pm 1$. For both values we have $\mathrm{Fix}(\tau) = X$. Since $|U| = q - 1$, Burnside's formula gives

$$\nu = \frac{1}{q-1}\,[(q-1) + (q-1)] = 2,$$

all other $\tau \in U$ contributing $|\mathrm{Fix}(\tau)|$=0 to the sum.

Case $q = 2^n$, $j = 0$: $X = \{E\colon y^2 + a_3 y = x^3 + a_4 x + a_6,\ a_3 \neq 0\}$, and $\Gamma = T = \{\tau = [s^2, s, t, u]\}$, hence

$$|X| = |T| = (q-1)q^2 = (2^n - 1)2^{2n}.$$

We must calculate $|\mathrm{Fix}(\tau)|$ for each $\tau \in T$.

Case $u = 1, s = 0$:

— If $t = 0$, we have $|\mathrm{Fix}(1)| = (2^n - 1)2^{2n}$.

— If $t \neq 0$, the transformation equations, with $a_i' = a_i \forall i$, reduce to $ta_3 + t^2 = 0$, hence a_3 is determined by τ while a_4, a_6 can be chosen arbitrarily. Thus $|\mathrm{Fix}(\tau)| = 2^{2n}$. The number of such τ is $2^n - 1$ and therefore they contribute a total of $(2^n - 1)2^{2n}$ to the sum in Burnside's formula.

Case $u = 1, s \neq 0$: the equations are $0 = sa_3 + s^4$ and $0 = s^2a_4 + s^6 + ta_3 + t^2$. With any t, these equations determine a_3 and a_4, while a_6 is arbitrary. Thus $|\mathrm{Fix}(\tau)| = 2^n$, and since the number of these τ is $(2^n - 1)2^n$, the contribution to the sum is $(2^n - 1)2^{2n}$.

Case $u = \zeta$, a primitive cube root of 1 in $\mathbf{F}_{2^n}$, which exists since n is even. After this case is completed, the contributions to the sum will be multiplied by 2 to cover the case $u = \zeta^2$.

— Subcase $s = 0$: The equations are $\zeta a_4 = a_4$, hence $a_4 = 0$, and $0 = ta_3 + t^2$. If $t = 0$ any a_3 ($\neq 0$ so that $\Delta \neq 0$) is allowed, while if $t \neq 0$ then a_3 is determined by τ. For any t, a_6 can be chosen arbitrarily, and the contribution of these two "subsubcases" to the sum is

$$1 \cdot (2^n - 1)2^n + (2^n - 1) \cdot 2^n = (2^n - 1)2^{n+1}.$$

— Subcase $s \neq 0$: the equations are

$$\begin{aligned} 0 &= (\zeta + 1)a_4 + sa_3 + s^4, \\ 0 &= s^2a_4 + s^6 + ta_3 + t^2. \end{aligned}$$

Solving the first equation for a_4, using $1/(\zeta+1) = \zeta$, and substituting in the second yields

$$0 = a_3(\zeta s^3 + t) + (\zeta + 1)s^6 + t^2.$$

— If $t = \zeta s^3$, the last equation is satisfied. since there are $2^n - 1$ choices for s, a_3 is arbitrary nonzero, a_4 is determined and a_6 is arbitrary, the contribution is $(2^n - 1) \cdot (2^n - 1)2^n$.

— If $t \neq \zeta s^3$, then a_3 is determined and the contribution is $(2^n - 1)^2 \cdot 2^n$.

Remembering to multiply the $u = \zeta$ contributions by 2 to account for $u = \zeta^2$, adding up all these contributions and dividing by $|T| = (2^n - 1)2^{2n}$ gives $\nu = 7$. ■

We can now easily obtain the total number of isomorphism classes of E over $\mathbf{F}_q$ by adding up appropriate ν's in the table in the proposition. For example when $p > 3$ (resp. $p = 2$ or 3), the first line in the table contributes $\nu = 2$ for each of the $q - 2$ (resp. $q - 1$) values of $j \neq 0$ or 1728.

Corollary 4.7.7 *The number N of isomorphism classes of elliptic curves E over $\mathbf{F}_q$ where $q = p^n$ is as follows:*

q	N
2^n, *n odd*	$2^{n+1}+1$
2^n, *n even*	$2^{n+1}+5$
3^n, *n odd*	$2(3^n+1)$
3^n, *n even*	$2(3^n+3)$
1 mod 12	$2q+6$
5 mod 12	$2q+2$
7 mod 12	$2q+4$
11 mod 12	$2q$ ■

When enumerating E of one type or another over finite fields, one often obtains more elegant results by assigning the "weighted" cardinality $1/|\mathrm{aut}_{\mathbf{F}_q} E|$ to E, rather than the natural cardinality 1. A formula for the sum of the relevant weighted cardinalities is referred to as a **mass formula**. The following mass formulas exhibit a striking simplification of the results of the previous corollary.

Corollary 4.7.8 *Let S denote the set of $\mathbf{F}_q$-isomorphism classes $[E]$ of elliptic curves $E_{/\mathbf{F}_q}$, and let S_j denote the subset of S represented by E with the given invariant j. Then*

$$\sum_{[E]\in S_j} \frac{1}{|aut_{\mathbf{F}_q} E|} = 1, \quad \textit{hence} \quad \sum_{[E]\in S} \frac{1}{|aut_{\mathbf{F}_q} E|} = q.$$

Remark. For an exquisite proof that does not plod through case by case, see ([How93]). It would take us too far afield at present to provide the background for that proof, and so it will be necessary for us to "get our hands dirty" – again.
Proof. We obtain a proof by combining results of the present proposition (made more explicit when $p = 2, 3$) with parts (a) and (c) of Proposition 4.5.1. For convenience we write a for $|\mathrm{aut}_{\mathbf{F}_q} E|$ and $\overline{a}$ for $|\mathrm{aut}_{\overline{\mathbf{F}}_q} E|$, so that $a = \overline{a}/i$ in the notation of Proposition 4.5.1. Also $|S_j|$ is denoted ν in the present proposition.

For each $j \neq 0$ or 1728 we have $\nu = 2$, and for each of these two $[E]$ we have $a = \overline{a} = 2$, which proves the mass formula for these j.

Next let $j = 1728$ and $p > 3$. For $q \equiv 1 \bmod 4$, resp. $q \equiv 3 \bmod 4$, we have $\nu = 4$, resp. 2, and $\sqrt{-1} \in K$, resp. $\notin K$, hence $a = 4/1 = 4$, resp. $a = 4/2 = 2$ — again the formula is verified.

Now let $j = 0$, $p > 3$. For $q \equiv 1$ resp. $5 \bmod 6$ we have $\nu = 6$, resp. 2, and $\sqrt{-3} \in K$ resp. $\notin K$, so $a = 6/1 = 6$ resp. $6/3 = 2$, and the formula is established. This completes the proof of the mass formulas for $p > 3$.

There remain the four cases $q = 3^n$, 2^n, n odd or even, all with $j = 0$. In each case we list representatives of the ν members of S_0. And in each case, rather than writing out all the details, we supply notes that make it easy to

(i) authenticate the list, that is, check that no two E in the list are $\mathbf{F}_q$-isomorphic by the tests given in Proposition 4.4.1, and

(ii) with the help of Proposition 4.5.1, determine the values of a for all the E in the list, hence prove the mass formula.

Case 3^n, *n odd.*

$$\begin{array}{ll} y^2 = x^3 + x\,(a=2) & y^2 = x^3 - x\,(a=6), \\ y^2 = x^3 - x + c\,(a=6), & y^2 = x^3 - x - c\,(a=6). \end{array}$$

Notes. Since $q - 1 \equiv 2 \bmod 4$, therefore ${\mathbf{F}_q}^{*4} = {\mathbf{F}_q}^{*2}$. Since $(-1/q) = -1$, it follows that $\{r^3 + r : r \in \mathbf{F}_q\}$ coincides with $\mathbf{F}_q$, whereas $\{r^3 - r : r \in \mathbf{F}_q\}$ is an additive subgroup of index 3 — we write coset representatives as $0, \pm c$.

Case 3^n, *n even.*

$$\begin{array}{ll} y^2 = x^3 + x\,(a=12) & y^2 = x^3 + x + d\,(a=6), \\ y^2 = x^3 + \rho x\,(a=4), & y^2 = x^3 + \rho^3 x\,(a=4), \\ y^2 = x^3 + \rho^2 x\,(a=12), & y^2 = x^3 + \rho^2 x + e\,(a=6). \end{array}$$

Notes. ρ denotes a primitive root: ${\mathbf{F}_q}^* = \langle \rho \rangle$; coset representatives of $\{r^3 + r\}$, resp. $\{r^3 + \rho^2 r\}$ are $0, \pm d$, resp. $0, \pm e$; for $k = 1$ and 3, the group $\{r^3 + \rho^k r\}$ is all of $\mathbf{F}_q$.

Notes for the two cases $q = 2^n$. We let Q_2 (resp. Q_4) denote the additive subgroup $\{x^2 + x : x \in \mathbf{F}_q\}$ (resp. $\{x^4 + x : x \in \mathbf{F}_q\}$). Of course Q_2 is the group of quadratic residues. Since

$$\left(\frac{x^4 + x}{2^n}\right) = \left(\frac{x^4}{2^n}\right)\left(\frac{x}{2^n}\right) = \left(\frac{x}{2^n}\right)\left(\frac{x}{2^n}\right) = 1,$$

Q_4 is a subgroup of Q_2. The group index is 1 (resp. 2) when n is odd (resp. even); this follows from the fact that $\mathbf{F}_q$ contains a primitive cube root of unity — an element ζ satisfying $\zeta^2 = \zeta + 1$ — iff n is even.

Case 2^n, *n odd.*

$$y^2 + y = \begin{cases} x^3, & a = 2, \\ x^3 + x, & a = 4, \\ x^3 + x + 1, & a = 4. \end{cases}$$

Notes. Since $q - 1 \equiv 1 \bmod 3$, every element of $\mathbf{F}_q$ has a unique cube root; coset representatives for $Q_2 = Q_4$ are 0 and 1.

Case 2^n, *n even.*

$$y^2 + \rho y = x^3,\ (a=6), \qquad y^2 + \rho y = x^3 + \alpha,\ (a=6),$$

$$y^2 + \rho^2 y = x^3,\ (a=6), \qquad y^2 + \rho y = x^3 + \beta,\ (a=6),$$

$$y^2 + y = x^3,\ (a=24), \qquad y^2 + y = x^3 + \gamma,\ (a=24),$$

$$y^2 + y = x^3 + \delta x,\ (a=4).$$

Notes. ρ denotes a primitive root; α (resp. β) denotes a nontrivial coset representative of $\{x^2+\rho x : x \in \mathbf{F}_q\}$ (resp. $\{x^2+\rho^2 x : x \in \mathbf{F}_q\}$) in $\mathbf{F}_q$. We have

A	ζA
Q_4	$\zeta^2 A$

$\zeta \in \mathbf{F}_q$ and $[Q_2 : Q_4] = 2$. Since $\zeta(x^4+x) = (\zeta x)^4 + \zeta x$, therefore $\zeta Q_4 = Q_4$. Let A denote the complement $Q_4 \backslash Q_2$. Then $\mathbf{F}_q$ is partitioned into the four cosets Q_4, A, ζA and $\zeta^2 A$.

In the list of curves, we can take any $\gamma \in \mathbf{F}_q \backslash Q_2$ and $\delta \in \mathbf{F}_q \backslash Q_4$. ■

4.7.3 The trace of Frobenius: continuation

For E over $\mathbf{F}_q$, any q, but with $j \neq 0, 1728$ we define

$$\chi(E) = \left(\frac{e}{q}\right) \quad \text{where} \quad e = \begin{cases} c_6 & \text{if } q \text{ is odd} \\ \dfrac{a_1 a_2 + a_3}{a_1^3} & \text{if } q \text{ is even.} \end{cases}$$

When $j = 1728$ and $q = p^n$ where $p \equiv 1 \bmod 4$, we define (using Lemma 4.7.1)

$$\chi(E) = c_4^{-(q-1)/4} \in \mathbf{F}_p;$$

and when $j = 0$ and $q = p^n$ where $p \equiv 1 \bmod 6$, we define

$$\chi(E) = c_6^{-(q-1)/6} \in \mathbf{F}_p.$$

There will be no need to make a more careful definition involving quartic and sextic reciprocity symbols, nor to define $\chi(E)$ in the remaining cases.

Lemma 4.7.9 *When defined, $\chi(E)$ is an $\mathbf{F}_q$-isomorphism invariant: if $E' = \tau E$ where τ is an isomorphism defined over $\mathbf{F}_q$, then (with equality in $\mathbf{F}_p$ understood in the cases $j = 0$ and 1728)*

$$\chi(E') = \chi(E).$$

Proof. If $\tau = [r, s, t, u]$ then, with e', c_4', c_6' referring to E', we have
— when $j \neq 0, 1728$:

$$e' = c_6 u^6 \text{ if } q \text{ is odd}, \quad e' = e + \left(\frac{s}{a_1}\right)^2 + \left(\frac{s}{a_1}\right) \text{ if } q \text{ is even;}$$

hence $(e/q) = (e'/q)$ in both cases;
— when $j = 1728$,

$$c_4'^{(q-1)/4} = (c_4 u^{-4})^{(q-1)/4} = c_4^{(q-1)/4},$$

and similarly when $j = 0$. ■

Proposition 4.7.10 *Let E be an elliptic curve defined over $\mathbf{F}_q$ where $q = p^n$, and as usual let j denote its invariant.*

(a) For any q and any $j \neq 0$ or 1728, $\chi(E)t(E,q)$ is an absolute invariant: if E' is also defined over $\mathbf{F}_q$ with invariant j, then

$$\chi(E)t(E,q) = \chi(E')t(E',q).$$

(b) $t(E,q) = 0$, hence $|E(\mathbf{F}_q)| = q+1$, in each of the following circumstances:

(i) *$j = 1728$ and $q \equiv 7$ or 11 mod 12, i.e., $p \equiv 3 \bmod 4$, $p > 3$ and n is odd;*

(ii) *$j = 0$ and $q \equiv 5 \bmod 6$, i.e., $p \equiv 5 \bmod 6$ and n is odd;*

(iii) *$j = 0$, $q = 3^n$, n odd, and E has a point of order 2 (which is the case, for example, when Δ is a quadratic nonresidue: $(\Delta/q) = -1$);*

(iv) *$j = 0$, $q = 3^n$, any n, and $(\Delta/q) = -1$.*

In the remaining cases we have analogous results, but only mod p :

(a′) If $j = 1728$ and $p \equiv 1 \bmod 4$, or $j = 0$ and $p \equiv 1 \bmod 6$, then

$$\chi(E)t(E,q) \equiv \chi(E')t(E',q) \bmod p.$$

(b′) $t \equiv 0 \bmod p$ in each of the following cases:

$j = 1728$ and $p \equiv 3 \bmod 4$, n even;

$j = 0$ and $p \equiv 5 \bmod 6$, n even;

$j = 0$, $q = 3^n$ or 2^n, any n. (This includes earlier cases of $q = 3^n$ where $t = 0$ — to simplify the statement.)

Remarks. Here is an example which shows that the 'mod p' qualification cannot be removed in general from statement (b').

$$y^2 + y = x^3 - 860x + 9707$$

has $\Delta = -43^3$ and $j = -2^{18}3^3 5^3$, hence defines an elliptic curve over $\mathbf{F}_p$ for $p \neq 43$. For the first few primes we have

$$j \equiv 0 \bmod p \quad \text{for} \quad p = 2,3,5 \quad \text{and} \quad j \equiv 6 \equiv 1728 \bmod 7;$$

$$t(E,p) = 0 \quad \text{and} \quad t(E,p^2) = -2p \quad \text{for} \quad p = 2,3,5,7.$$

The values of $t(E,p)$ at least can be verified without too much trouble by hand (see §4.7.4 for hints and useful tables), and there is a theoretical reason (the Riemann Hypothesis for $E_{/\mathbf{F}_q}$ — see §4.7.5) why $t(E,p) = 0 \Rightarrow t(E,p^2) = -2p$. In fact for every $E_{/\mathbf{Q}}$ there are infinitely many p satisfying $t(E,p) = 0$ (Elkies' theorem, [Elk89]).

$$y^2 + y = x^3 - 7x + 6$$

has $\Delta = 5077$ (a prime) and $j = 2^{12}3^3 7^3/5077$, hence

$$j \equiv 0 \bmod p \quad \text{for} \quad p = 2, 3, 7 \quad \text{and} \quad j \equiv 3 \equiv 1728 \bmod 5.$$

One finds $t(E, 2) = -2$, $t(E, 4) = 0$; $t(E, 3) = -3$, $t(E, 9) = 3$, $t(E, 27) = 0$ (E has a point of order 2 over $\mathbf{F}_{27}$, but not over $\mathbf{F}_3$ or $\mathbf{F}_9$; also $(\Delta/3) = 1$); $t(E, 5) = -4$, $t(E, 25) = 6$; $t(E, 7) = -4$, $t(E, 49) = 2$.

We will encounter both these curves again because of their special properties; in terms to be explained later, the first curve has everywhere good reduction in an abelian extension of $\mathbf{Q}$ and has integral j-invariant (*cf.* [Con93]), while the second played a key role in solving a problem of Gauss (*cf.* [BGZ85] and [Zag84]).

Proof. (a) Let E have invariant $j \neq 0, 1728$, there being no restriction on q. We know (Proposition 4.3.2) that E is $\mathbf{F}_q$-isomorphic to some twist E_j^d of the generic-j curve, and by lemmata 4.7.2 and 4.7.9, we can assume that $E = E_j^d$:

$$E : \begin{cases} y^2 + xy = x^3 + \dfrac{d-1}{4}x^2 - \dfrac{36d^2}{j-1728}x - \dfrac{d^3}{j-1728} & (q \text{ odd}) \\ y^2 + xy = x^3 + dx^2 + 1/j & (q \text{ even}) \end{cases}$$

By the previous proposition,

$$\left(\frac{d}{q}\right) t(E, q) = t(E_j, q).$$

This quantity depends only on j. Since

$$\chi(E) = \begin{cases} \left(\dfrac{-jd^3/(j-1728)}{q}\right) = \left(\dfrac{d}{q}\right)\left(\dfrac{-j(j-1728)}{q}\right) & (q \text{ odd}) \\ \left(\dfrac{d}{q}\right) & (q \text{ even}) \end{cases}$$

we see that $\chi(E)/\left(d/q\right)$ depends only on j. Hence $\chi(E)t(E, q)$ depends only on j, and we have the desired conclusion

$$\chi(E)t(E, q) = \chi(E')t(E', q).$$

(b) When $j = 1728$, $p > 2$ and E has a point of order 2 (which is automatic when $p > 3$ — *cf.* Proposition 4.2.2), we can take E in the form $y^2 = x^3 + ax$. When $q \equiv 3 \bmod 4$, which is so in (b)(i) and (iii), we have $(-1/q) = (-1)^{(q-1)/2} = -1$, hence

$$\left(\frac{x^3 + ax}{q}\right) = -\left(\frac{-x^3 - ax}{q}\right)$$

and therefore the $q-1$ terms with $x \neq 0$ in the sum $t = t(E,q)$ occur in cancelling pairs. Thus $t = 0$.

Similarly in cases (ii) and (iv) we get cancelling pairs. We can assume that E has the form $y^2 = f(x)$ where $f(x) = x^3 + a$, respectively $x^3 + ax + b$. Both cases follow from the fact that $x \mapsto f(x)$ is a bijection, so the quadratic residues cancel the quadratic nonresidues in the sum $t(E,q)$. In case (ii) this is because $q \equiv 2 \bmod 3$ and therefore every element has a unique cube root: $x \mapsto x^3$ is a bijection. In the other case, if $c \in \mathbf{F}_q$ and $f(x) = c$ has a root ρ then, because the characteristic is 3, the other roots are $\rho \pm \sqrt{-a}$. Since $\Delta = -a^3$, therefore by assumtion the other roots are not in $\mathbf{F}_q$. Thus for each c, $f(x) = c$ has at most one root and therefore exactly one root since there are q values of x and q values of c. Incidentally, $f(x) = 0$ has a unique root, *i.e.,* $E(\mathbf{F}_q)$ has a unique point of order 2, as remarked in the proposition.

(a') and (b'): We state a sequence of results which together complete the proof, without specifically labelling each case of (a') and (b').

First let us dispose of characteristic 2. When $p = 2$ and $j = 0$, by Proposition 1.7.10, $E(\mathbf{F}_{2^n})$ has no point of order 2 so $|E(\mathbf{F}_{2^n})| = 2^n + 1 - t$ is odd, hence t is even. Thus assume $p \geq 3$.

Consider the case $j = 0$, $q = 3^n$, E has no point of order 2, and $(\Delta/q) = 1$. We can take E in the form $y^2 = f(x) = x^3 + ax + b$ where $\Delta = -a^3$ and therefore $(-a/q) = 1$. If $c \in \mathbf{F}_q$ and $f(x) = c$ has a root $\rho \in \mathbf{F}_q$, then the other roots $\rho \pm \sqrt{-a}$ are also in $\mathbf{F}_q$, and the three roots are distinct since the derivative $f'(x) = a \neq 0$. The values c which do occur do not include 0 since E has no point of order 2, and therefore the 3 values of x giving $f(x) = c$ contribute either $+3$ or -3 to the sum t, according as c is a quadratic residue or not. Hence $t \equiv 0 \bmod 3$.

We recall that for $k \in \mathbf{Z}$

$$\sum_{x \in \mathbf{F}_q^*} x^k = \begin{cases} -1 & \text{if } k \equiv 0 \bmod q-1 \\ 0 & \text{otherwise} \end{cases} \qquad \P$$

This is a simple consequence of the fact that ${\mathbf{F}_q}^*$ is cyclic.

When E is of the form $y^2 = x^3 + ax$, ¶ can be applied to

$$t = -\sum_{x \in \mathbf{F}_q^*} \left(\frac{x^3 + ax}{q} \right) \equiv -\sum_{x \in \mathbf{F}_q^*} (x^3 + ax)^{(q-1)/2} \bmod p$$

by expanding the terms on the right and adding like powers of x:

$$t \equiv -\sum_{i=0}^{(q-1)/2} \sum_{x \in \mathbf{F}_q^*} \binom{(q-1)/2}{i} x^{2i+(q-1)/2} a^{(q-1)/2-i} \bmod p.$$

By ¶, the only contribution occurs when $2i + (q-1)/2 = q-1$, which requires $q \equiv 1 \bmod 4$ and then

$$t \equiv \binom{(q-1)/2}{(q-1)/4} a^{(q-1)/4} \bmod p.$$

Case $p \equiv 3 \bmod 4$, n even: we can choose a such that $a^{(q-1)/4} \notin \mathbf{F}_p$, for instance a a generator of ${\mathbf{F}_q}^*$; *c.f.* Lemma 4.7.1. Since t and the binomial coefficient are integers, we draw the following purely 'combinatorial' corollary:

$$\binom{(q-1)/2}{(q-1)/4} \equiv 0 \bmod p \quad \text{for } q = p^n,\ n \text{ even and } p \equiv 3 \bmod 4.$$

(I do not know an attractive, direct proof of this congruence; but at least I now know 'the reason why', for example, $\binom{4}{2}$ is divisible by 3.) Going back to arbitrary a, we have $t \equiv 0 \bmod p$ for all E occuring in this case.

Case $p \equiv 1 \bmod 4$: Since $c_4 = -48a$, we see that $\chi(E)t(E,q)$ has a constant value mod p for E with $j = 1728$:

$$\chi(E)t(E,q) \equiv (-3)^{-(q-1)/4} \binom{(q-1)/2}{(q-1)/4} \bmod p.$$

Similarly, when E has the form $y^2 = x^3 + a$ and $q \equiv 1 \bmod 6$, ¶ implies

$$t(E,q) \equiv \binom{(q-1)/2}{(q-1)/6} a^{(q-1)/6} \bmod p,$$

which leads in the same way to

$$\binom{(q-1)/2}{(q-1)/6} \equiv 0 \bmod p \quad \text{for } q = p^n,\ n \text{ even and } p \equiv 5 \bmod 6,$$

hence $t \equiv 0 \bmod p$ in such cases; and when $p \equiv 1 \bmod 6$, $\chi(E)t(E,q)$ has a constant value mod p. ■

Corollary 4.7.11 ([Cox89, p.320]) *Let E, E' be defined over $\mathbf{F}_q$ where $q = p^n$, and suppose $t(E,q) \not\equiv 0 \bmod p$. Then E, E' are $\mathbf{F}_q$-isomorphic iff*

$$j(E) = j(E') \quad \textit{and} \quad t(E,q) = t(E',q).$$

Remarks. In Chapter 1 we classified elliptic curves defined over fields of characteristic $p > 0$ into two types: *supersingular* and *ordinary*. (The definitions are given just before Proposition 1.7.10.) We will prove in Chapter 6 that the E considered in this proposition are precisely the ordinary ones, in other words, for E defined over $\mathbf{F}_q$,

- E is supersingular iff $t(E,q) \equiv 0 \bmod p$, hence

- E is ordinary iff $t(E,q) \not\equiv 0 \bmod p$.

This will amount to proving the following two facts:

(a) if $t(E,q) \equiv 0 \bmod p$ then $t(E,q^f) \equiv 0 \bmod p$ for all $f > 0$, hence E is supersingular as defined in Chapter 1;

(b) if $t(E,q) \not\equiv 0 \bmod p$ then $t(E,q^f) \equiv 1$ for some $f > 0$, so $|E(\mathbf{F}_{q^f})| \equiv 0 \bmod p$, hence $E(\mathbf{F}_{q^f})$ contains a point of order p, and E is ordinary as defined in Chapter 1.

Accepting these statements, from the proposition, E is supersingular when $j = 0$ and $p = 2, 3$ or $q \equiv 5 \bmod 6$; and when $j = 1728$ and $q \equiv 3 \bmod 4$. An example where $j \neq 0$ or 1728 is $y^2 = x^3 - 4x + 4$ over $\mathbf{F}_{13}$. This curve has $j = 5$; note that $1728 \equiv 12 \bmod 13$. The computer tells us that $E(\mathbf{F}_{13})$ has order 14 and that one generator of this cyclic group is the point $(x, y) = (2, 2)$.

Thus the corollary can be stated:

> *two ordinary elliptic curves defined over $\mathbf{F}_q$ are isomorphic over $\mathbf{F}_q$ iff they have the same invariant and the same trace.*

Proof. Obviously the two conditions are necessary. Conversely assume these two conditions.

If $j \neq 0$ or 1728 the result follows from Proposition 4.7.5: since the j's are equal, $E' = E^d$ for some d, and since the t's are equal and nonzero, $(d/q) = 1$. This means that the curves are $\mathbf{F}_q$-isomorphic.

The remaining cases follow from the present proposition and Proposition 4.4.1. For suppose $j = 1728$ (resp. 0). Since $t \not\equiv 0 \bmod p$, we have $p \equiv 1 \bmod 4$ (resp. mod 6) and therefore $\chi(E) \equiv \chi(E') \bmod p$. Hence ${c_4'}^{(q-1)/4} \equiv c_4^{(q-1)/4} \bmod p$ (resp. ${c_6'}^{(q-1)/6} \equiv c_6^{(q-1)/6} \bmod p$). By Lemma 4.7.1 this implies that $\sqrt[4]{c_4/c_4'} \in \mathbf{F}_q$ (resp. $\sqrt[6]{c_6/c_6'} \in \mathbf{F}_q$), hence E and E' are $\mathbf{F}_q$-isomorphic. ■

4.7.4 Examples

We present some tables that allow us to calculate $|E(\mathbf{F}_q)|$ quickly for the first few values of q. We write $\mathbf{F}_q = \mathbf{F}_p(\theta)$. Reminder: *quadratic residue* has an unconventional meaning when q is even — see Section 4.7.1.

q	θ	quadratic residues
2	–	0
3	–	1
4	$\theta^2 = \theta + 1$	$0, 1$
5	–	$1, 4$
7	–	$1, 2, 4$
8	$\theta^3 = \theta + 1$	$0, \theta, \theta^2, \theta^2 + \theta$
9	$\theta^2 = 2$	$1, 2, \theta, 2\theta$
11	–	$1, 4, 5, 9, 10$
etc.		

We denote the entry for j in the following tables by $t^{(j)}$. When $j \neq 0$ or 1728, $t^{(j)}$ is the trace of Frobenius of a curve with the given j as invariant and with $\chi(E) = 1$. ‡ It follows from Proposition 4.7.10 that

$$\text{when } j \neq 0, 1728, \quad t(E, q) = \chi(E) t^{(j)}.$$

The entries for $j = 0$ and $j = 1728$ (the latter being distinguished by the symbol †) are, when $p > 3$, $t(E_0, q)$ for a curve with $c_6 = 1$, resp. $c_4 = 1$ and therefore

$$\text{when } j = 0, \quad t(E, q) \equiv c_6^{(q-1)/6} t^{(j)} \bmod p;$$

$$\text{when } j = 1728, \quad t(E, q) \equiv c_4^{(q-1)/4} t^{(j)} \bmod p.$$

When $p = 2$ resp. 3 the entries for $j = 0$ are the traces of Frobenius of the two special curves $y^2 + y = x^3$ and $y^2 = x^3 + x$.

When Proposition 4.7.10 does not state that $t = 0$, one must use some additional information to obtain t exactly. Various points are illustrated by examples after the tables.

We present the tables of $t^{(j)}$ for the prime fields up to $p = 17$ followed by tables for $q = 4, 8, 9$.

‡Specifically,

$$t^{(j)} = \left(\frac{1728 - j}{q}\right) t(E_j, q) \text{ when } p \neq 2, \quad t(E_j, q) \text{ when } p = 2.$$

Table of $t^{(j)}$. A † means $j \equiv 1728 \bmod p$.

$p\backslash j$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	0†	−1															
3	0†	1	−2														
5	0	2	3	−4†	−1												
7	−1	−3	−4	−1	−2	2	0†										
11	0	0†	4	−6	3	5	1	−2	4	−2	3						
13	−5	6	1	5	−2	0	2	4	−3	1	−4	2	6†				
17	0	−6	−3	−1	4	4	6	−3	0	2	−6	8†	−7	−3	−2	−5	2

$q = 4$:
$\theta^2 = \theta + 1$

j	$t^{(j)}$
0	−4
1	−3
θ	1
$\theta + 1$	1

$q = 8$:
$\theta^3 = \theta + 1$

j	$t^{(j)}$
0	0
1	5
θ	−3
$\theta + 1$	1
θ^2	−3
$\theta^2 + 1$	1
$\theta^2 + \theta$	−3
$\theta^2 + \theta + 1$	1

$q = 9$:
$\theta^2 = 2$

j	$t^{(j)}$
0	−6
1	−5
2	−2
θ	1
$\theta + 1$	−2
$\theta + 2$	4
2θ	1
$2\theta + 1$	−2
$2\theta + 2$	4

We illustrate the use of the tables by considering the curve

$$y^2 + xy + y = x^3 - 3x + 1. \qquad \textbf{A105}$$

We will use E loosely to denote the curve given by this equation interpreted over various fields K. As a curve over $\mathbf{Q}$ it has data

$$c_4 = 11^2, \quad c_6 = -13 * 97, \quad \Delta = 3 * 5 * 7, \quad j = \frac{11^6}{3 * 5 * 7},$$

and is therefore singular when $\operatorname{char} K = 3, 5$ and 7; we considered this example at the end of Chapter 1 where we saw that the singularity in each of these three characteristics is split multiplicative. We notice the point $P = (1, -1) \in E(\mathbf{Q})$ of order 2. (As an exercise in Nagell-Lutz and simple 2-decent, the reader may verify that $E(\mathbf{Q}) = \{O, (1, -1)\}$.)

When we regard E as an elliptic curve over $\mathbf{F}_q$ (with $p \neq 3, 5, 7$), $(1, -1)$ will remain as a point of order 2, and therefore we can predict that $|E(\mathbf{F}_q)|$ is even.

We simplify the notation $t(E,q)$ to t_q.

Let $q=2$. Then $j=1$ in $\mathbf{F}_2$ and in the notation used to define $\chi(E)$, $e=1$ and $(e/q)=(1/2)=-1$ from the table of quadratic residues. The t-table entry for $j=1$ is -1, hence $t_2=-1*(-1)=1$ and $|E(\mathbf{F}_2|=2$. This is confirmed by direct inspection:

$$E(\mathbf{F}_2)=\{O,(1,1)\}.$$

Of course here $(1,-1)=(1,1)$. When we enlarge the field to $\mathbf{F}_4$, e stays the same but it is now a quadratic residue, *i.e.,* $(1/4)=1$. We have $t_4=1*(-3)=-3$ and $|E(\mathbf{F}_4)|=8$. Similarly over $\mathbf{F}_8$ we find $t_8=-1*5=-5$ and $|E(\mathbf{F}_4)|=14$; this does not contradict the previous result since $\mathbf{F}_4$ is not a subfield of $\mathbf{F}_8$.

Next, $j\equiv 0 \bmod 11$ and $11\equiv 5 \bmod 6$, so $t_{11}=0$.

$j\equiv 12\equiv 1728 \bmod 13$ and $c_4\equiv 4 \bmod 13$ and therefore

$$t_{13}\equiv 4^3*6\equiv 7 \bmod 13.$$

We also know that t_{13} is even and, as a very crude estimate, $|t_{13}|\le 12$ which is obtained by taking an isomorphic curve $y^2=x^3+ax$ so that

$$|t_{13}|\le \sum_{x\in\mathbf{F}_{13}}\left|\left(\frac{x^3+ax}{13}\right)\right|\le 12.$$

This is sufficient information to determine that $t_{13}=-6$ and $|E(\mathbf{F}_{13})|=20$.

$j\equiv 14\not\equiv 1728 \bmod 17$, also $c_6\equiv 14 \bmod 17$. Using quadratic reciprocity,

$$\chi(E)=(e/q)=(2/17)(7/17)=(3/7)=-1,\quad \text{hence}\quad t_{17}=-t^{(14)}=2.$$

Thus $|E(\mathbf{F}_{17})|=16$.

Finally we consider the supersingular examples over $\mathbf{F}_3$ and $\mathbf{F}_2$. Over $\mathbf{F}_3$, the three curves $y^2=x^3+x+a$, $a\in\{0,1,-1\}$ are isomorphic and have $t=0$; the remaining three curves with $j=0$ are $y^2=x^3-x+a$ where $a=0,1,-1$ and they have $t=0,-3,3$ respectively. Over $\mathbf{F}_2$, $y^2+y=x^3+a$, $a\in\{0,1\}$ are isomorphic and have $t=0$, while $y^2+y=x^3+x+a$, $a=0,1$ have $t=-2,2$ respectively.

4.7.5 A preview of some future topics

There are a number of interesting matters concerning the trace of Frobenius whose discussion must be postponed.

— The Riemann Hypothesis (R.H.) for $E_{/\mathbf{F}_q}$: for any E over $\mathbf{F}_q$, we have $|t|\le 2\sqrt{q}$. This was conjectured by E. Artin in his thesis and subsequently proved by Hasse [Has33]. Let us elaborate a little.

In Chapter 6 we define the Frobenius endomorphism of E and show that it has a representation as a linear operator on a 2-dimensional vector space over the complex field. The characteristic polynomial is x^2-tx+q, and therefore

the trace of this operator is t, hence the name *trace of Frobenius*. Thus the R.H. is the statement that the two eigenvalues are conjugate complex numbers, or both equal $2\sqrt{q}$. Let us denote these two numbers π and $\overline{\pi}$. Then for all positive integers k

$$t(E, q^k) = \pi^k + \overline{\pi}^k.$$

Hence the group order $|E(\mathbf{F}_q)|$ immediately determines $|E(\mathbf{F}_{q^k})|$ for all overfields.

Manin [Man56] gave an elementary proof of the R.H. in the case $q = p$. Subsequently Chahal [Cha95] gave a simplified proof extended to $q = p^n$, but only for $p \geq 5$. The proof given in Chapter 6, based on the general notion of isogeny, is shorter and applies to all q.

— When does there exist an E with a given t within the range allowed by the R.H.? For now we only quote the answer; references are [Wat69] and [Rüc87].

The t satisfying $|t| \leq 2\sqrt{q}$ that actually occur are those that satisfy one of the following conditions; $q = p^n$ as usual.

(a) $\gcd(t, p) = 1$;

(b) *n even, $t = \pm 2\sqrt{q}$;*

(c) *n even, $p \not\equiv 1 \bmod 3$, $t = \pm\sqrt{q}$;*

(d) *n odd, $p = 2$ or 3, $t = \pm p^{(n+1)/2}$;*

(e) *n odd, $t = 0$;*

(f) *n even, $p \not\equiv 1 \bmod 4$, $t = 0$.*

Thus all values allowed by the Riemann Hypothesis occur for the prime field $\mathbf{F}_p$; however no E defined over $\mathbf{F}_8$ has $t = \pm 2$, *i.e.*, $|E(\mathbf{F}_8)| = 7$ and 11 do not occur.

— How many E for a given (allowable) t? For the prime field $\mathbf{F}_p$ the number of equations $y^2 = x^3 + bx + c$ with the given t (where $|t| < 2\sqrt{p}$) is given by a precise formula involving the "weighted Hurwitz class number" — see [Cox89, p.319].

— When $j = 0$, resp. 1728 and $q = p \geq 5$ there are explicit formulas for t that involve sextic, resp. quartic reciprocity symbols; the proofs involve Gauss and Jacobi sums. See the last chapter in the text by Ireland and Rosen [Ire-Ro82].

— Let E be defined over $\mathbf{Z}$ (for simplicity) so that apart from finitely many 'bad' primes, E mod p is an elliptic curve and on the basis of the R.H. we can write $t(E, p) = 2\sqrt{p}\cos\theta_p$ where $0 < \theta_p < \pi$. Suppose that E does not have complex multiplication; this term will be defined in Chapter 9. Then the Sato-Tate conjecture is that if $\mathrm{ST}(x, a, b)$ denotes the number of $p < x$ such that $a < \theta_p < b$ (where $0 \leq a < b \leq \pi$), and $\pi(x)$ denotes the number of $p < x$ then

$$\lim_{x\to\infty} \mathrm{ST}(x, a, b) \Big/ \pi(x) \int_a^b \sin^2\theta \, d\theta = \frac{2}{\pi}.$$

The number $2/\pi$ on the right is $\left(\int_0^\pi \sin^2\theta\, d\theta\right)^{-1}$. In other words, θ_p follows a $\sin^2$ distribution: for large x and small positive $d\theta$, the number of $p < x$ with θ_p between θ and $\theta + d\theta$ is approximately

$$\frac{2}{\pi}\pi(x)\sin^2\theta\, d\theta.$$

See [Cas66] for original references.

— Elkies [Elk89] solved a long outstanding problem by proving that there are infinitely many p for which $\theta_p = \pi/2$, *i.e.,* for which E is supersingular.

Chapter 5

Minimal Weierstrass Equations

The basic ideas of this chapter are illustrated by the example of the elliptic curve

$$E: \quad y^2 = x^3 + \frac{1}{4}x^2 + \frac{1}{2}x + \frac{1}{4}, \quad \Delta = -26$$

defined over $\mathbf{Q}$. The $\mathbf{Q}$-isomorphism $[0, 0, 0, 2^{-1}]$ (using the notation introduced in §4.1) can be used to "clear denominators" — the transformed equation E' has all $a_i \in \mathbf{Z}$, hence we say that it is defined over $\mathbf{Z}$:

$$E' = [0, 0, 0, 2^{-1}]E: \quad y^2 = x^3 + x^2 + 8x + 16, \quad \Delta = -2^{13}13.$$

Among all the E' isomorphic to E and defined over $\mathbf{Z}$ there are those with minimal $|\Delta|$; these are called $\mathbf{Z}$-minimal (Weierstrass) models. For this example one such model is

$$[0, 1/2, 1/2, 1]E: \quad y^2 + xy + y = x^3, \quad \Delta = -26. \qquad \textbf{A26}$$

This minimal model is unique if we also impose the requirements $a_1, a_3 \in \{0, 1\}$ and $a_2 \in \{-1, 0, 1\}$.

To analyze this concept in greater generality, one first considers *local* minimal models: for a valuation v, Weierstrass equations with all $v(a_i) \geq 0$ and with $v(\Delta)$ minimal. Then one attempts to globalize, say to a Krull domain A: find a Weierstrass equation with coefficients $a_i \in A$ which is simultaneously v-minimal for all essential valuations v. When the ring is not a PID, such global minimal models do not always exist.

The key tool, at least for fields of characteristic 0, is a result of Kraus that is presented in Section 2. It will be used to obtain a much improved version of Laska's algorithm — we call it the Laska-Kraus algorithm — which finds global

minimal models of E defined over number fields, or determines that none exists. The simplicity of the algorithm in the case of $\mathbf{Z}$ is particularly striking. Kraus's result will also be essential to the solution of several other problems.

5.1 Some definitions

Let E and E' be elliptic curves over the field K. If A is a subring of K, or more generally a subring of an extension field of K, we say that E and E' are ***A*-isomorphic** if there exists a transformation $\tau = [r, s, t, u]$ as in the previous chapter with $r, s, t \in A$ and $u \in A^*$ such that $\tau E = E'$. We also call such a τ an ***A*-isomorphism**.

This definition can be expressed in terms of a more primitive concept: a transformation $\tau = [r, s, t, u]$ is ***A*-integral** if $r, s, t, u \in A$. Since

$$\tau^{-1} = [-u^{-2}r, -u^{-1}s, u^{-3}(rs - t), u^{-1}],$$

τ is an A-isomorphism iff both τ and τ^{-1} are A-integral.

We say that E is **defined over *A***, or that E is ***A*-integral**, when all the $a_i \in A$; this can be indicated by the notation $E_{/A}$. The quantities $b_2, \ldots, b_8, c_4, c_6, \Delta$ are then all in A; however $j = c_4^3/\Delta$ of course need not be in A.

We define two polynomials:

$$\begin{aligned} \Psi_2(x) &:= x^3 - 3c_4x - 2c_6, \\ \Psi_3(x) &:= x^4 - 6c_4x^2 - 8c_6x - 3c_4^2; \end{aligned}$$

Their discriminants are

$$\mathrm{Dis}\,(\Psi_2) = 2^8 3^6 \Delta, \qquad \mathrm{Dis}\,(\Psi_3) = -2^{24} 3^9 \Delta^2,$$

and their relations to the division polynomials are

$$\begin{aligned} \Psi_2(12x + b_2) &= 432\psi_2^2, \quad \text{hence} \quad \Psi_2(b_2) = 2^4 3^3 b_6, \\ \Psi_3(12x + b_2) &= 6912\psi_3, \quad \text{hence} \quad \Psi_3(b_2) = 2^8 3^3 b_8. \end{aligned}$$

Here are some identities that we will use; the last three can be checked on the computer.

$$\Psi_3(x)' = 4\Psi_2(x), \tag{\#i}$$

$$\Psi_3(x) = 4x\Psi_2(x) - 3(x^2 - c_4)^2, \tag{\#ii}$$

$$(x^2 - c_4)^3 = \Psi_2(x)(x^3 + 2c_6) + 3(c_4x + c_6)^2 + 1728\Delta, \tag{\#iii}$$

$$(x^2 - 3c_4)\Psi_3(x) + (x^2 - c_4)^3 = 2\Psi_2(x)^2 + 24^3\Delta, \tag{\#iv}$$

Let v be a valuation on K, let E be an elliptic curve defined over K, and let a, b, c denote $v(c_4), v(c_6), v(\Delta)$ respectively. Since $c_4^3 - c_6^2 = 1728\Delta$, the

minimum among $3a, 2b, c + v(1728)$ occurs at least twice. To discuss candidate triples of values we define a **signature** to be an ordered triple a, b, c where $a, b \in \{0, 1, 2, \ldots\} \cup \{\infty\}$ with a, b not both ∞, and $c \in \{0, 1, 2, \ldots\}$; and we define a signature a, b, c to be **admissible** when the minimum among $3a, 2b, c+v(1728)$, with natural definitions concerning ∞, occurs at least twice. For example, when $v(3) = 1$, no signature of the form $a, 1, c$ is admissible since $v(1728) = 3v(3) = 3$ and the minimum among $3a, 2, c + 3$, which is either 0 or 2, occurs only once. Most, but not all, admissible signatures actually occur; see Proposition 5.2.3.

By elementary reasoning one compiles the list of admissible signatures (w denotes $v(1728)$):

$$\begin{array}{rl} 2d, 3d, c, & c, d \in \{0, 1, 2, \ldots\},\ c \geq 6d - w, \\ a, b, 3a - w, & a, b \in \{0, 1, 2, \ldots\} \cup \{\infty\},\ w/3 \leq a < \infty,\ 3a/2 < b, \\ a, b, 2b - w, & a, b \in \{0, 1, 2, \ldots\} \cup \{\infty\},\ w/2 \leq b < \infty,\ 2b/3 < a. \end{array}$$

The following notation is convenient.

Let a and m be integers with m positive. The **symmetric residue** of a mod m is the unique integer denoted $\operatorname{mods}(a, m)$ satisfying

$$\operatorname{mods}(a, m) \equiv a \bmod m, \quad \text{and} \quad -m/2 < \operatorname{mods}(a, m) \leq m/2.$$

5.2 Kraus's theorem

Let V be a discrete valuation ring with quotient field K, and let $v : K \longrightarrow \mathbf{Z} \cup \{\infty\}$ denote the valuation map as usual. Thus a K-isomorphism $\tau = [r, s, t, u]$ is V integral when $v(r), v(s), v(t)$ and $v(u)$ are all ≥ 0, and is a V-isomorphism when additionally $v(u) = 0$.

Now let K be a field of characteristic $\neq 2$ or 3. Then $1728 \neq 0$ in K, and a pair of elements $c_4, c_6 \in K$ for which $\Delta := (c_4^3 - c_6^2)/1728 \neq 0$ unambiguosly determines a K-isomorphism class of elliptic curves by the c-form

$$\eta^2 = \xi^3 - \frac{c_4}{48}\xi - \frac{c_6}{864};$$

(Lemma 4.2.1 explains to what extent conversely the isomorphism class of these curves determines the pair c_4, c_6.)

Question: When are two given elements c_4, c_6 of V realized as the covariants of some Weierstrass equation defined over V?

First, if $\operatorname{char} \tilde{v} \neq 2, 3$, equivalently $v(2) = v(3) = 0$, the c-form makes it clear that arbitrarily chosen $c_4, c_6 \in V$, subject only to $\Delta \neq 0$, occur for some E defined over V.

In contrast there is considerable linkage between $v(c_4)$ and $v(c_6)$ when $\operatorname{char} K = 0$ and the residue field characteristic is 2 or 3.[†] The next proposition is due to Kraus [Kra89]; we have written the criteria in the case $p = 2$ in a different way and we have added uniqueness statements that will be needed in applications.

Proposition 5.2.1 *Let* $\operatorname{char} K = 0$*, let* v *be a valuation on* K *with ring* V *and residue field characteristic* p*, and let* c_4, c_6, Δ *be elements of* V *satisfying*

$$c_4^3 - c_6^2 = 1728\Delta \neq 0. \tag{¶}$$

Then c_4, c_6 *occur as the covariants of some* $E_{/V}$ *iff:*

- *when* $\boxed{p = 3}$ $\exists \vartheta \in V$ *satisfying*

$$\Psi_2(\vartheta) \equiv 0 \bmod 27;$$

this is always true when $v(c_4) = 0$ *(take* $\vartheta = -c_6/c_4$*) and when* $v(c_6) \geq v(27)$ *(take* $\vartheta = 0$*); moreover for any such* ϑ*,*

$$\vartheta^2 \equiv c_4 \bmod 3 \quad \textit{and} \quad \vartheta c_4 + c_6 \equiv 0 \bmod 3;$$

when a solution exists it is unique mod 3*: if* ϑ *is a solution then so is* $\vartheta + 3\alpha$*,* $\forall \alpha \in V$*, and conversely the difference between any two solutions is of the form* 3α*,* $\alpha \in V$*; if* $E_{/V}$ *is a model with these covariants then a solution is* $\vartheta = b_2$*;*

- *when* $\boxed{p = 2}$ $\exists \theta, \tau \in V$ *satisfying*

$$\Psi_3(\theta^2) \equiv 0 \bmod 256 \quad \textit{and} \quad \Psi_2(\theta^2) \equiv -16\tau^2 \bmod 64;$$

when a solution exists it is unique mod 2*: if* θ, τ *is a solution then so is* $\theta + 2\alpha, \tau + \theta\alpha(\theta + \alpha) + 2\beta$ $\forall \alpha, \beta \in V$ *and all solutions are obtained this way; if* $E_{/V}$ *is a model with these covariants then a solution is* $\theta = a_1, \tau = a_3 + a_1 a_2$*. These requirements simplify in the following cases:*

(i) *when* $v(c_4) = 0$*: iff* $\exists x \in V$ *satisfying* $c_6 \equiv -x^2 \bmod 4$ *(take* $\theta = c_4/x, \tau = -(x^2 + c_6)/(4x)$*);*

(ii) *when* $v(c_4) \geq v(16)$*: iff* $\exists x \in V$ *satisfying* $c_6 \equiv 8x^2 \bmod 32$ *(take* $\theta = 0, \tau = x$*).*

Proof. $\boxed{p = 3}$ If $E_{/V}$ exists with the specified covariants then $\vartheta = b_2$ satisfies

$$\Psi_2(\vartheta) \equiv 0 \bmod 27, \quad \vartheta^2 \equiv c_4 \bmod 3, \quad \vartheta c_4 + c_6 \equiv 0 \bmod 3.$$

[†]The logically remaining cases are $\operatorname{char} K = 2$ and $\operatorname{char} K = 3$. But then the question is not meaningful; or perhaps one should say, this is not the right question in these cases.

Conversely if $\vartheta \in V$ satisfies $\Psi_2(\vartheta) \equiv 0 \bmod 27$, applying the transformation $[\vartheta/12, 0, 0, 1]$ to $y^2 = x^3 - (c_4/48)x - (c_6/864)$ gives

$$y^2 = x^3 + (\vartheta/4)x^2 + (\vartheta^2 - c_4)/48x + \Psi_2(\vartheta)/1728.$$

This has covariants c_4, c_6 and we must prove that $\vartheta^2 \equiv c_4 \bmod 3$, so this equation is defined over V, and that $c_4\vartheta + c_6 \equiv 0 \bmod 3$. Using $\Psi_2(\vartheta) \equiv 0 \bmod 27$ and $c_4^3 \equiv c_6^2 \bmod 27$ which comes from (¶), we obtain

$$(c_4\vartheta + c_6)^3 \equiv 3c_6(c_4\vartheta + c_6)^2 \bmod 27,$$

hence $c_4\vartheta + c_6 \equiv 0 \bmod 3$. Now the identity (#iii) implies $\vartheta^2 \equiv c_4 \bmod 3$.

If ϑ is a solution and $\alpha \in V$ then Taylor's expansion gives

$$\begin{aligned}\Psi_2(\vartheta + 3\alpha) &= \Psi_2(\vartheta) + (3\vartheta^2 - 3c_4)(3\alpha) + (3\vartheta)(3\alpha)^2 + (3\alpha)^3 \\ &\equiv 0 \bmod 27\end{aligned}$$

since $\vartheta^2 \equiv c_4 \bmod 3$.

Conversely if ϑ and $\vartheta + 3\alpha$ are solutions, we wish to prove that $v(\alpha) \geq 0$. Denoting $v(3)$ and $v(\alpha)$ by e and β, the values of the terms in the above congruence are

$$\geq 3e,\ \geq 3e + \beta,\ \geq 3e + 2\beta,\ 3e + 3\beta,\ \geq 3e,$$

meaning that $v(\Psi_2(\vartheta)) \geq 3e, \ldots, v((3\alpha)^3) = 3e + 3\beta$, and the last value $\geq 3e$ corresponding to the term on the right represented by 0 mod 27. The minimum value must occur at least twice in an equation, and therefore $3e + 3\beta \geq$ at least one of the other values. Hence $\beta \geq 0$.

$\boxed{p = 2}$

If there exists such an E then, since 3 is invertible in V, we can apply the transformation $[-a_2/3, 0, 0, 1]$ to obtain $a_1' = a_1, a_2' = 0, a_3' = a_3 - a_1a_2/3$ and with no change in c_4, c_6, Δ. Dropping the primes from the notation, $a_2 = 0$ hence $b_2 = a_1^2$. Taking $\theta = a_1$ and $\tau = a_3$ we have

$$\begin{aligned}\Psi_3(\theta^2) &= 2^8 3^3 b_8 \equiv 0 \bmod 256, \\ \Psi_2(\theta^2) &= 2^4 3^3 b_6 \equiv -16\tau^2 \bmod 64.\end{aligned}$$

Conversely suppose we have θ, τ. To prove the existence of $E_{/V}$ we begin by choosing $a_1 = \theta$ and $a_2 = 0$; we must prove that v-integral a_3, a_4, a_6 can be chosen so that

$$\begin{aligned}c_4 &= \theta^4 - 24(\theta a_3 + 2a_4), \\ c_6 &= -\theta^6 + 36\theta^2(\theta a_3 + 2a_4) - 216(a_3^2 + 4a_6).\end{aligned}$$

The assumptions imply that

$$16\theta^2\Psi_2(\theta^2) - \Psi_3(\theta^2) \equiv 0 \bmod 256$$

which can be rewritten by a simple calculation as

$$(\theta^4 - c_4 - 8\tau\theta)(\theta^4 - c_4 + 8\tau\theta) \equiv 0 \bmod 256.$$

Thus at least one factor on the left is divisible by 16. We choose the sign of τ so it is the left factor and then

$$a_3 := \tau, \qquad a_4 := (\theta^4 - c_4 - 24\tau\theta)/48$$

are v-integral. Now a_6 is determined by the equation for c_6 and it remains to check that $v(a_6) \geq 0$. Multiplying the equation for c_6 by 2 this follows from

$$\begin{aligned} 1728a_6 &= -2c_6 - 2\theta^6 + 72\tau\theta^3 + 3\theta^2(\theta^4 - c_4 - 24\tau\theta) - 432\tau^2 \\ &\equiv -2c_6 + \theta^6 - 3c_4\theta^2 + 27(\theta^6 - 3c_4\theta^2 - 2c_6) \bmod 64 \\ &\equiv 28(\theta^6 - 3c_4\theta^2 - 2c_6) \bmod 64 \\ &\equiv 0 \bmod 64. \end{aligned}$$

Suppose that θ, τ is a solution and $\alpha, \beta \in V$. The two congruences and the identity (#ii) imply $\theta^4 - c_4 \equiv 8\theta\tau \bmod 16$. Using also the identity $\Psi_3(x)' = 4\Psi_2(x)$ and writing $(\theta + 2\alpha)^2 = \theta^2 + 4\gamma$, Taylor expansion of $\Psi_3(\theta^2 + 4\gamma)$ yields

$$\Psi_3(\theta^2) + 4(-16\tau^2)(4\gamma) + (6\theta^4 - 6c_4)(4\gamma)^2 + (4\theta^2)(4\gamma)^3 + (4\gamma)^4$$

$$\equiv 0 \bmod 256.$$

An equally straightforward calculation shows that the second congruence is satisfied by $\theta + 2\alpha$ and $\tau + \theta\alpha(\theta + \alpha) + 2\beta = \tau + \theta\gamma + 2\beta$.

Conversely suppose θ, τ and $\theta + 2\alpha, \tau + \theta\alpha(\theta + \alpha) + 2\beta$ are both solutions; we wish to prove that $v(\alpha) \geq 0$, equivalently $v(\gamma) \geq 0$ where $\gamma = \alpha^2 + \theta\alpha$ as above, and $v(\beta) \geq 0$. Denoting $v(2)$ and $v(\gamma)$ by e and δ, the values of the terms in the above congruence are

$$\geq 8e,\ \geq 8e + \delta,\ \geq 8e + 2\delta,\ \geq 8e + 3\delta,\ 8e + 4\delta,\ \geq 8e,$$

where the last value is that of the term represented by 0 mod 256. As in the case $p = 3$, $8e + 4\delta \geq$ at least one of the other values. Hence $\delta \geq 0$. Secondly, the Ψ_2 congruences imply

$$\begin{aligned} -16((\tau + \theta\gamma + 2\beta)^2 - \tau^2) &\equiv 3(\theta^4 - c_4)(4\gamma) + 3\theta^2(4\gamma)^2 \\ &\equiv 32\theta\tau\gamma + 48\theta^2\gamma^2 \bmod 64 \end{aligned}$$

which reduces to

$$64\beta^2 + 64\beta(\tau + \theta\gamma) \equiv 0 \bmod 64.$$

Again the principle that the minimum value must be attained at least twice yields the result $v(\beta) \geq 0$.

Now suppose $v(c_4) = 0$, equivalently $v(c_6) = 0$ by (¶). If $E_{/V}$ exists then

$$-c_6 \equiv b_2^3 \equiv a_1^6 \bmod 4$$

is a square mod 4. Conversely suppose $-c_6 \equiv x^2 \bmod 4$. We can write this as

$$x^2 = -c_6(1 + 4a), \qquad a \in V.$$

Secondly, (¶) and $v(c_6) = 0$ imply

$$c_4^3 = c_6^2(1 + 64b), \qquad b \in V.$$

Defining $\theta = c_4/x$ we have

$$c_4 = \frac{\theta^4(1+4a)^2}{1+64b}, \quad c_6 = -\frac{\theta^6(1+4a)^3}{(1+64b)^2}$$

and substituting these values we immediately find

$$\Psi_3(\theta^2) \equiv 0 \bmod 256, \quad \Psi_2(\theta^2) \equiv -16a^2\theta^6 \bmod 64,$$

so we can take $\tau = a\theta^3 \equiv (x^2 + c_6)/(4x) \bmod 2$.

Finally suppose $v(c_4) \geq 4e$ where e denotes $v(2)$. If $E_{/V}$ exists then from the formulas for the b's and c's we see in succession that $v(a_1) \geq e$, $v(b_4) \geq e$, $v(b_2) \geq 2e$, hence

$$c_6 \equiv -216b_6 \equiv 8a_3^2 \bmod 32.$$

Conversely if $c_6 \equiv 8x^2 \bmod 32$ then $\theta = 0, \tau = x$ satisfy the requirements. ■

Here is an example from [Kra89] that illustrates the non-redundancy of the conditions in the case $p = 2$: let $\pi^6 = 2$, $K = \mathbf{Q}(\pi)$ or $\mathbf{Q}_2(\pi)$, $c_4 = 8\pi^2$, and $c_6 = 8\pi^3$. Thus $v(2) = 6$, $v(c_4) = 20$ and $v(c_6) = 21$. Then $\theta = \pi^5$ satisfies $\Psi_3(\theta^2) \equiv 0 \bmod 256$ but

$$\Psi_2(\theta^2) = -16\pi^3(1 + \pi^9)$$

is not of the form $-16\tau^2 \bmod 64$.

For the globalization of Kraus's theorem to number fields see §5.2.2; for the special case $\mathbf{Q}$ see Proposition 5.6.1.

5.2.1 The case $v(p) = 1$

When $v(p) = 1$ the statements of Kraus's theorem simplify considerably, and since this includes the important special cases $K = \mathbf{Q}$ and $\mathbf{Q}_p$, $p = 2, 3$, we put the results in a corollary.

The admissible signatures in the two cases at hand are

$\boxed{p=3}$

$$\begin{array}{rl} 2d, 3d, c, & c, d \in \{0, 1, 2, \ldots\},\ c \geq 6d-3, \\ a, b, 3(a-1), & 1 \leq a < \infty,\ 3a/2 < b \leq \infty, \\ a, b, 2b-3, & 2 \leq b < \infty,\ 2b/3 < a \leq \infty. \end{array}$$

For any triple $c_4, c_6, \Delta \in V$ satisfying (¶), we have either $v(c_4) = 0$, equivalently $v(c_6) = 0$, or $v(c_6) \geq 2$.

$\boxed{p=2}$

$$\begin{array}{rl} 2d, 3d, c, & c, d \in \{0, 1, 2, \ldots\},\ c \geq 6(d-1), \\ a, b, 3(a-2), & 2 \leq a < \infty,\ 3a/2 < b \leq \infty, \\ a, b, 2(b-3), & 3 \leq b < \infty,\ 2b/3 < a \leq \infty. \end{array}$$

For any triple $c_4, c_6, \Delta \in V$ satisfying (¶), we have either $v(c_4) = 0$ or $v(c_4) \geq 2$.

Corollary 5.2.2 *Suppose $v(p) = 1$ (for example when v is the p-adic valuation on $K = \mathbf{Q}$ or $\mathbf{Q}_p$). Then elements c_4, c_6, Δ of V satisfying (¶) are the covariants of some $E_{/V}$*

- *when $p = 3$: iff $v(c_6) \neq 2$;*
- *when $p = 2$: iff either of the special cases* (i) *or* (ii)*, that is, iff either*

  $v(c_4) = 0$ and $c_6 \equiv -x^2 \bmod 4$ has a solution $x \in V$, or

  $v(c_4) \geq 4$ and $c_6 \equiv 8x^2 \bmod 32$ has a solution $x \in V$.

Proof. First let $p = 3$. By the proposition, the covariants occur for some $E_{/V}$ when $v(c_4) = 0$ (equivalently $v(c_6) = 0$) or when $v(c_6) \geq v(27) = 3$. From our list the only undecided admissible signatures are $1, 2, 0$ and $a, 2, 1$, $a \geq 2$. In these cases $\Psi_2(\vartheta) \equiv 0 \bmod 27$ has no solution as we see by writing it as an equation $\vartheta^3 - 3c_4\vartheta - 2c_6 = 27z, v(z) \geq 0$ and applying the principle that the minimum value must be attained at least twice.

The case $p = 2$ is treated in the same way. When $v(c_4) = 2$ or 3, $\Psi_3(\theta^2) \equiv 0 \bmod 256$ is seen to have no solution, and all the remaining admissible cases fall under special cases (i) or (ii) of the proposition. ■

The following proposition clears up a minor point.

Proposition 5.2.3 *When $v(2) = 1$ or $v(3) = 1$ all admissible signatures a, b, c occur as the values of some triple $c_4, c_6, \Delta \in V$ satisfying (¶), with one exception: when the residue field is $\mathbf{F}_2$ (so $p = 2$), the admissible signatures $2d, 3d, 6d-6$, $1 \leq d < \infty$ do not so occur.*

Proof. For convenience we say that the signature a, b, c *occurs via* c_4, c_6 when $a = v(c_4)$, $b = v(c_6)$ and $c = v(\Delta)$ where $\Delta = (c_4^3 - c_6^2)/1728$.

$\boxed{p=3}$

1. Given an admissible triple a, b, c of the type where $c + 3 = \min\{3a, 2b\}$, so that $a \geq 1$ and $b \geq 2$, one may take, for example,

$$c_4 = -4 * 3^a, \quad c_6 = 8 * 3^b, \quad \Delta = -3^{3a-3} - 3^{2b-3}.$$

This is valid also when one of a, b is ∞, provided one interprets 3^∞ as 0.

2. $0, 0, 0$ is the signature of **A11** for which

$$c_4 = 16, \quad c_6 = -8 * 19, \quad \Delta = -11.$$

3. If $0, 0, c$ occurs via c_4, c_6, then a simple calculation that we leave to the reader shows that $0, 0, c + 1$ occurs via $c_4, c_6 + 3\delta$ for appropriate $\delta \in V$.

4. For $d \geq 1$ the signature $2d, 3d, 6d - 3$ occurs: take *e.g.*

$$c_4 = -4 * 3^{2d}, \quad c_6 = 8 * 3^{3d}, \quad \Delta = -2 * 3^{6d-3}.$$

5. For $d \geq 1$ the signature $2d, 3d, 6d - 2$ occurs: take *e.g.*

$$c_4 = 4 * 3^{2d}, \quad c_6 = 16 * 3^{3d}, \quad \Delta = -3^{6d-2}.$$

6. If $2d, 3d, c$ occurs via c_4, c_6 where $d \geq 1$ and $c > 6d - 3$, say $c = 6d - 3 + e$, $e \geq 1$, then simple calculation shows that $2d, 3d, c{+}1$ occurs via $c_4, c_6{+}\delta$ where δ is chosen to satisfy the following condition, where $c_6 = \gamma 3^{3d}$ and $\Delta = \Delta_0 3^c$:

$$\delta \equiv 32\frac{\Delta_0}{\gamma}3^{e-1} + \epsilon 3^e \bmod 3^{e+1}.$$

This concludes the proof that when $v(3) = 1$, all admissible 3-adic signatures occur as the covariant signatures of some $E_{/K}$, and in fact for some $E_{/V}$ except when $v(c_6) = 2$ by the previous corollary.

$\boxed{p = 2}$

1. The admissible signature a, b, c where $3a \neq 2b$ (and therefore $c + 6 = \min\{3a, 2b\}$, $a \geq 2$, $b \geq 3$) occurs:

$$c_4 = 3 * 2^a, \quad c_6 = 9 * 2^b, \quad \Delta = 2^{3a-6} - 3 * 2^{2b-6}.$$

Again one of a, b can be ∞ provided 2^∞ is understood to be 0.

2. The signature $2d, 3d, 6d{-}6$ for $d \geq 1$ does *not* occur when the residue field is $\mathbf{F}_2$ because $c_4 = 2^{2d}\gamma_4$, $c_6 = 2^{3d}\gamma_6 \implies \Delta = 2^{6d-6}(\gamma_4^3 - \gamma_6^2)/27$ and the image of $\gamma_4^3 - \gamma_6^2$ in the residue field is obliged to be $1 - 1 = 0$, hence $v(\gamma_4^3 - \gamma_6^2) \geq 1$. On the other hand, if the residue field properly contains $\mathbf{F}_2$ then we can choose $\gamma_4 = 1$ and take any $\gamma_6 \in V$ whose image in the residue fied is not 0 or 1.

3. The curve **A15**: $y^2 + xy + y = x^3 + x^2$ has $c_4 = 1$, $c_6 = -161$, $\Delta = -15$, hence $0, 0, 0$ occurs.

4. For $d \geq 1$, the signature $2d, 3d, 6d - 5$ occurs: take $c_4 = -17 * 2^{2d}$, $c_6 = 2^{3d}$, $\Delta = -91 * 2^{6d-5}$.

5. If $2d, 3d, 6d-6+e$ occurs, where $d \geq 0$, $e \geq 1$ ($e \geq 6$ when $d = 0$), then $2d, 3d, 6d-5+e$ occurs. For let $c_4 = \gamma_4 2^{2d}$, $c_6 = \gamma_6 2^{3d}$ and $\Delta = \Delta_0 2^{6d-6+e}$, and define $c_4' = c_4 + \delta 2^e$ for any $\delta \in V$ satisfying

$$\delta \equiv -\frac{9\Delta_0}{\gamma_4^2} + 2\epsilon \bmod 2^{e+2}, \text{ where } \epsilon = 0 \text{ if } e = 1,\ \epsilon = 1 \text{ if } e > 1. \qquad \blacksquare$$

5.2.2 The globalization of Kraus's theorem to number fields

Let K be a number field of degree n over $\mathbf{Q}$ with ring of integers $\mathcal{O}$ and with a chosen integral basis $\omega_1, \ldots, \omega_n$. Thus a typical member of $\mathcal{O}$ is

$$\alpha = \sum n_i \omega_i, \quad n_i \in \mathbf{Z}.$$

Let m be a positive integer. Then $\alpha \equiv 0 \bmod m$, *i.e.*, $\alpha = m\beta$ for some $\beta \in \mathcal{O}$, iff $m | n_i\ \forall i$. We define

$$\operatorname{mods}(\alpha, m) = \sum \operatorname{mods}(n_i, m)\, \omega_i$$

(the symmetric residue $\operatorname{mods}(n_i, m)$ was defined in §5.1), and say that α is **restricted mod m** or **m-restricted** if $\operatorname{mods}(\alpha, m) = \alpha$. Clearly

$$\operatorname{mods}(\alpha, m) \equiv \alpha \bmod m.$$

A Weierstrass equation $E_{/\mathcal{O}}$ is of **restricted type** if a_1 and a_3 are 2-restricted and a_2 is 3-restricted. Of course this property depends on the choice of integral basis. In the case $K = \mathbf{Q}$, $\mathcal{O} = \mathbf{Z}$, restricted type means

$$a_1, a_3 \in \{0, 1\} \quad \text{and} \quad a_2 \in \{-1, 0, 1\}.$$

Proposition 5.2.4 *Given $E_{/\mathcal{O}}$ there exists a unique $\mathcal{O}$-isomorphism of the form $\tau = [r, s, t, 1]$ such that $E' = \tau E$ is of restricted type; E' and τ can be determined by the following sequence of calculations.*

$$\begin{aligned}
a_1' &= \operatorname{mods}(a_1, 2),\\
s &= (a_1' - a_1)/2,\\
a_2' &= \operatorname{mods}(a_2 - sa_1 - s^2, 3),\\
r &= (a_2' - a_2 + sa_1 + s^2)/3,\\
a_3' &= \operatorname{mods}(a_3 + ra_1', 2),\\
t &= rs + (a_3' - a_3 - ra_1')/2,\\
a_4' &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st,\\
a_6' &= a_6 + ra_4 + r^2 a_2 + r^3 - ta_3 - t^2 - rta_1.
\end{aligned}$$

Proof. The transformation rules of Proposition 4.1.1 show that the Weierstrass coefficients of τE are a'_i as stated. By definition of a'_1, a'_2, a'_3, E' is of restricted type. Since τ is restricted to have 4-th term $u = 1$, it is unique. ■

Note that E' in general is not unique when we relax the definition of $\tau = [r, s, t, u]$ to allow $u \in \mathcal{O}^*$. For example when $\mathcal{O} = \mathbf{Z}[i]$, so $\mathcal{O}^* = \{\pm 1, \pm i\}$, with integral basis $\{1, i\}$, then $E : y^2 = x^3 + 1$ and $[0, 0, 0, i]E : y^2 = x^3 - 1$ are $\mathbf{Z}[i]$-isomorphic and are both of restricted type. However these two curves have different c_6. Uniqueness of the restricted model is restored when the covariants are prescribed, as will be explained in the next proposition.

Lemma 5.2.5 *Let $b \in \mathcal{O}$ and suppose $x^2 \equiv b \bmod 4$ has a solution $x = a \in \mathcal{O}$. Then the congruence has a unique 2-restricted solution in $\mathcal{O}$, namely $x = \operatorname{mods}(a, 2)$.*

Proof. Set $\overline{a} = \operatorname{mods}(a, 2)$. Then $\overline{a} = a + 2c$ for some $c \in \mathcal{O}$, so $\overline{a}^2 \equiv a^2 \equiv b \bmod 4$. If a' is another solution of the congruence and we set $a' = \overline{a} + d$, then $2\overline{a}d + d^2 \equiv 0 \bmod 4$. Hence if $P|2$, *i.e.,* if P is any prime ideal divisor of $2\mathcal{O}$, then $v_P(d(2\overline{a} + d)) \geq 2v_P(2)$ which implies $v_P(d) \geq v_P(2)$. Thus $d/2 \in \mathcal{O}$ and therefore if a' is also restricted then $d = 0$. ■

The next proposition is the globalization of Kraus's theorem to number fields.

Proposition 5.2.6 (a) *Let c_4, c_6 be elements of $\mathcal{O}$ such that $\Delta = (c_4^3 - c_6^2)/1728$ is a nonzero element of $\mathcal{O}$. Then $\exists$ Weierstrass model $E_{/\mathcal{O}}$ with these covariants iff $\exists\, \vartheta, \theta, \tau \in \mathcal{O}$ satisfying*

$$\Psi_2(\vartheta) \equiv 0 \bmod 27, \quad \Psi_3(\theta^2) \equiv 0 \bmod 256, \quad \Psi_2(\theta^2) \equiv -16\tau^2 \bmod 64.$$

When ϑ exists it is unique mod 3*: $\vartheta + 3\alpha$ is also a solution $\forall \alpha \in \mathcal{O}$ and all solutions are obtained this way. When θ, τ exist they are unique* mod 2 *in the following sense: the other solutions are $\theta + 2\beta, \tau + \theta\beta(\theta + \beta) + 2\gamma$ where $\beta, \gamma \in \mathcal{O}$.*

(b) *When ϑ, θ and τ satisfying the three congruences exist, there exists a unique model $E_{/\mathcal{O}}$ of restricted type with the given covariants (that is, unique once an integral basis $\omega_1, \ldots, \omega_n$ has been chosen). It can be calculated from ϑ and θ by the following steps.*

1. *Let $a_1 = \operatorname{mods}(\theta, 2)$, and then let $a_2 = \operatorname{mods}(\vartheta - a_1^2, 3)$.*
2. *Define $b_2 = a_1^2 + 4a_2$, then $b_6 = \Psi_2(b_2)/432$.*
3. *Then $x^2 \equiv b_6 \bmod 4$ has a solution $x \in \mathcal{O}$. Let $a_3 = \operatorname{mods}(x, 2)$.*
4. *Let $a_4 = (b_2^2 - c_4 - 24a_1a_3)/48$ and $a_6 = (b_6 - a_3^2)/4$.*

Proof. (a) Let P denote a typical prime ideal divisor of $3\mathcal{O}$. If $E_{/\mathcal{O}}$ exists then it is defined over $\mathcal{O}_P \supset \mathcal{O}$, hence by the local result there exists $\vartheta_P \in \mathcal{O}_P$ satisfying $\Psi_2(\vartheta_P) \equiv 0 \bmod 27\mathcal{O}_P$. By the approximation theorem for Dedekind domains we choose $\vartheta \in \mathcal{O}$ such that

$$\vartheta \equiv \vartheta_P \bmod 3\mathcal{O}_P, \quad \forall P|3,$$

where the notation $\forall P|3$ is short for "every prime divisor P of 3". Again by the local result, $\Psi_2(\vartheta) \equiv 0 \bmod 27\mathcal{O}_P$. Hence if we write $\Psi_2(\vartheta) = 27\lambda$, then $v_P(\lambda) \geq 0$ $\forall P|3$, and also $\forall P \nmid 3$ since $\Psi_2(\vartheta) \in \mathcal{O}$. Therefore $\lambda \in \mathcal{O}$ and so $\Psi_2(\vartheta) \equiv 0 \bmod 27\mathcal{O}$.

Similarly by the approximation theorem we choose $\theta, \tau \in \mathcal{O}$ so that

$$\theta \equiv \theta_P \bmod 4\mathcal{O}_P \quad \forall P|2,$$

$$\tau \equiv \tau_P \bmod 2\mathcal{O}_P \quad \forall P|2.$$

(If one makes the coarser approximation

$$\theta \equiv \theta_P \bmod 2\mathcal{O}_P \quad \forall P|2, \quad \text{say} \quad \theta = \theta_P + 2\alpha_P,$$

then one has the more complicated looking

$$\tau \equiv \tau_P + \alpha_P\tau_P(\alpha_P + \tau_P) \bmod 2\mathcal{O}_P.)$$

Conversely if ϑ, θ, τ exist then the local criteria are satisfied for the finite set S of prime ideal divisors P of 2 and 3. Hence for each $P \in S$ $\exists$ a transformation $[r_P, s_P, t_P, 1]$ defined over K from

$$E: \ y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864}$$

to an $\mathcal{O}_P$-integral curve with the given covariants. By the approximation theorem, we pick $r, s, t \in K$ P-adically integral for all $P \notin S$, and sufficiently P-adically close to r_P, s_P, t_P respectively so that the transformed curve is still $\mathcal{O}_P$-integral $\forall P \in S$. Then $[r, s, t, 1]$ applied to E gives an $\mathcal{O}$-integral model with the given covariants.

If ϑ and $\vartheta + 3\alpha$ are two solutions in $\mathcal{O}$, we wish to prove that $\alpha \in \mathcal{O}$. The local results imply that $v_P(\alpha) \geq 0$ $\forall P|3$, and $v_P(\alpha) = v_P(3\alpha) \geq 0$ $\forall P \nmid 3$, hence the result. Similarly the statements concerning the uniqueness of θ and τ follow from details given in Proposition 5.2.1.

(b) Suppose ϑ, θ and τ exist. Then by part (a) an $E_{\mathcal{O}}$ exists with the given covariants, and by the previous proposition we can assume that E is of restricted type. Let $a_1 = \theta + 2d$; we wish to prove that $d \in \mathcal{O}$. By the local Kraus theorem, since one solution of the Ψ_3 congruence is a_1 and other solutions differ by a multiple of 2, we have $v_P(d) \geq 0$ $\forall P|2$, hence $d \in \mathcal{O}$ as required. Since a_1 is 2-restricted it follows that $a_1 = \text{mods}\,(\theta, 2)$. By part (a), $\theta \in \mathcal{O}$ is unique mod $2\mathcal{O}$ hence a_1 is unique.

Similarly the local result for primes over 3 states that one solution of the Ψ_2 congruence is $b_2 \equiv a_1^2 + a_2 \bmod 3$, and as in the case a_1, we find that a_2 restricted mod 3 is unique and is given by $a_2 = \text{mods}\,(\vartheta - a_1^2, 3)$.

To complete the proof of uniqueness of E it remains to prove that a_3 restricted mod 2 is unique, for then a_4 and a_6 are determined by the formulas defining c_4 and c_6. If there is a prime P over 2 whose residue field is larger than $\mathbf{F}_2$ then it is incorrect to argue as follows.

> One solution is $\tau = a_3 + a_1a_2$ and since "solutions are unique mod 2" therefore $a_3 = \text{mods}\,(\tau + a_1a_2, 2)$; and since a_3 is 2-restricted it is unique.

The trouble is the general solution has the form $\tau = a_3 + a_1a_2 + \beta\tau(\beta+\tau) + 2\gamma$ and the term $\beta\tau(\beta+\tau)$ can be nonzero mod 2. In other words, one may start with the 'wrong' τ.

However we can prove uniqueness as follows. Since a_1 and a_2 are unique at this point, so are $b_2 = a_1^2 + 4a_2$ and $b_6 = \Psi_2(b_2)/432$. Thus $a_3^2 = b_6 - 4a_6$ is unique mod 4 and therefore by the Lemma, a_3 is unique mod 2, hence unique since it is 2-restricted. ■

5.3 Local minimal models and δ_v

We return to the local situation of a discrete valuation ring V with quotient field K of any characteristic.

E is **minimal** (with respect to the valuation v) if E is defined over V and the nonnegative integer $v(\Delta)$ is minimal among all K-isomorphic E' defined over V. Then E is called a **minimal Weierstrass model** (for v), the name Weierstrass being inserted when necessary to distinguish it from the Néron minimal model which is a much more sophisticated construction to be discussed later.

Lemma 5.3.1 *If E and E' are defined over V and are K-isomorphic with $v(\Delta) \geq v(\Delta')$,* e.g. *$E'$ minimal, then any K-isomorphism $\tau = [r, s, t, u]$ from E to E' is V-integral.*

Proof. From $v(\Delta') = 12v(u) + v(\Delta) < \infty$ and $v(\Delta') \leq v(\Delta)$ we deduce $v(u) \geq 0$, *i.e.,* $u \in V$. The relations from § 3.1

$$\begin{aligned} u^8 b_8' &= b_8 + 3rb_6 + 3r^2b_4 + r^3b_2 + 3r^4 \\ u^6 b_6' &= b_6 + 2rb_4 + r^2b_2 + 4r^3 \end{aligned}$$

and the fact that V is integrally closed imply in succession that $3r, 4r \in V$ hence $r \in V$. Next

$$\begin{aligned} u^2 a_2' &= a_2 - sa_1 + 3r - s^2 \\ u^6 a_6' &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1 \end{aligned}$$

show that $s, t \in V$. ■

Proposition 5.3.2 *Every elliptic curve defined over K has a minimal Weierstrass model for v. It is unique up to V-isomorphism. If* char $\tilde{v} \neq 2$ *then it exists in b-form*

$$y^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4};$$

if char $\tilde{v} \neq 2$ *or* 3 *then it exists in c-form*

$$y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864}.$$

Proof. If $n = \min\{v(a_i)\}$ then a transformation $[0, 0, 0, \pi^n]$, where π denotes a uniformizer for v, clears denominators so there is a model defined over V. The existence and uniqueness statements now follow from the lemma. When char $\tilde{v} \neq 2$ (and $\neq 3$) then 2 (and 3) are invertible in V, so completing the square (and cube) as in the proof of Lemma 4.2.1 can be done over V. ■

Corollary 5.3.3 *The signature $v(c_4), v(c_6), v(\Delta)$ is the same for all v-minimal models of an elliptic curve defined over K.*

For another signature has the form $v(u^{-4}c_4), v(u^{-6}c_6), v(u^{-12}\Delta)$ and $v(u) = 0$ by the proposition. ■

We fix the notation

$$\delta_v = \delta_v(E) = v(\Delta)$$

where Δ is the discriminant of any v-minimal model of E.

We also use the notation δ_π for δ_v where π is a uniformizer, for example δ_2 when $K = \mathbf{Q}$, when that is more convenient. We do not introduce special notation for the other two members of the signature of a v-minimal model.

The transformation equation

$$u^{12}\Delta' = \Delta$$

shows that for any model

$$v(\Delta) \equiv \delta_v \bmod 12$$

and if $0 \leq v(\Delta) < 12$ then E is minimal. For example over $K = \mathbf{Q}$

$$y^2 = x^3 + 16, \quad \text{with} \quad c_4 = 0,\ c_6 = -2^9 3^3,\ \Delta = -2^{12}3^3$$

is transformed by $\tau = [0, 0, 4, 2]$ to

$$y^2 + y = x^3, \quad \text{with} \quad c_4 = 0,\ c_6 = -2^3 3^3,\ \Delta = -3^3 \qquad (\mathbf{A27})$$

which is a 2-minimal model, *i.e.,* minimal for the 2-adic valuation. In fact **A27** is **Z**-*minimal*, that is, simultaneously minimal for all the p-adic valuations, a matter that we will return to shortly. In this case there is no 2-minimal model in b-form, *a fortiori* none in c-form.

The converse criterion is not true: a minimal model may have $v(\Delta) \geq 12$, and in fact it is easy to produce examples with arbitrarily large $v(\Delta)$. For the transformation equation

$$u^4 c_4' = c_4$$

shows that E is minimal when $v(c_4) < 4$. Thus

$$y^2 + xy + y = x^3 - 454x - 544 \qquad (\mathbf{H30})$$

$$\text{with} \quad c_4 = 21769,\ c_6 = 437147,\ \Delta = 2^3 * 3 * 5^{12}$$

is **Z**-minimal. In general one can take for instance (say $\operatorname{char} \tilde{v} \neq 2, 3$ for simplicity)

$$y^2 = x^3 - 3(1 + \pi^N)x - 2(1 + \pi^N),$$

where $v(\pi) = 1$, which has $v(c_4) = 0$, hence is v-minimal, and

$$v(\Delta) = v(2^6 3^3 (1 + \pi^N)^2 \pi^N) = N.$$

Proposition 5.3.4 *Let V be a valuation ring with quotient field K, valuation map v, uniformizer π, and residue field k. Let E be an elliptic curve defined over V with signature*

$$a = v(c_4), \quad b = v(c_6), \quad c = v(\Delta).$$

(a) *E is v-minimal iff the following condition, labelled (α), is not true.*

(α) *E is K-isomorphic to an E' defined over V with signature $a - 4$, $b - 6$, $c - 12$ (which implies $a \geq 4$, $b \geq 6$, $c \geq 12$).*

Consequently, if any of $a < 4$, $b < 6$ or $c < 12$ is true, then E is v-minimal.

(b) *Let $\operatorname{char} K \neq 2$ or 3. Then condition (α) is equivalent to*

(β) *$c_4' = c_4 \pi^{-4}$, $c_6' = c_6 \pi^{-6}$ occur as the covariants of some E' defined over V.*

That is, E is v-minimal iff (β) is not true, and then

$$v(\Delta) < \max\{12 + 12v(2) + 3v(3) - v(j)\,,\ 12 + 10v(2) + 6v(3) - v(j - 12^3)\}.$$

(c) *If $\operatorname{char} k \neq 2$ or 3 (hence $\operatorname{char} K \neq 2$ or 3 and $v(2) = v(3) = 0$), then E is minimal iff*

$$\textit{either} \quad v(\Delta) < 12 \quad \textit{or} \quad v(c_4) < 4$$

(equivalently, either $v(c_4) < 4$ or $v(c_6) < 6$). Therefore if E is any model then

$$\delta_v = v(\Delta) - 12 \min\{\lfloor v(c_4)/4 \rfloor, \lfloor v(\Delta)/12 \rfloor\}.$$

Proof. (a) If such an E' exists then it is V-isomorphic to E by Lemma 5.3.1, and so E is not v-minimal. The converse follows from the trivial technicality: if E has v-minimal model E'' with $c_4'' = c_4\pi^{-4N}$, $c_6'' = c_6\pi^{-6N}$, then $E' = [0,0,0,\pi^{-(N-1)}]E''$ is defined over V and has covariants $c_4\pi^{-4}$, *etc.*, hence signature $a-4, b-6, c-12$.

(b) If E' satisfies condition (α) then, for an appropriate unit $u \in V^*$, $[0,0,0,u]E'$ serves as the E' in (β). Conversely if E' satisfies (β), then E and E' have the same j-invariant, hence are K-isomorphic by Proposition 4.4.1, (i)–(iii). Therefore E' satisfies (α).

If we suppose the inequality not satisfied then we have both

$$v(j\Delta) = v(c_4^3) \geq 12 + 12v(2) + 3v(3),$$

and

$$v((j-1728)\Delta) = v(c_6^2) \geq 12 + 10v(2) + 6v(3).$$

Then
$$y^2 = x^3 - \frac{c_4}{2^4 3\pi^4}x - \frac{c_6}{2^5 3^3 \pi^6}$$

is V-isomorphic to E and has smaller $v(\Delta)$.

(c) When $\operatorname{char} k \neq 2$ or 3, any two $c_4, c_6 \in V$ such that $c_4^3 \neq c_6^2$ are the covariants of an E' defined over V, *e.g.* $y^2 = x^3 - c_4x/48 - c_6/864$. Thus (b) follows from (c). ■

Here are two $\mathbf{Z}$-minimal examples illustrating part (b) of the proposition.

$$y^2 = x^3 + 4x, \;\; c_4 = -2^6 3, \; \Delta = -2^{12}, \; j = 1728; \qquad (\mathbf{B32})$$

$$y^2 + xy + y = x^3 - x^2 + 25x + 1 \qquad (\mathbf{H162})$$

$$c_4 = -3^5 5, \; \Delta = -2 * 3^{12}, \; j = 3^3 5^3/2, \; j - 1728 = -3^4/2.$$

5.3.1 The special cases $v(3) = 1$ and $v(2) = 1$

Under either assumption $v(3) = 1$ or $v(2) = 1$, the field K is of characteristic 0; of particular interest are $K = \mathbf{Q}, \mathbf{Q}_2$ and $\mathbf{Q}_3$.

Proposition 5.3.5 *Supposing $v(p) = 1$, where $p = 2$ or 3, the complete list of possible signatures of v-minimal $E_{/V}$ is as follows. Conversely any $E_{/V}$ with one of these signatures is v-minimal, and such $E_{/V}$ actually occur for each of the signatures, unless noted otherwise. (The footnotes refer to certain 2-adic signatures only.)*

$p=3$		$p=2$		
$0,0,c$	$c \geq 0$	$0,0,c$	$c \geq 0$	†
$2,3,c$	$c \geq 3$	$4,6,c$	$c \geq 6$	§‡
$4,6,c$	$c = 9,10,11$	$6,9,c$	$c \geq 12$	§
$1,b,0$	$3 \leq b \leq \infty$	$4,b,6$	$7 \leq b \leq \infty$	
$2,b,3$	$4 \leq b \leq \infty$	$5,b,9$	$8 \leq b \leq \infty$	
$3,b,6$	$5 \leq b \leq \infty$	$6,b,12$	$10 \leq b \leq \infty$	
$4,b,9$	$7 \leq b \leq \infty$	$7,b,15$	$11 \leq b \leq \infty$	
$5,8,12$		$a,3,0$	$4 \leq a \leq \infty$	†
$a,3,3$	$3 \leq a \leq \infty$	$a,5,4$	$4 \leq a \leq \infty$	
$a,4,5$	$3 \leq a \leq \infty$	$a,6,6$	$5 \leq a \leq \infty$	
$a,5,7$	$4 \leq a \leq \infty$	$a,7,8$	$5 \leq a \leq \infty$	
$a,6,9$	$5 \leq a \leq \infty$	$a,8,10$	$6 \leq a \leq \infty$	
$a,7,11$	$5 \leq a \leq \infty$	$a,9,12$	$7 \leq a \leq \infty$	¶
$a,8,13$	$6 \leq a \leq \infty$	$a,10,14$	$7 \leq a \leq \infty$	

†*A covariant triple* c_4, c_6, Δ *of some* $E_{/K}$ *with the given signature might not pass the criterion of Proposition* 5.2.1 (*special case* (i) *or* (ii)), *and such a triple will not occur as the covariants of any* $E_{/V}$*; however there do exist (other)* $E_{/V}$ *with this signature, and any such* $E_{/V}$ *is* 2*-minimal.*
§*The lowest value of* c *is not included when the residue field is* $\mathbf{F}_2$.
‡*An* $E_{/V}$ *with this signature is not necessarily* 2*-minimal when* $c \geq 12$.
¶*An* $E_{/V}$ *with this signature is not necessarily* 2*-minimal when* $a \geq 8$.

Proof. First we verify that there exist $E_{/V}$ with the stated signatures. Since the signatures satisfy $v(c_6) \neq 2$ when $p = 3$ and $v(c_4)$ is 0 or ≥ 4 when $p = 2$, as they must by Corollary 5.2.2, much of this follows from Proposition 5.2.3. The only cases that require additional comment are the two tagged with †. For the case of 2-adic signature $0, 0, c$, the curve **A15** serves as an example of $E_{/V}$ when $c = 0$ — see step 3 in the proof of Proposition 5.2.3, case $p = 2$ — and in the induction to $0, 0, c+1$ as explained there in step 5, only c_4 is altered: c_6 keeps the value -161, hence the congruence $c_6 \equiv -x^2 \bmod 4$ has a solution, as required by Proposition 5.2.1, special case (i). The other case is $a, 3, 0$ where $4 \leq a \leq \infty$. We use the construction in step 1 of the proof referred to above. In this case it gives $c_4 = 3 * 2^a$ ($= 0$ if $a = \infty$), $c_6 = 9 * 8$, hence the requirements of special case (ii) are met.

Next we prove the minimality of an $E_{/V}$ with one of these signatures, apart from the exceptions allowed by the tags ‡ and ¶. Let c_4, c_6, Δ be the covariants of E. We wish to prove that $2^{-4}c_4$, $2^{-6}c_6$, $2^{-12}\Delta$ are not the covariants of some $E'_{/V}$. For many cases this is obvious from either $a < 4$, $b < 6$ or $c < 12$. The explanation for the 3-adic signatures $5, 8, 12$ and $6, 8, 13$ is that the reduced signature would have $v(c_6) = 2$, which is disallowed by Corollary 5.2.2. Similarly,

2-adic signatures with $a = 5, 6$ or 7 are minimal since the reduced signature would have $v(c_4) = 1, 2$ or 3, which are disallowed. Also special case (ii) of Proposition 5.2.1 disallows 2-adic signatures of the form $a-4, 4, 2$, $(7 \le a \le \infty)$, hence all $a, 10, 14$, $(7 \le a \le \infty)$, are 2-minimal. However special case (ii) does permit some $a - 4, 3, 0$, and therefore $a, 9, 12$ may or may not be 2-minimal, depending on the particular c_6 that is involved. Similarly, $4, 6, c$ with $c \ge 12$ may or may not be 2-minimal, depending on c_6.

Finally we check that no 2 or 3-minimal signature is missing from the lists. This amounts to verifying that if an admissible signature a, b, c (as listed in § 5.2.1) satisfies the requirements of Corollary 5.2.2 and is not in the present list, then $a-4, b-6, c-12$ is in the present list and appears without the tag †. This is straightforward. ■

Here are a few examples of the earliest appearance of an admissible signature in the catalogs [AntIV] and [Cre92]. This requires calculation since the catalogs do not contain the values of c_4, c_6; in some cases one has to go quite far in the catalog.

3-adic signatures: $0, 0, 0$: A11; $0, 0, 1$: A15; $2, 3, 4$: E162; $1, \infty, 0$: A32; $2, 4, 3$: $\mathrm{A}_1 459$; $\infty, 3, 3$: A27.

2-adic signatures: $0, 0, 0$: A15; $0, 0, 1$: B14; $0, 0, 39$: $\mathrm{D}_2 858$; $4, 6, 6$: does not occur; $4, \infty, 6$: A32; $\infty, 6, 6$: 2∗A27 (beyond the catalog listing).

5.4 Unramified base change

For a given field K with valuation v a **base change** or **extension** is a pair K', v' consisting of an extension field K' of K and a valuation v' extending v: for $x \in K$

$$v'(x) = e(v', v)v(x).$$

Typically π, π' denote uniformizers and k, k' denote residue fields of v, v' respectively.

The base change is **unramified** if

(i) the residue field extension k'/k is separable and

(ii) $e(v', v) = 1$; this allows us to choose $\pi' = \pi$ and often we will simplify the notation to the mildly inaccurate $v(x)$ $\forall x \in K'$.

Examples of unramified base change are: $K' = \widehat{K}_{\mathrm{nr}}$, the maximal unramified extension of the completion of K; or, more mundanely, a finite extension where v' is an unramified extension of v — of course some of the other extensions of v to K' may be ramified.

Let S be a set of representatives of k in V so that the elements of K are uniquely represented as sums $\sum s_i \pi^i$, $s_i \in S$. We can choose a set $S' \subset V'$ of representatives of k' that contains S; when $e(v', v) = 1$, an element $\sum s'_i \pi^i$ of K' is in K iff all $s'_i \in S$.

Proposition 5.4.1 *Let K', v' be an unramified extension of K, v and let E be an elliptic curve defined over V. Then E is v'-minimal iff it is v-minimal.*

Proof. Suppose $\tau = [r, s, t, u]$ is defined over V' and transforms E to a curve E' defined over V' with smaller $v(\Delta)$, *i.e.*, $v(u) > 0$; let us call such a τ *reducing*. We must prove that there is a reducing transformation defined over V. We can assume that $u = \pi$: with the given r, s, t, clearly $[r, s, t, \pi]$ is reducing. Let $\operatorname{char} k = p \geq 0$, so $\operatorname{char} K = 0$ or p.

If $p \neq 2, 3$ the result is trivial: since 2 and 3 are invertible in V, E is V-isomorphic to its c-form

$$y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864},$$

and similarly E' is V'-isomorphic to

$$y^2 = x^3 - \frac{c_4\pi^{-4}}{48}x - \frac{c_6\pi^{-6}}{864}.$$

With the Weierstrass equations in this form the transformation is simply $[0, 0, 0, \pi]$ and the reducing property amounts to the two conditions $v(c_4) \geq 4$, $v(c_6) \geq 6$.

Secondly suppose $p = 3$. Since 2 is invertible in V we can take the equation of E in b-form

$$y^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4},$$

and similarly for E'. The transformation equation for a_1', namely $\pi a_1' = a_1 + 2s$ with $a_1' = a_1 = 0$, implies $s = 0$, and similarly the equation for a_3' implies $t = 0$. Thus $\tau = [r, 0, 0, \pi]$. Next, by analyzing

$$\begin{aligned}
\pi^2 b_2' &= b_2 + 12r, \\
\pi^4 b_4' &= b_4 + rb_2 + 6r^2, \\
\pi^6 b_6' &= b_6 + 2rb_4 + r^2 b_2 + 4r^3,
\end{aligned}$$

we prove in successsion

(a) for all $\rho \in V'$, $[r + \rho\pi^2, 0, 0, \pi]$ is reducing, hence we can assume $r = r_0 + r_1\pi$, $r_i \in S'$;

(b) $r_0 \in S$ and $r_1 \in S$ so $r \in V$ and τ is defined over V, hence the result.

Proof of (a). Let us denote the right sides of the above transformation equations by g_6, g_4, g_2, regarded as functions of r. We must show that $g_i(r + \rho\pi^2) \equiv 0 \bmod \pi^i$ for $i = 2, 4, 6$ so that the new b_2', b_4', b_6' determined by this modified r are in V'.

$$\begin{aligned}
g_2(r + \rho\pi^2) &= g_2(r) + 12\rho\pi^2 \\
&= (b_2' + 12\rho)\pi^2 \equiv 0 \bmod \pi^2; \\
g_4(r + \rho\pi^2) &= g_4(r) + \rho\pi^2(b_2 + 12r) + 6\rho^2\pi^4 \\
&= (b_4' + b_2'\rho + 6\rho^2)\pi^4 \equiv 0 \bmod \pi^4; \\
g_6(r + \rho\pi^2) &= (b_6' + 2b_4'\rho + b_2'\rho^2 + 4\rho^3)\pi^6 \equiv 0 \bmod \pi^6.
\end{aligned}$$

Proof of (b). Write $3 = \phi\pi$ where $v(\phi) \geq 0$. (When $\operatorname{char} K = 3$, $\phi = 0$ and $v(\phi) = \infty$.) The equation for b_2 shows that $b_2 = \beta_2\pi$, $\beta_2 \in V$, and now the equation for b_4 shows that $b_4 = \beta_4\pi$, $\beta_4 \in V$. Using these facts in $g_6(r) \equiv 0 \bmod \pi^6$, we deduce

$$g_6(r) \equiv 4r_0^3 + b_6 \equiv 0 \bmod \pi.$$

At the residue field level this is a purely inseparable equation for $\widetilde{r}_0$, hence $\widetilde{r}_0 \in k$. This implies $r_0 \in S$.

Writing

$$b_6 + 2r_0b_4 + r_0^2b_2 + 4r_0^3 = \alpha \in V,$$

we find

$$g_6(r) = \alpha + 2\pi r_1(\pi^4 b_4') - \pi^3(8r_1^3 + r_1^2(\phi r_0 + \beta_2)) \equiv 0 \bmod \pi^6$$

hence $\alpha = \beta\pi^3$ where $\beta \in V$ and

$$8r_1^3 + r_1^2(\phi r_0 + \beta_2) \equiv -\beta \bmod \pi.$$

But

$$\phi r_0 + \beta_2 \equiv \pi b_2' \equiv 0 \bmod \pi,$$

hence the purely inseparable equation $r_1^3 \equiv \beta \bmod \pi$. As with r_0, this implies $r_1 \in V$.

The case $p = 2$ proceeds similarly except it is more intricate. This time we use the transformation equations for the a_i. Since 3 is invertible in V we can assume $a_2 = 0$, and as explained above, we are dealing with a reducing transformation $[r, s, t, \pi]$ defined over V'. This time there are four steps to obtain a reducing transformation defined over V:

(a) for any $\rho, \sigma, \tau \in V'$,

$$[r + \rho\pi^3, s + \sigma\pi, t + \tau\pi^3, \pi]$$

is also reducing, hence we can assume

$$r = r_0 + r_1\pi + r_2\pi^2, \quad s = s_0, \quad t = t_0 + t_1\pi + t_2\pi^2;$$

(b) $r_0, r_1, s_0, t_0, t_1 \in S$;

(c) λ can be chosen in V' so that $r_2 + \lambda\pi^2$ and $t_2 - s_0\lambda\pi^2$ are in S;

(d) $\forall \lambda \in V'$, $[r + \lambda\pi^2, s_0, t - s_0\lambda\pi^2, \pi]$ is reducing.

The transformation equations are

$$\begin{aligned}
\pi a_1' &= a_1 + 2s \\
\pi^2 a_2' &= -sa_1 + 3r - s^2 \qquad (a_2 = 0) \\
\pi^3 a_3' &= a_3 + ra_1 + 2t \\
\pi^4 a_4' &= a_4 - sa_3 - (t + rs)a_1 + 3r^2 - 2st \\
\pi^6 a_6' &= a_6 + ra_4 + r^3 - ta_3 - t^2 - rta_1
\end{aligned}$$

Let $f_1, \dots f_6$ denote the right sides of these equations. The first three equations imply

$$a_1 = \alpha_1\pi,\ a_3 = \alpha_3\pi, \quad \text{where} \quad \alpha_i \in V, \quad r \equiv s^2 \bmod \pi.$$

We write $2 = \phi\pi$ where $v(\phi) \geq 0$.

Proof of (a). First we change s: we wish to prove that $f_i(s+\sigma\pi) \equiv f_i(s) \bmod \pi^i$ for $i = 1, \dots, 6$; we have temporarily suppressed r and t from the notation — we have simplified $f_i(r, s, t)$ to $f_i(s)$. These are easy and we only indicate the calculation for one:

$$\begin{aligned} f_4(s+\sigma\pi) &= f_4(s) - \sigma\pi^2(\alpha_3 + r\alpha_1 + \phi t) \\ &= f_4(s) - \sigma\pi^2 a_3'\pi^2 \equiv f_4(s) \bmod \pi^4. \end{aligned}$$

Next we change r; to shorten the exposition we only work through the most complicated of the five cases:

$$\begin{aligned} f_6(r+\rho\pi^3) &= f_6(r) + \rho\pi^3 a_4 + 3r^2\rho\pi^3 + 3r\rho^2\pi^6 + \rho^3\pi^9 - \rho\pi^3 t a_1 \\ &\equiv f_6(r) + \rho\pi^3(a_4 + 3r^2 - ta_1) \\ &\equiv f_6 + \rho\pi^3(\pi^4 a_4') + \rho\pi^3 s(a_3 + ra_1 + 2t) \\ &\equiv f_6(r) + \rho\pi^3 s(\pi^3 a_3') \equiv f_6(r) \bmod \pi^6. \end{aligned}$$

Similarly it is easily verified that $f_i(t + \tau\pi^3) \equiv f_i(t) \bmod \pi^i$.

Proof of (b). From the a_4 equation, remembering that a_1, a_3 and 2 are $\equiv 0 \bmod \pi$, we get $3r^2 + a_4 \equiv 0 \bmod \pi$, hence $r_0^2 + a_4 \equiv 0 \bmod \pi$. Since this is a purely inseparable equation over the residue field, therefore $r_0 \in S$.

Next, from the a_6 equation, $-t^2 + r^3 + ra_4 + a_6 \equiv 0 \bmod \pi$ hence $t_0^2 + r_0^3 + r_0 a_4 + a_6 \equiv 0$, so we have a purely inseparable equation for t_0 at the residue field level and therefore $t_0 \in S$. Similarly from f_2 we find $s_0^2 + r_0 \equiv 0 \bmod \pi$ hence $s_0 \in S$. Let $3r_0 - s_0^2 = \pi\lambda,\ \lambda \in V$.

Again f_2 implies

$$-s_0\pi\alpha_1 + 3r_1\pi + \lambda\pi \equiv 0 \bmod \pi^2$$

hence $r_1 \equiv$ an element of k, therefore $r_1 \in S$.

If we write $q = r_0 + r_1\pi$ then taking all the π^2 terms in f_6 yields

$$r_2(a_4 + 3q^2) - t_1(\alpha_3 + \phi t_0 + q\alpha_1) - t_1^2 \equiv \gamma \bmod \pi^3$$

for some $\gamma \in V$. But $a_4 + 3q^2 \equiv 0 \bmod \pi$ by the a_4 equation, and the coefficient of t_1 is also 0 modπ by the a_3 equation. Thus by the usual argument, $t_1 \in S$.

Proof of (c). At this point the a_6 equation gives for some $\delta \in V$ (the congruences are modπ^6)

$$\begin{aligned} \delta &\equiv r_2\pi^2(a_4'\pi^4 + sa_3 + rsa_1 + 2st - 3rr_2\pi^2) \\ &\qquad -t_2\pi^3(\pi^2 a_3' - r_2\pi a_1) - t_2^2\pi^4 \\ &\equiv r_2\pi^2 s(a_3 + ra_1 + 2t) - 3rr_2^2\pi^4 \\ &\qquad -t_2\pi^5(a_3' - r_2\alpha_1) - t_2^2\pi^4 \\ &\equiv r_2\pi^2 s(\pi^3 a_3') + \&\text{c}. \end{aligned}$$

We deduce $rr_2^2 + t_2^2 \equiv \epsilon \bmod \pi$ for some $\epsilon \in V$. Using $r \equiv s_0^2$ as noted earlier and the purely inseparable form of $(s_0 r_2 + t_2)^2 \equiv \epsilon \bmod \pi$, we find $s_0 r_2 + t_2 \equiv \zeta \bmod \pi$, some $\zeta \in V$.

Proof of (d). It remains to check that $f_i(r+\lambda\pi^2, t-s_0\lambda\pi^2) \equiv f_i(r,t) \bmod \pi^i$. Only the cases $i = 4$ and 6 require any work; we set out the former case:

$$\begin{aligned} f_4(r',t') &= f_4(r,t) + 3(2r\lambda\pi^2 + \lambda^2\pi^4) + 2s_0^2\lambda\pi^2 \\ &\equiv f_4(r,t) + 2\pi^2\lambda(r_0 + s_0^2) \\ &\equiv f_4(r,t) \bmod \pi^4 \\ &\quad \text{since } r_0 + s_0^2 \equiv 2s_0^2 \bmod \pi. \qquad \blacksquare \end{aligned}$$

Corollary 5.4.2 *Let v be a valuation on K with ring V.*

(a) *Let E be an elliptic curve defined over K and let K', v' be an unramified extension. Then*

$$\delta_v(E) = \delta_{v'}(E).$$

(b) *Let $u \in K^*$ and suppose that v is unramified in the extension $K(\sqrt{u})/K$. Let E be defined over K, and let E^u denote the twist of E by u. Then*[†]

$$\delta_v(E) = \delta_v(E^u).$$

(c) *Let* $\mathrm{char} K \neq 2$ *or* 3. *The status of a pair of elements $c_4, c_6 \in K$, in the sense of Kraus, is unchanged in an unramified extension K', v': the pair occurs as the covariants of a Weierstrass equation defined over V iff it does so over V'.*

Proof. (a) We can take E as a v-minimal model, so $\delta_v = v(\Delta)$. By the proposition, E is also v'-minimal and $\delta_{v'} = v(\Delta)$.

(b) Let v' be an extension of v to $K(\sqrt{u})$. By part (a), $\delta_v(E) = \delta_{v'}(E)$ and $\delta_v(E^u) = \delta_{v'}(E^u)$. But E and E^u are isomorphic over $K(\sqrt{u})$, hence $\delta_{v'}(E) = \delta_{v'}(E^u)$.

(c) We apply Proposition 5.3.4(b).

Suppose the pair c_4, c_6 is realized over V' but not over V. By replacing it with the pair $c_4\pi^{4i}, c_6\pi^{6i}$ for appropriate i we can assume in addition that $c_4\pi^4, c_6\pi^6$ is realized by a model E defined over V. Then E is v-minimal hence, by the proposition, v'-minimal; but this contradicts the fact that c_4, c_6 is realized over V'. $\blacksquare$

[†]The general question of how twisting affects δ_v will be considered in §5.7.

5.5 Global minimal equations and $\mathcal{A}_E$

Let A be a Krull domain with quotient field K and let E be an elliptic curve defined over K. If d is a common denominator of the Weierstrass coefficients, then the transformation $[0, 0, 0, d^{-1}]$ produces a model defined over A. Thus let us assume that E is defined over A.

The set $\mathcal{S}$ of essential valuations v such that $\delta_v > 0$ is a subset of $\{v : v(\Delta) > 0\}$, and is therefore a finite set. If $v(\Delta) = \delta_v \ \forall v$ we say that E (and Δ) is ***A*-minimal**, or that E is a **global minimal model**. In this section we develop criteria for such models to exist.

Of particular interest in arithmetical algebraic geometry are the Dedekind domains A that arise as the integral closure of $\mathbf{Z}$ (resp. $k[t]$ where k is a field and t a transcendental) in a finite extension of $\mathbf{Q}$ (resp. $k(t)$). In the first case — the case when K is a number field — we will prove in a later chapter the following theorem of Shafarevich: given a finite set $\mathcal{S}_0$ of valuations there are up to isomorphism only finitely many E for which $\mathcal{S} \subset \mathcal{S}_0$. For example, with $A = \mathbf{Z}$, since

$$y^2 = x^3 - 3(1 + 5^N)x - 2(1 + 5^N)$$

has $\delta_5 = N$, these curves are are not isomorphic over $\mathbf{Q}$, hence as we increase N we are bound to introduce more and more primes into $\mathcal{S}$. In the same vein, here is the complete list of $\mathbf{Z}$-minimal Δ that contain only the prime 2:

$$\pm 2^6,\ 2^7,\ -2^8,\ \pm 2^9,\ \pm 2^{12},\ 2^{13},\ -2^{14},\ \pm 2^{15}.$$

This list was determined by Ogg [Ogg66] and then expanded by Coghlan (see [AntIV, p.123–134]) to include all $\mathbf{Z}$-minimal Δ of the form $\pm 2^a 3^b$. Note that a $\mathbf{Z}$-minimal isomorphism class has a unique associated Δ since the only units in $\mathbf{Z}$ are ± 1 and $(\pm 1)^{12}\Delta = \Delta$.

Recall that $\forall v,\ v(\Delta) \equiv \delta_v \bmod 12$. The **Weierstrass divisor** of E is

$$\mathcal{A} = \mathcal{A}_E = \sum \frac{1}{12}(v(\Delta) - \delta_v)P_v,$$

where the sum is over all essential valuations. Let cl indicate the class of a divisor in the divisor class group. Then, following Silverman [Sil84], $\mathrm{cl}\,\mathcal{A}$ is the **Weierstrass class** of E. It follows that

$$\mathrm{cl} \sum \delta_v P_v = \mathrm{cl}\ ((\Delta) - [12]\mathcal{A}) = [12]\mathrm{cl}\,(-\mathcal{A})$$

is a multiple of 12; when A is a Dedekind domain we often switch to multiplicative notation, so the latter is described as a 12-th power in the ideal class group. When a global minimal model exists then $\sum \delta_v P_v$ is realized as the divisor of the discriminant Δ of a globally minimal Weierstrass equation.

For *minimal twists* and the associated divisor $\mathcal{A}_E^*$, see §5.7.3.

Proposition 5.5.1 *Let E be defined over the Krull domain A.*

(i) *If an A-minimal model exists then it is unique up to A-isomorphism and the divisor $\mathcal{A}$ is principal.*

(ii) *Partial converse: if $\mathcal{A}$ is principal and $6 \in A^*$ then E has a minimal model.*

(iii) *If A is a Dedekind domain we can remove the extra condition in* (ii): *an A-minimal model of E exists iff the ideal*

$$\mathcal{A}_E = \prod P^{(v(\Delta)-\delta_v)/12}$$

is principal.

Remarks. For a K-isomorphism $\tau = [r, s, t, u]$, the discriminant of τE is $u^{-12}\Delta$, hence

$$\mathcal{A}_{\tau E} = \mathcal{A}_E - \mathrm{div}(u).$$

This verifies that whether $\mathcal{A}_E$ is principal or not depends only on the K-isomorphism class of E.

There is no difficulty allowing non-integral models defined over K : then the divisor $\mathcal{A}$ is not effective, that is, not all the coefficients are ≥ 0; when divisors are written multiplicatively this means $\mathcal{A}$ is a fractional ideal.

I don't know if the converse of (i), that is (ii) with the extra condition removed, is true for UFD's, even in particular cases such as $\mathbf{Z}[t]$.

Proof. (i) If E has an A-minimal model E' then $\mathcal{A}_{E'} = 0$ and from $\Delta = u^{12}\Delta'$ we see that $\mathcal{A}_E = \mathrm{div}(u)$ is principal.

If E' and E'' are both A-minimal models of E then there exists a K-isomorphism τ from E' to E''. By Lemma 5.3.1 τ is A_P-integral for each minimal prime ideal P of A, and therefore is defined over $\cap A_P = A$ since A is Krull.

(ii) When 2 and 3 are both units in A we can assume that all Weierstrass equations are in c-form. In particular, let E be given by $y^2 = x^3 + bx + c$. Any transformation $[r, s, t, u]$ between two such equations has $r = s = t = 0$. Thus for each essential valuation v, E has a v-minimal model of the form $[0, 0, 0, d_v]E : y^2 = x^3 + bd_v^{-4} + cd_v^{-6}$, for some $d_v \in K^*$ with $v(d_v) = v(\Delta) - \delta_v$. When $\mathcal{A} =$div(u) is principal, $v(u) = v(d_v)$, and therefore d_v can be replaced by u for each v. Thus $[0, 0, 0, u]E$ is an A-minimal form.[‡]

(iii) Suppose that $\mathcal{A} = Au$ is principal. Let S be the finite set of v for which $v(\Delta) > 0$, and for each $v \in S$ let $[r_v, s_v, t_v, u]$ denote a v-isomorphism from E to a v-minimal form. Let $r \in A$ satisfy $v(r - r_v) \geq N$ where N is large, for all $v \in S$; such an r exists by the approximation theorem for Dedekind domains. Similarly choose s and t in A v-adically close to s_v, t_v $\forall v \in S$. With N large enough to ensure that the right sides of the transformation equations

[‡]If we assume only that 2 is invertible, then the equations are only in b-form, and we have to contend with transformations of the form $[r, 0, 0, u]$.

for $\tau = [r, s, t, u]$, namely $a_1 + 2s$, etc. have for each $v \in S$ the same v-adic value as the "local" versions $a_1 + 2s_v$, etc., it is clear that $E' := \tau E$ is defined over A and is A-minimal. ■

Here is an amusing example due to Setzer [Set78].

Corollary 5.5.2 *If all $\delta_v = 0$, and A is a Dedekind domain whose class group contains no elements of orders 2 or 3 then E has an A-minimal model.*[§]

Proof. $(\Delta) = \mathcal{A}_E^{12}$ is principal and therefore so is $\mathcal{A}_E$. ■

Examples E_6 and E_8 from the table in §4.4 are defined over the Dedekind domain $A = \mathbf{Z}[(1 + \sqrt{65})/2]$ and have no A-minimal model. For in both cases all $v(\Delta) = 0$ except for the ramified prime P over 5 and there $v(\Delta) = 12$. Thus by Proposition 5.3.4(c), $\mathcal{A}_E = P$ which is not principal. Setzer's result doesn't apply to these examples since the class number of $\mathbf{Q}(\sqrt{65})$ is 2.

5.6 The Laska-Kraus algorithm

We now describe an improved form of Laska's algorithm [Las82] which produces for a given E defined over a number field K a minimal model with respect to the ring of integers $\mathcal{O} = \mathcal{O}_K$ when it exists. Because of the substantial simplification that occurs, we present separately the case $\mathcal{O} = \mathbf{Z}$. The cases implemented in **aPecs**, at the time of writing, are $\mathbf{Z}$ and quadratic fields of class number 1, so that $\mathcal{O}$ is a PID.

5.6.1 The case Z

Proposition 5.6.1 *Let c_4, c_6, Δ be elements of $\mathbf{Z}$ such that*

$$c_4^3 - c_6^2 = 1728\Delta \neq 0.$$

In order that there exist a Weierstrass equation $/\mathbf{Z}$ with these covariants, it is necessary and sufficient that

(i) $v_3(c_6) \neq 2$, *and*

(ii) *either* (iia) $c_6 \equiv -1 \bmod 4$

or (iib) $c_4 \equiv 0 \bmod 16$ *and* $c_6 \equiv 0$ *or* $8 \bmod 32$.

Proof. If such an equation exists then the necessity of (i) and (ii) is an immediate consequence of Corollary 5.2.2. It is only the converse that requires a little

[§] As will be explained in Proposition 6.1.1, the condition $\delta_v = 0$ can be expressed as 'E has good reduction at v'. Thus Setzer's result can be stated: if E has good reduction everywhere and A has class number prime to 6 then E has a global minimal model.

explanation. By that corollary there exist transformations $[r_p, s_p, t_p, 1],\ p = 2, 3$ from

$$E : y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864}.$$

to $\mathbf{Z}_{(p)}$-integral curves each with the given covariants. By the approximation theorem we can pick r, s and t sufficiently p-adically close to r_p, s_p and t_p respectively for $p = 2, 3$ and p-adically integral for $p \geq 5$ so that $[r, s, t, 1]$ applied to E gives a curve over $\mathbf{Z}$ with the given covariants. ■

Here are some examples which fail one or both criteria of the proposition.

c_4	c_6	Δ	condition(s) violated
± 12	0	± 1	(ii)
0	72	-3	(i)
-15	9	-2	(i) and (ii)
52	± 368	3	(ii)
33	-207	-4	(i) and (ii)
76	± 640	17	(ii)

Thus starting with any model over $\mathbf{Z}$, to find a $\mathbf{Z}$-minimal model one could go through the finite list of integers u such that c_4u^{-4}, c_6u^{-6} are integers and test each pair according to the above proposition. This was the original approach of Laska (in general one needs to take only non-associated u's, *i.e.,* one from each coset modulo $\mathcal{O}^*$); however now we have efficient ways of calculating *all* the δ_v, hence $\mathcal{A}_E$ and when principal, say $\mathcal{A}_E = \mathcal{O}u$, we need only deal with one u.

As defined in §5.2.2, an $E_{/\mathbf{Z}}$ is of restricted type if

$$a_1, a_3 \in \{0, 1\}, \quad \text{and} \quad a_2 \in \{-1, 0, 1\}.$$

All the E in [AntIV] and [Cre92] are of restricted type.

Proposition 5.6.2 *Each E_1 defined over $\mathbf{Z}$ is $\mathbf{Z}$-isomorphic to a unique $E_{/\mathbf{Z}}$ of restricted type. Given $c_4, c_6, \Delta \in \mathbf{Z}$ satisfying the conditions of the previous proposition, the construction of this curve of restricted type is as follows:*

$$\begin{aligned}
a_1 &= \operatorname{mods}(c_4, 2), \\
a_2 &= \operatorname{mods}(-c_6 - a_1, 3), \\
a_3 &= \operatorname{mods}(\Psi_2(b_2)/16, 2) \\
&= \operatorname{mods}((b_2^3 - 3c_4b_2 - 2c_6)/16, 2), \ \textit{where } b_2 = a_1 + 4a_2, \\
a_4 &= (b_2^2 - 24a_1a_3 - c_4)/48, \\
a_6 &= (-b_2^3 - c_6 + 36b_2(a_1a_3 + 2a_4) - 216a_3)/864.
\end{aligned}$$

Remark. It is the special circumstance $\mathbf{Z}^* = \{\pm 1\}$ that gives uniqueness; *cf.* the example following Proposition 5.2.4.

Proof. Applying Proposition 5.2.4, let $E = [r_1, s_1, t_1, 1]E_1$ be of restricted type.

Now $\mathbf{Z}$-isomorphisms between elliptic curves have the form $\tau = [r, s, t, u]$ where $u \in \mathbf{Z}^* = \{1, -1\}$. Suppose $E' = \tau E$ is also of restricted type. We can assume that $u = 1$: if $u = -1$, transform the two sides with $[-1]_{E'} = [0, -a_1', -a_3', -1]$ using $[-1]_{E'}E' = E'$. By the uniqueness statement of Proposition 5.2.4, we have $\tau = [r_1, s_1, t_1, 1]$, hence E' coincides with E.

The formulas to calculate $a_1, \ldots, a_6$ from given c_4, c_6 are immediate from the relations stated in § 1.1, noticing that $a_1^2 = a_1$ since $a_1 \in \{0, 1\}$ and similarly $a_3^2 = a_3$. For example

$$a_2 = b_2 - a_1^2 - 3a_2 \equiv b_2 - a_1^2 = b_2 - a_1 \equiv b_2^3 - a_1 \equiv -c_6 - a_1 \bmod 3,$$

hence $a_2 = \operatorname{mods}(-c_6 - a_1, 3)$, and similarly for a_1, a_3. The formulas for a_4 and a_6 are just the definitions from § 1.1 of c_4 and c_6 rewritten. ■

Corollary 5.6.3 *Each $E_{/\mathbf{Q}}$ has a unique $\mathbf{Z}$-minimal model of restricted type.*

Proof. Since $\mathbf{Z}$ is a PID a $\mathbf{Z}$-minimal model always exists. Putting it in restricted form does not destroy the minimality since the transformation used has $u = 1$. The uniqueness follows from the proposition. ■

With these results we can now present the complete algorithm for $\mathbf{Z}$; only a few details are not immediately clear from the preceding, and these will be explained after stating the algorithm. In the pseudo-code, a statement such as `Let` $u = up^d$ is computerese for "replace u by up^d".

The input to the algorithm consists of integers $\Delta \neq 0, c_4, c_6$ satisfying $1728\Delta = c_4^3 - c_6^2$; it is not required that the criteria of Proposition 5.6.1 be satisfied — see the notes after the algorithm.[†] The output consists of the Weierstrass coefficients, here denoted $a_1', \ldots, a_6'$, of the unique $\mathbf{Z}$-minimal equation of restricted type E' that is $\mathbf{Q}$-isomorphic with

$$y^2 = x^3 - \frac{c_4}{48}x - \frac{c_6}{864}.$$

[†]To allow the input of rational covariants, preface the algorithm with commands that replace the input with $d^{12}\Delta, d^4c_4, d^6c_6$ for appropriate $d \in \mathbf{Z}$ to convert these three quantities into integers.

```
Let u = 1, g = gcd(c_6^2, Δ)
Loop on {p : p|6g}
    Let v = v_p, d = ⌊v(g)/12⌋, u = up^d
    If p = 2 then
        Let a = c_4 2^{-4d}, b = c_6 2^{-6d}
        If v(a) = 0 and mods(b, 4) = -1 then
            Let τ = -(b + 1)/4
        Else if v(a) > 3 and mods(b, 32) ∈ {0, 8} then
            Let τ = b/8
        Else
            Let u = u/2, τ = 0
        End if
    Else if p = 3 then
        If v(c_6) = 6d + 2 then let u = u/3 End if
    End if
End loop
Let
        c'_4  =  c_4 u^{-4},
        c'_6  =  c_6 u^{-6},
        a'_1  =  mods(c'_4, 2),
        a'_2  =  mods(-a'_1 - c'_6, 3),
        a'_3  =  mods(τ + a'_1 a'_2, 2),
        b'_2  =  a'_1 + 4a'_2,
        a'_4  =  (b'_2^2 - c'_4 - 24a'_1 a'_3)/48,
        a'_6  =  (36b'_2(a'_1 a'_3 + 2a'_4) - b'_2^3 - c'_6 - 216a'_3)/864
Return(a'_1, ..., a'_6)
```

The denominator of u is a divisor of 6 and can be > 1 when the integers Δ, c_4, c_6 do not satisfy the criteria (i) and (ii) of Proposition 5.6.1; that is the reason for including $p = 2$ and 3 in the p-loop even when $p \nmid g$. In many applications one begins with an $E_{/\mathbf{Z}}$, with Δ, c_4, c_6 calculated, and then the criteria are automatically satisfied. In that situation one may wish to append the calculation of the transformation $T = [r, s, t, u]$ from E to E': u is determined by the algorithm and then s, r, t are calculated from the transformation equations of Proposition 4.1.1:

$$s = (ua'_1 - a_1)/2,$$

$$\begin{aligned} r &= (u^2 a_2' - a_2 + s a_1 + s^2)/3, \\ t &= (u^3 a_3' - a_3 - r a_1)/2. \end{aligned}$$

T can be used to transfer points between E and E'. (If there is an initial transformation $a_i \mapsto d^i a_i$ to clear denominators then this T should be multiplied by $[0, 0, 0, d^{-1}]$.)

The method used to calculate a_3' is different from the one stated in the previous proposition, and depends on the fact that the residue field of v_2 is $\mathbf{F}_2$, rather than some extension of $\mathbf{F}_2$: in the notation of Proposition 5.2.1, case $p = 2$, the general form of τ is

$$\tau = a_3 + a_1 a_2 + \theta\alpha(\theta + \alpha) + 2\beta \equiv a_3 + a_1 a_2 \bmod 2$$

for all $\theta, \alpha \in \mathbf{Z}$. Moreover every case over $\mathbf{Z}$ is one of the special cases (i) or (ii) of that proposition (*cf.* Corollary 5.2.2). The algorithm calculates τ according to the information given for these special cases in Proposition 5.2.1, and then calculates $a_3' = \operatorname{mods}(\tau + a_1' a_2', 2)$.

We illustrate the algorithm with two examples of particular interest: $E_{/\mathbf{Q}}$ with $j = 1728 = 12^3$ and $j = 0$. We also include the case $j = 66^3$ since such curves occur naturally in the discussion of $j = 12^3$. ($E_{/\mathbf{Q}}$ with $j = 66^3$ arise from applying 2-isogenies to curves with $j = 12^3$, as will be explained in Chapter 6.)

First we make a convenient definition: for $s \in \mathbf{Q}^*$ and a positive integer n, the ***n*-th power free part** of s is the unique n-th power free integer N such that $s = Nt^n$ where $t \in \mathbf{Q}^*$ and also $N > 0$ when n is odd.

Corollary 5.6.4 *Let $E_{/\mathbf{Q}}$ have invariant $j = 12^3$, 66^3 or 0 and covariants c_4, c_6, Δ. Then the $\mathbf{Z}$-minimal model of restricted type is as follows.*

$\boxed{j = 12^3}$

$$\begin{aligned} E_{\min} &: \quad y^2 = x^3 + Bx, \\ \textit{covariants} &: \quad -48B,\ 0,\ -64B^3 \end{aligned}$$

where B is the 4-th power free part of $-27c_4$. Hence a Weierstrass equation $y^2 = x^3 + a_4 x$ is $\mathbf{Z}$-minimal iff a_4 is 4-th power free.

$\boxed{j = 66^3}$

$$\begin{aligned} E_{\min} &: \quad y^2 = x^3 - 11D^2 x - 14D^3, \\ \textit{covariants} &: \quad 2^4 \cdot 3 \cdot 11D^2,\ 2^6 \cdot 3^3 \cdot 7D^3,\ 2^9 D^6 \end{aligned}$$

where D is the square free part of $77c_4c_6$.

$\boxed{j = 0}$

$$\begin{aligned} E_{\min} &: \quad y^2 = x^3 + C, \\ \textit{covariants} &: \quad 0,\ -864C,\ -432C^2 \end{aligned}$$

where C is the 6-th power free part of $-54c_6$ except when C has the form $16(4c+1)$, $c \in \mathbf{Z}$, and then the minimal model is obtained by the transformation $[0,0,4,2]$, i.e., *the substitutions $x = 4x'$, $y = 8y' + 4$:*

$$\begin{aligned} E_{\min} &: \quad y'^2 + y' = x'^3 + c, \\ \textit{covariants} &: \quad 0,\ -216(4c+1),\ -27(4c+1)^2. \end{aligned}$$

Proof. Let $j = 1728$. Then E is $\mathbf{Q}$-isomorphic to all curves of the form

$$y^2 = x^3 - \frac{u^4 c_4}{48} x, \quad u \in \mathbf{Q}^*,$$

and in particular to $y^2 = x^3 + Bx$ where B is the 4-th power free part of $-27c_4$. With this curve as input, that is, with $c_4 = -48B, c_6 = 0, \Delta = -64B^3$, it is a simple matter to trace through the algorithm to find that this curve is $\mathbf{Z}$-minimal. We mention two points:

(i) it would not improve the algorithm to replace $g = \gcd(0, \Delta) = 64|B|^3$ with $\gcd(c_4^3, \Delta)$;

(ii) in the loop when $p = 3$, we have $v(c_6) = \infty \neq 6d + 2$.

The curve

$$E_1 : y^2 = x^3 - 11x - 14$$

has $j = 66^3$ and every $E_{/\mathbf{Q}}$ with $j = 66^3$ is isomorphic to a quadratic twist of E, hence to a curve of the type displayed in the corollary. Again the algorithm immediately shows that these curves are $\mathbf{Z}$-minimal.

When $j = 0$, E is isomorphic to $y^2 = x^3 + C$ which has covariants $c_4 = 0$, $c_6 = -2^5 3^3 C$, $\Delta = -2^4 3^3 C^2$. With this input, the algorithm validates the stated results. (Of course one can also use Proposition 5.6.1 directly.) ■

5.6.2 The general number field case

We use the notation of §5.2.2. It is assumed that algorithms for ordinary and ideal arithmetic in the number field K are in place. In particular our algorithms assume that an integral basis has been chosen, so that $\operatorname{mods}(\alpha, m)$ is unambiguously defined for $\alpha \in \mathcal{O}$. Also they call two special procedures. The first, which we label **Factor**, factorizes a given ideal into prime ideal powers; the second, labelled **Gen**, determines whether a given ideal is principal, and when so finds a generator. Specifically, Factor($\Delta\mathcal{O}$) returns a list **factors** consisting of quadruples P, c, v, p where $\Delta\mathcal{O} = \prod P^c$ is the factorization, v is the P-adic valuation on K and p is the residue field characteristic; when P or v is chosen, the other three items in the quadruple are assigned the corresponding values. Thus a sequence of pseudo-code such as

```
Call Factor(ΔO)
Loop on v ∈ factors
  If p = 2 then ...
```

should be self-explanatory. Calling Gen($\mathcal{A}_E$) assigns to u a value which is either the word *FAIL*, meaning $\mathcal{A}_E$ is not principal hence E does not have a global minimal model, or a generator: $\mathcal{A}_E = u\mathcal{O}$. As in the algorithm for $\mathbf{Z}$, `Let` $m = m - 1$ is pseudo-code for "replace m by $m - 1$", for instance.

The following algorithm takes as input the Weierstrass coefficients $a_1, \ldots, a_6$ of an $E_{/\mathcal{O}}$ (we don't take the trouble to allow the input of any $c_4, c_6, \Delta \in \mathcal{O}$ satisfying $c_4^3 - c_6^2 = 1728\Delta$ as we did in the case $\mathcal{O} = \mathbf{Z}$) and outputs either the message that E does not have a global minimal model, or the Weierstrass coefficients $a'_1, \ldots, a'_6$ of a global minimal model E' of restricted type, together with a transformation $[r, s, t, u] : E \longrightarrow E'$. The algorithm first determines all δ_v, hence $\mathcal{A}_E$. In order to make the presentation not too cumbersome, when the residue characteristic $p = 3$ or 2, to calculate δ_v the algorithm calls the procedure labelled L3 or L2; these procedures are presented after the algorithm.

The coding is for the most part clear from the preceding results of this chapter, with frequent references to the boxed relations in §1.1 and the transformation formulas in Proposition 4.1.1. Thus we feel it necessary to add only one explanatory remark concerning L3. In an actual implementation there are some obvious programming details that are not spelled out in the code, *e.g.* a preliminary preparation for table look-up of the unique 2-restricted solution x of $x^2 \equiv b \bmod 4$ (or possibly $x =$ FAIL) — at least when more than one E is being considered (as opposed to considering a fixed E over varying K).

The main algorithm:

```
Input c4, c6, Δ of E/O
Let A_E = 1
Call Factor(ΔO)
Loop on v ∈ factors
  Let a, b, c = v(c4), v(c6), v(Δ)
  Let m = min{⌊a/4⌋, ⌊b/6⌋, ⌊c/12⌋}
  If m = 0 Then Let δ_v = c
  Else if p ≥ 5 Then Let δ_v = c − 12m
  Else if p = 3 Then Call L3
  Else Let c''_4 = c_4, c''_6 = c_6, N = 0; Call L2; Let δ_v = δ_v − N
  End if
  Let A_E = A_E * P^((c−δ_v)/12)
End loop
Call Gen(A_E)
If u = FAIL Then Return("Global min. model does not exist")
Else
  Let c'_4 = c_4 u^−4,  c'_6 = c_6 u^−6
  Loop on 2-restricted soln.'s a'_1 of a'_1^4 ≡ c'_4 mod 8
    Let s = (u a'_1 − a_1)/2
    If s ∉ O Then Next a'_1 End if
    Loop on 3-restricted soln.'s a'_2 of a'_2^3 ≡ −c'_6 − a'_1^6 mod 3
```

```
        Let r = (u^2 a'_2 - a_2 + s a_1 + s^2)/3
        If r ∉ O Then Next a'_2 End if
        Let
          b'_2 = a'_1^2 + 4a'_2,
          b'_4 = (b'_2^2 - c'_4)/24,
          b'_6 = (-b'_2^3 + 36 b'_2 b'_4 - c'_6)/216
        If b'_4 ∉ O Or b'_6 ∉ O Then Next a'_2 End if
        If x^2 ≡ b'_6 mod 4 has a sol.'n Then
          Let a'_3 be the unique 2-restricted sol.'n
        Else Next a'_2
        End if
        Let t = (u^3 a'_3 - a_3 - r a_1)/2
        Let a'_4 = (b'_4 - a'_1 a'_3)/2
        If t ∉ O Or a'_4 ∉ O Then Next a'_2 End if
        Let a'_6 = (b'_6 - a'_3^2)/4
        If a'_6 ∈ O Then Return(a'_1, ..., a'_6, r, s, t, u) End if
      End a'_2-loop
    End a'_1-loop
  End if
```

The procedure L3:

```
  If v(3) = 1 Then
    If b = 2 + 6m Then Let δ_v = c - 12(m - 1)
    Else Let δ_v = c - 12m
    End if
  Else
    If a = 4m Or b ≥ 6m + v(27) Then
      Let δ_v = c - 12m
    Else
      Let n = 0
      Loop on α in a set of coset rep.'s of O/P^{6m+v(3)-min{2v(b_2),v(9),4m}}
        Let ϑ = b_2 + 3α, d = ⌊v(Ψ_2(ϑ)/27)/6⌋
        If d ≥ m Then
          Let n = m
          Exit this loop
        Else
          If d > n Then Let n = d End if
        End if
      End loop
      Let δ_v = c - 12n
    End if
  End if
  Return(δ_v)
```

Note. Let us explain the exponent on P in the α-loop. Of course the objective is to get as small an exponent as possible; it may well be that this exponent can be improved. Let π denote a uniformizer for v. Since E is defined over the valuation ring of v, by Proposition 5.2.1 we know that $\Psi_2(b_2) \equiv 0 \bmod 27$. We wish to find $\vartheta = b_2 + 3\alpha$ which maximizes n, subject to $n \le m$, in

$$\Psi_2(\vartheta) = \vartheta^3 - 3c_4\vartheta - 2c_6 = 27\pi^{6n}\lambda, \quad v(\lambda) \ge 0.$$

Then, since $v(c_4) \ge 4m \ge 4n$ and $v(c_6) \ge 6m \ge 6n$, we have $v(\vartheta) \ge 2n$ and defining $\vartheta_1 = \vartheta\pi^{-2n}$, $c_4' = c4\pi^{-4n}$, $c_6' = c_6\pi^{-6n}$,

$$\vartheta_1^3 - 3c_4'\vartheta_1 - 2c_6' \equiv 0 \bmod 27.$$

When n is maximal this means that $\delta_v = c - 12n$.

To avoid duplication of effort in the loop we wish to know what $i \ge 0$ guarantees that if $\alpha' = \alpha + \pi^i\lambda$, $v(\lambda) \ge 0$, and $\vartheta' = \vartheta + 3\pi^i\lambda$, then $\Psi_2(\vartheta') \equiv \Psi_2(\vartheta) \bmod 27\pi^{6m}$. Now

$$\Psi_2(\vartheta') - \Psi_2(\vartheta) = 9\pi^i\lambda\left[\vartheta^2 + 3\pi^i\lambda\vartheta + 3\pi^{2i}\lambda^2 - c_4\right]$$

$$= 9\pi^i\lambda\mu \quad \text{say,}$$

and using $v(\vartheta) = v(b_2 + 3\alpha) \ge \min\{v(b_2), v(3)\}$, one can check that $i = 6m + v(3) - \min\{2v(b_2), v(9), 4m\}$ implies that the value of each of the four terms in μ is at least $\min\{2v(b_2), v(9), 4m\}$, which guarantees that $v(\Psi_2(\vartheta') - \Psi_2(\vartheta)) \ge v(27) + 6m$.

The procedure L2:

```
Let π ⊂ 𝒪 be a uniformizer for v
Let
  c'_6 = c''_6 π^{-6m}, δ_v = -1
  Ψ'_2 = x^3 - 3c''_4 x - 2c''_6
  Ψ'_3 = x^4 - 6c''_4 x^2 - 8c''_6 x - 3c''_4^2
If a = 4m And x^2 ≡ -c'_6 mod 4 has a sol.'n x ∈ 𝒪
  Or a ≥ 4m + v(16) And 8x^2 ≡ c'_6 mod 32 has a sol.'n x ∈ 𝒪
  Then Let δ_v = c - 12m
Else if v(2) > 1 Then
  Let n = 0, δ_v = c
  Loop on α in a set of coset rep.'s of 𝒪/P^{8m+v(16)}
    Let θ = a_1 + 2α, d = min{⌊v(Ψ'_3(θ^2)/256)/8⌋, m}
    If d > n Then
      Let n' = 0
      Loop on β in a set of coset rep.'s of 𝒪/P^{6m+v(2)}
        Let τ = a_3 + a_1 a_2 + θα(θ + α) + 2β
        Let e = min{⌊v((Ψ'_2(θ^2) + 16τ^2)/64)/6⌋, m}
```

```
          If e > n' Then
            Let n' = e
            If e > n Then
              Let n = min{d, e}, δ_v = c − 12n
              If n = m Then Return(δ_v)
                Exit both α and β loops
              End if
            End if
          End if
        End β loop
      End if
    End α loop
  End if
  If δ_v = −1 Then
    If m = 1 Then δ_v = c Return(δ_v)
    Else
      Let a = a − 4, m = m − 1, c''_4 = c''_4 π^{−4}, c''_6 = c''_6 π^{−6}, N = N + 12
      Call L2
      Return(δ_v)
    End if
  End if
  Return(δ_v)
```

5.7 How twisting affects minimal models

We assume throughout this section that char $K \neq 2$, for twisting in characteristic 2 takes a different turn, as it were, and requires separate treatment; *cf.* §4.3.

5.7.1 The local case

Recall from §5.3 that $\delta_v(E)$ denotes $v(\Delta_{\min})$ where $\Delta_{\min}$ is the discriminant of a v-minimal model of E. How do $\delta_v(E)$ and $\delta_v(E^u)$ compare? A particular case was already treated in Corollary 5.4.2. We make two simple observations:

(i) Since twisting by a square gives a K-isomorphic curve, we can assume that $v(u) = 0$ or 1.

(ii) E and E^u appear symmetrically in the discussion: $(E^u)^{u^{-1}} = E$. Thus in part (c) of the next proposition, we can choose the notation so that $\delta_v(E^u) \geq \delta_v(E)$.

In the case of residue field characteristic 2, it is very confusing to deal with δ_v alone; it is much better to consider the whole v-adic signature to sort out the different cases. We include a detailed statement of the possibilities when $v(2) = 1$ in the following

Proposition 5.7.1 *Let* char $K \neq 2$, *let* v *be a valuation on* K *with residue field* k, *let* $u \in K^*$ *with* $v(u) = 0$ *or* 1, *and let* E *be an elliptic curve defined over* K; *let the covariants of a* v*-minimal model of* E *be* c_4, c_6, Δ *and define* $\lambda_v = \min\{3v(c_4), 2v(c_6), v(\Delta)\}$.

(a) *If* $\operatorname{char} k \neq 2$ *then*

$$\delta_v(E^u) = \delta_v(E) \pm 6v(u).$$

In particular, $\delta_v(E^u) = \delta_v(E)$ *when* $v(u) = 0$. *If* $\operatorname{char} k \neq 2$ *or* 3, *then the sign is* $+$ *when* $\lambda_v < 6$ *and* $-$ *when* $\lambda_v \geq 6$.

(b) *If* char $K = 0$ *and* $v(3) = 1$, *hence* $\operatorname{char} k = 3$, *then when* $v(u) = 1$

$$\delta_v(E^u) = \begin{cases} \delta_v(E) + 6 & \text{if } \lambda_v < 6 \text{ or } v(c_6) = 5, \\ \delta_v(E) - 6 & \text{if } \lambda_v \geq 6 \text{ and } v(c_6) \neq 5. \end{cases}$$

(c) *If* char $k = 2$ *(hence* char $K = 0$ *since we have assumed* char $K \neq 2$*), and let us say* $\delta_v(E^u) \geq \delta_v(E)$, *then*

$$\delta_v(E^u) = \delta_v(E) + 6v(u) + 12\nu \quad \text{where } \nu \in \mathbf{Z} \text{ and } 0 \leq \nu \leq v(2).$$

In particular, $\delta_v(E^u) = \delta_v(E)$ *when* v *is unramified in the extension* $K(\sqrt{u})$ (*e.g.* when $z^2 \equiv u \bmod 8$ has a solution $z \in K$).

(d) *Let* $v(2) = 1$, *so* $\operatorname{char} K = 0$ *and* $\operatorname{char} k = 2$, *assume (for convenience) that* E *is* v*-minimal, and let* $\operatorname{sig}(E) = v(c_4), v(c_6), v(\Delta)$ *denote the* 2*-adic signature of* E *(so the third member is* $v(\Delta) = \delta_v(E)$*). Then the value, or two possible values, of* $\delta_v(E^u)$ *can be read from the following table, with the natural conventions for the symbol* ∞. *If* $\operatorname{sig}(E) = a, b, c$ *then the symbol* $\bullet$ *in the column headed* $v(u) = 0$ (resp. $v(u) = 1$) *stands for* a, b, c (resp. $a + 2, b + 3, c + 6$).

$\operatorname{sig}(E)$	$\operatorname{sig}(E^u)$ *when* $v(u) = 0$	$\operatorname{sig}(E^u)$ *when* $v(u) = 1$
$0, 0, c\,(c \geq 0)$	$\bullet$ *or** $4, 6, c + 12$	$6, 9, c + 18$
$4, 6, c\,(c \geq 6)$	$\bullet$ *or*† $0, 0, c - 12$	$\bullet$
$6, 9, c\,(c \geq 12)$	$\bullet$	$4, 6, c - 6$ *or*§ $0, 0, c - 18$
$4, b, 6\,(7 \leq b \leq \infty)$	$\bullet$	$\bullet$
$5, b, 9\,(8 \leq b \leq \infty)$	$\bullet$	$\bullet$
$6, b, 12\,(10 \leq b \leq \infty)$	$\bullet$	$4, b - 3, 6$
$7, b, 15\,(11 \leq b \leq \infty)$	$\bullet$	$5, b - 3, 9$
$a, 3, 0\,(4 \leq a \leq \infty)$	$\bullet$ *or** $a + 4, 9, 12$	$\bullet$
$a, 5, 4\,(4 \leq a \leq \infty)$	$\bullet$	$\bullet$
$a, 6, 6\,(5 \leq a \leq \infty)$	$\bullet$	$\bullet$ *or*¶ $a - 2, 3, 0$
$a, 7, 8\,(5 \leq a \leq \infty)$	$\bullet$	$\bullet$
$a, 8, 10\,(6 \leq a \leq \infty)$	$\bullet$	$a - 2, 5, 4$
$a, 9, 12\,(7 \leq a \leq \infty)$	$\bullet$ *or*‡ $a - 4, 3, 0$	$a - 2, 6, 6$
$a, 10, 14\,(7 \leq a \leq \infty)$	$\bullet$	$a - 2, 7, 8$

* *The first alternative if* $u \equiv z^2 \bmod 4$ *has a solution, the second if not.*

† *The second alternative if* $c \geq 12$ *and* $2^{-6}c_6u \equiv -z^2 \bmod 4$ *has a solution, the first otherwise.*
‡ *The second alternative if* $a \geq 8$ *and* $2^{-9}c_6u \equiv z^2 \bmod 4$ *has a solution, the first otherwise.*
§ *The second alternative if* $c \geq 18$ *and* $c_6u^{-9} \equiv -z^2 \bmod 4$ *has a solution, the first otherwise.*
¶ *The second alternative if* $a \geq 6$ *and* $2^{-9}c_6u^3 \equiv x^2 \bmod 4$ *has a solution, the first otherwise.*

Remark. In the situation of footnote † or ‡ to the table in part (d), the congruence cannot have a solution when $u \equiv 1 \bmod 4$ since otherwise $\mathrm{sig}(E)$ could be reduced and E would not be v-minimal. It follows that the first alternative obtains when $u \equiv 1 \bmod 4$.

Proof. Replacing E with a K-isomorphic curve replaces E^u with a K-isomorphic curve, so we can assume that E is v-minimal, *i.e.,* defined over V with $v(\Delta(E)) = \delta_v(E)$. Since $\mathrm{char}\, K \neq 2$, K-isomorphisms change the discriminant by 12-th power factors. Thus

$$\delta_v(E^u) \equiv v(\Delta(E^u)) \equiv v(\Delta(E)) + 6v(u) = \delta_v(E) + 6v(u) \bmod 12,$$

say

$$\delta_v(E^u) = \delta_v(E) + 6v(u) + 12\nu, \quad \text{some } \nu \in \mathbf{Z}. \tag{\#}$$

(a) By Proposition 5.3.2 we can assume that E is in b-form. Then, since $\mathrm{char}, k \neq 2$, 2 is invertible in V and E^u is V-integral. Hence

$$\delta_v(E^u) \leq v(\Delta(E^u)) = v(\Delta(E)) + 6v(u) = \delta_v(E) + 6v(u).$$

Since the roles of E and E^u can be interchanged, the result just obtained implies $\delta_v(E) \leq \delta_v(E^u) + 6v(u)$. Hence $\nu = 0$ or -1 in (#). When $\mathrm{char}\, k \neq 2, 3$, then $\lambda_v = \min\{3v(c_4), v(\Delta\}$ and we can apply Proposition 5.3.4(c):

$$\delta_v(E^u) = \delta_v(E) + 6v(u) - \min\{\lfloor (v(c_4) + 2v(u))/4 \rfloor, \lfloor (v(\Delta) + 6v(u))/12 \rfloor\}.$$

Hence $\nu = 0$ when $\lambda_v < 6$, *i.e.,* either $c_4 < 2$ or $v(\Delta) < 6$, and $\nu = 1$ otherwise.

(b) Again we can assume that $E = E'$, so if the signature of E is a, b, c then that of E^u is $a+2, b+3, c+6$. This can be reduced to $a-2, b-3, c-6$, that is, $c_4 3^{-2}, c_6 3^{-3}$ are the covariants of some $E'_{/V}$ which is then a v-minimal model of E^u, precisely when $\lambda_v \geq 6$ (in order that all of $a-2, b-3, c-6$ are ≥ 0) and $b - 3 \neq 2$ (to satisfy Corollary 5.2.2).

(c) Define

$$c'_4 = (4u)^2 c_4, \quad c'_6 = (4u)^3 c_6, \quad \Delta' = (4u)^6 \Delta.$$

These are the covariants of the curve $[0, 0, 0, 1/2]E^u$ which is K-isomorphic with E^u. Since $v(c'_4) \geq v(16)$ and $v(c'_6) \geq v(64)$, so that $c'_6 \equiv 8x^2 \bmod 32$ has the

solution $x = 0$, special case (ii) of Proposition 5.2.1 is satisfied, hence c_4', etc. occur as the covariants of some E_1 defined over V. This implies that

$$\delta_v(E^u) \leq v(\Delta(E_1)) = v(\Delta') = v(\Delta) + 6v(u) + 12v(2).$$

This proves that $\nu \leq v(2)$.

When $z^2 \equiv u \bmod 8$ has a solution then Hensel's lemma shows that $z^2 = u$ has a solution in the local field K_v. Hence the local degree $ef = [K_v(\sqrt{u}) : K_v] = 1$, so no extension of v to $K(\sqrt{u})$ can be ramified. We mention this case because the details concerning the congruences in Proposition 5.2.1 can be made explicit: if θ, τ are solutions of the congruences for the covariants c_4, c_6 of E, then applying Newton's method once to $z^2 \equiv u \bmod 8$ to obtain $z_1^2 \equiv u \bmod 64$, an easy calculation shows that $\theta z_1, \tau z_1^3$ are solutions of the congruences for $c_4 u^2, c_6 u^3$.

(d) We can assume that E is v-minimal, hence $c_4, c_6, \Delta \in V$ and therefore the covariants $u^2 c_4, u^3 c_6, u^6 \Delta$ of E^u are also in V. If the signature of E is a, b, c then that of E^u is $a+2, b+3, c+6$. Referring to the table for $p = 2$ in Proposition 5.3.5, the table of the present proposition is produced case by case, testing whether the signature of E^u can be reduced a notch to $a-2, b-3, c-6$, or can stay the same, or must be jacked up to $a+6, b+9, c+18$. And in any particular case where there are two possibilities, it is clear which prevails. For instance when $a, b, c = 0, 0, c$ and $v(u) = 0$, the alternative $4, 6, c+12$ obtains when $u \equiv z^2 \bmod 4$ has no solution, since $c_6 \equiv -x^2 \bmod 4$ has a solution by Proposition 5.2.1, and therefore $c_6' = u^3 c_6 \equiv -x^2 \bmod 4$ does not.

When v is unramified in $K(\sqrt{u})$ then $v(u) = 0$ (since $v(u)$ was assumed to be 0 or 1), and $\nu = 0$ by Corollary 5.4.2. ■

5.7.2 The global case

Let A be a Krull domain and E an elliptic curve defined over A. It is not generally true that the existence of an A-minimal model for E implies the same for twists E^u. In fact one can expect quite the opposite, as is explained in the next proposition; the number theory case is due to Silverman [Sil84].

As in §2.2, $\mathrm{Cl}(A)$ denotes the divisor class group, and $\mathrm{cl}(D)$ the class of a divisor D.

Proposition 5.7.2 *Let A be a Krull domain and $E_{/A}$ an elliptic curve. Then the map $A - \{0\} \longrightarrow \mathrm{Cl}(A)$ defined by*

$$u \longmapsto \mathrm{cl}\,(\mathcal{A}_{E^u})$$

is surjective in the following two cases:

(i) *2 is invertible ($2 \in A^*$) and $\mathrm{Cl}(A)$ is finite.*

(ii) *A is the ring of integers in a number field.*

Hence in these cases when the class number > 1, every $E_{/A}$ has a twist for which there is no global minimal model.

Proof. (i) Let K denote the quotient field of A. Suppose $u \in A$ is such that

$$v(\Delta(E)) > 0 \implies v(u) = 0.$$

Since $2 \in A^*$, residue field characteristic 2 does not occur, and by the previous proposition, for all v

$$\delta_v(E^u) = \delta_v(E) + 6v(u) - 12\lfloor v(u)/2 \rfloor.$$

(In the local calculation to determine $\delta_v(E^u)$, when $v(u) > 1$ we can apply the transformation $[0, 0, 0, \pi^{\lfloor v(u)/2 \rfloor}]$ to E^u, where π is a uniformizer for v, to reduce $v(\Delta(E^u))$.) Since

$$v(\Delta(E^u)) = v(\Delta(E)) + 6v(u),$$

and

$$\mathcal{A}_E = \sum_v \frac{1}{12}(v(\Delta(E)) - \delta_v(E))P_v,$$

therefore

$$\mathcal{A}_{E^u} = \mathcal{A}_E + \sum \lfloor v(u)/2 \rfloor P_v = \mathcal{A}_E + D, \quad \text{say},$$

and it remains to show that u can be chosen so that $\mathrm{cl}(D)$ falls in a given class c.

Since c can be written as $c_1 + c_2$ where c_1 has odd order $2m+1$ and c_2 has order of the form 2^n, and since the composition of twists is a twist, the general statement will follow from the two particular cases $c = c_1$, $c = c_2$.

By Proposition 2.2.3(b), choose a squarefree $D_1 = \sum P_i \in [-2]c_1$ with support disjoint from that of $\mathrm{div}(2\Delta(E))$, so that $[2m+1]D_1 = \mathrm{div}(u)$ is principal. Then since $v_{P_i}(u) = 2m+1$ and $v(u) = 0$ for all other v, in the formula above $D = \sum [m]P_i \in c_1$.

Secondly choose squarefree $D_i \in [2^i]c_2$, $0 \le i \le n-1$, with mutually disjoint support and also with support disjoint from the support of $\mathrm{div}(2\Delta(E))$. Thus

$$[2]P_0 + P_1 + \cdots + P_{n-1} = \mathrm{div}(u)$$

is principal. Then $D = D_1 \in c_2$.

(ii) For Silvermans's proof (which I have not personally verified to the last detail) see [Sil84]. ■

When $K = \mathbf{Q}$ Proposition 5.7.1 can be globalized in another way. We write δ_p and λ_p for δ_{v_p} and λ_{v_p}.

Proposition 5.7.3 *Let E be defined over $\mathbf{Z}$ and $\mathbf{Z}$-minimal, let the covariants of E be c_4, c_6, Δ, let u be a square-free integer and let Δ' be the minimal discriminant of the quadratic twist $E * u$. Then Δ and Δ' have the same sign and the same p-adic values except for the prime divisors p of u and, when u is odd, possibly also $p = 2$:*

- *If p is an odd prime divisor of u then $v_p(\Delta') = v_p(\Delta) \pm 6$; the sign is $+$ if $\lambda_p < 6$ or if $p = 3$ and $v(c_6) = 5$, otherwise the sign is $-$.*

- *The value of $D := v_2(\Delta') - v_2(\Delta)$ is as follows:*

  $u \equiv 1 \bmod 4 : D = 0;$

  $u \equiv 3 \bmod 4 : D$ *is*

  12 *if* $\mathrm{sig}(E) = 0, 0, c$ *or* $a, 3, 0,$

  -12 *if* $\mathrm{sig}(E) = 4, 6, c$ *or* $a, 9, 12,$

  0 *otherwise;*

  $u \equiv 2 \bmod 4$, *say* $u = 2w : D$ *is*

  18 *if* $\mathrm{sig}(E) = 0, 0, c;$

  -18 *if* $\mathrm{sig}(E) = 6, 9, c$ *with* $c \geq 18$ *and* $2^{-9} c_6 w \equiv -1 \bmod 4;$

  6 *if* $v(c_4) = 4$ *or* 5, *or* $v(c_6) = 3, 5,$ *or* 7, *or* $\mathrm{sig}(E) = a, 6, 6$ *with* $a \geq 6$ *and* $2^{-6} c_6 w \equiv -1 \bmod 4;$

  -6 *otherwise.*

Proof. Since $\Delta = u^6 w^{12} \Delta'$ for some $w \in \mathbf{Q}^*$, the statement that Δ, Δ' have the same sign is clear. The rest is a straightforward application of the Proposition 5.7.1. ■

Here are two particular examples of the corollary. If E is the curve

$$y^2 + y = x^3 - x^2, \quad c_4 = 2^4,\, c_6 = -2^3 * 19,\, \Delta = -11, \qquad (\mathbf{A11})$$

u is a square-free integer, and Δ' denotes the discriminant of a $\mathbf{Z}$-minimal model of the twist E^u, then

$$\Delta' = \begin{cases} u^6 \Delta & \text{if } u \not\equiv 3 \bmod 4, \\ 2^{12} u^6 \Delta & \text{if } u \equiv 3 \bmod 4. \end{cases}$$

The results come out slightly differently for the curve

$$y^2 + xy + y = x^3 + x^2, \quad c_4 = 1,\, c_6 = -7 * 23,\, \Delta = -15. \qquad (\mathbf{A15})$$

Then

$$\Delta' = \begin{cases} u^6 \Delta & \text{if } u \equiv 1 \bmod 4, \\ 2^{12} u^6 \Delta & \text{if } u \not\equiv 1 \bmod 4. \end{cases}$$

In both these examples we notice that we always have $v_p(\Delta') \geq v_p(\Delta)$ for *all* p. The curves **A11** and **A15** are examples of *global minimal twists*, which are discussed in the next section.

5.7.3 Minimal twists

We are led inexorably to the definition

$$\begin{aligned}\delta_v^*(E) \quad = \quad & \min\{\delta_v(E') : E' \text{ is a twist of } E\text{, } i.e.,\ E' \text{ is defined over } K \\ & \qquad \text{and has the same } j\text{-invariant as } E\}\end{aligned}$$

A twist E' of E which has $v(\Delta(E')) = \delta_v^*(E)$ is a ***v*-minimal twist** of E. Then E' is v-minimal, but it need not be a v-minimal model of E since E and E' are not necessarily K-isomorphic. Moreover, different v-minimal twists of E need not be K-isomorphic, though of course they are twists of one another. For example, if $E' : y^2 = x^3 + bx + c$ is a v-minimal twist (hence $\operatorname{char} \tilde{v} \neq 2$) and u is an element of K such that $v(u) = 0$ and $\sqrt{u} \notin K$, then $y^2 = x^3 + bu^2x + cu^3$ is another v-minimal twist of E' that is not K-isomorphic with E'.

With K fixed, $\delta_v^*(E)$ depends on E only up to twists, and therefore a valid alternative notation, which we will use occasionally, is $\delta_v^*(j)$.

For example Corollary 5.6.4 yields the results for $K = \mathbf{Q}$:

$$\delta_2^*(1728) = 6, \quad \delta_p^*(1728) = 0 \quad \text{for all} \quad p > 2;$$

$$\delta_2^*(66^3) = 9, \quad \delta_p^*(66^3) = 0 \quad \text{for all} \quad p > 2;$$

$$\delta_3^*(0) = 3, \quad \delta_p^*(0) = 0 \ \text{ for all } \ p \neq 3.$$

When $j \neq 1728$ or 0, all twists are quadratic (Proposition 4.3.2) and Proposition 5.7.1 implies various relationships between δ_v and δ_v^* which we repeat for future reference.

Proposition 5.7.4 *Let* $\operatorname{char} K \neq 2$*, let v be a valuation on K with residue field k, let E be a v-minimal elliptic curve defined over K with invariant $j \neq 1728$ or 0 and covariants c_4, c_6, Δ and, as in the previous proposition, define* $\lambda_v = \min\{3v(c_4), 2v(c_6), v(\Delta)\}$.

(a) *If* $\operatorname{char} k \neq 2$ *then* $\delta_v^*(E) = \delta_v(E) - 6$ *or* $\delta_v(E)$. *If* $\operatorname{char} k \neq 2$ *or* 3 *then* $\delta_v^*(E) = \delta_v(E) - 6$ *when* $\lambda_v \geq 6$, *and* $\delta_v^*(E) = \delta_v(E)$ *when* $\lambda_v < 6$.

(b) *If* $\operatorname{char} K = 0$ *and* $v(3) = 1$ *then*

$$\delta_v^*(E) = \begin{cases} \delta_v(E) - 6 & \textit{when } \lambda_v \geq 6 \textit{ and } v(c_6) \neq 5 \\ \delta_v(E) & \textit{otherwise} \end{cases}$$

(c) *If* $\operatorname{char} k = 2$ *then* $\delta_v^*(E) = \delta_v(E) - 6x$ *for some integer x satisfying* $0 \leq x \leq 2v(2) + 1$. *(When $v(2) = 1$ the detailed conditions for the four possible values for x are given in the previous proposition.)* ■

Now let E be defined over the quotient field K of a Krull domain A. Let $\mathcal{V}$ denote the set of essential valuations and let $\mathcal{P}$ denote the set of minimal prime ideals of A; it is convenient to have available notation for both bijections $\mathcal{V} \leftrightarrow P$: let them be denoted $v \mapsto P_v$ and $P \mapsto v_P$. Then a twist E^* of E is a **global minimal twist** or an ***A*-minimal twist** if E^* is defined over A and (*cf.* Proposition 5.5.1)

$$\text{for all } v \in \mathcal{V}, \quad v(\Delta(E)) = \delta_v^*(E).$$

Such an E^* is automatically A-minimal.

For example we have the $\mathbf{Z}$-minimal twists (*cf.* Corollary 5.6.4)

$$y^2 = x^3 - x,\ \Delta = 64\ (\mathbf{A32}) \quad \text{and} \quad y^2 + y = x^3,\ \Delta = -27\ (\mathbf{A27})$$

with $j = 1728$ and 0 respectively. Other easy cases are:

- if 2 is invertible in A, hence $\operatorname{char} K \neq 2$, then $y^2 = x^3 - x/4$ has $j = 1728$ and $\Delta = 1$, hence is a global minimal twist, and $\delta_v^*(1728) = 0$ for all v; this includes the case $j = 0$, $\operatorname{char} K = 3$;
- if 3 is invertible in A, hence $\operatorname{char} K \neq 3$, then $y^2 + y = x^3$ has $j = 0$ and $\Delta = -27$, hence is a global minimal twist, and $\delta_v^*(1728) = 0$ for all v.

If an A-minimal twist exists, it is not usually unique (up to A-isomorphism); this was already mentioned above in the local case $A = V$. In the case of $A = \mathbf{Z}$, Proposition 5.7.3 shows that in many cases E and $E * (-1)$ have the same p-signature for all primes p but are not $\mathbf{Q}$-isomorphic. For example,

$$y^2 = x^3 - x^2 + x, \quad c_4 = -32, \quad c_6 = -224, \quad \Delta = -48 \qquad \mathbf{A24}$$

$$y^2 = x^3 + x^2 + x, \quad c_4 = -32, \quad c_6 = 224, \quad \Delta = -48 \qquad \mathbf{A24} * (-1) = \mathbf{A48}$$

are both $\mathbf{Z}$-minimal twists with $j = 2^{11}/3$, but they are not isomorphic over $\mathbf{Q}$.

Apart from this lack of uniqueness, and technical complication caused by non-quadratic twists when $j = 1728$ or 0, we can imitate what was done in Proposition 5.5.1. Define the divisor

$$\mathcal{A}^* = \mathcal{A}_E^* = \mathcal{A}_j^* = \sum_{v \in \mathcal{V}} \frac{1}{6}(v(\Delta) - \delta_v^*(E))P_v.$$

Note the fraction 1/6 instead of 1/12 that occurs in the sum for $\mathcal{A}_E$. If τE is a twist where $\tau = [r, s, t, u]$, and $j \neq 1728$ or 0 to ensure that $u^2 \in K^*$, then

$$\mathcal{A}_{\tau E}^* = \mathcal{A}_E^* - \operatorname{div}(u^2).$$

Thus the property of $\mathcal{A}^*$ being principal is preserved under twisting (when $j \neq 1728$ or 0).

Proposition 5.7.5 *Let A be a Krull domain with quotient field K of characteristic $\neq 2$, and let E be an elliptic curve defined over K with $j \neq 1728$ or 0.*

(i) *If an A-minimal twist of E exists then $\mathcal{A}_E^*$ is principal.*

(ii) *Partial converse: if $\mathcal{A}_E^*$ is principal and $6 \in A^*$, then E has an A-minimal twist.*

(iii) *If A is a Dedekind domain we can remove the extra condition in* (ii): *an A-minimal twist of E exists iff the ideal*

$$\mathcal{A}_E^* = \prod_{v \in \mathcal{V}} P_v^{(v(\Delta)-\delta_v^*)/6}$$

is principal.

Proof. (i) Suppose E^* is an A-minimal twist. Since $j \neq 1728$ or 0, by Proposition 3.3.2, E^* is K-isomorphic with E^d for some $d \in K^*$, say $E^* = [r, s, t, u]E^d$. Then the discriminant of E^* is $u^{-12}d^6\Delta$, hence

$$\mathcal{A}_E^* = \operatorname{div}(u^2 d^{-1})$$

is principal.

(ii) Since 2 and 3 are invertible, we can take the equation of E in c-form $y^2 = x^3 + bx + c$. Then $\mathcal{A}_E^*$ is still principal, say $\mathcal{A}_E^* = \operatorname{div}(u)$. For $v \in \mathcal{V}$, there is a v-minimal twist E_v which we can take in c-form. Then the transformation from E to E_v is necessarily of the form $[0, 0, 0, \sqrt{d_v}]$ for some $d_v \in K^*$, *i.e.*, E_v has the equation $y^2 = x^3 + bd_v^{-2}x + cd_v^{-3}$. It follows that $v(d_v) = v(u)$ and therefore each d_v can be replaced by u. Hence $E * (u^{-1})$ is an A-minimal twist.

The proof of (iii) is similarly an obvious adaptation of the proof of Proposition 5.5.1(iii). ■

Bibliography for Chapters 1–5

[**Ada-Ra80**] W. Adams and M. Razar, Multiples of points on elliptic curves and continued fractions, *Proc. London Math. Soc.* **41**(1980), 481–498.

[**Alt72**] A. Altman, The size function on abelian varieties, *Trans. Amer. Math. Soc.* **164**(1972), 153–161.

[**AntIV**] B. Birch and W. Kuyk (editors), Modular functions of one variable IV, *Lecture notes in Mathematics* **476**, Springer Verlag, 1975.

[**Aya92**] M. Ayad, Points S-entiers des courbes elliptiques, *Manu. Math.* **76**(1992), 305–324.

[$\cdots$] B. Birch and W. Kuyk (editors) — see [AntIV].

[**BACn**] N. Bourbaki, Algèbre commutative, chap. n, Hermann.

[**BAn**] ————, Algèbre, chap. n, Hermann.

[**BGZ85**] J.P. Buhler, B.H. Gross and D.B. Zagier, On the conjecture of Birch and Swinnerton-Dyer for an elliptic curve of rank 3, *Math. Comp.* **44**(1985), 473–481.

[**Bil38**] G. Billing, Beiträge zur arithmetischen Theorie der ebenen Kubischen Kurven vom Geschlecht eins, *Nova Acta Regiae Soc. Sci. Upsaliensis* (4)**11**(1938), No.1, 1–165.

[**Bil-Ma40**] G. Billing and K. Mahler, On exceptional points on cubic curves, *J. London Math. Soc.* **15**(1940), 32–43.

[$\cdots$] B. Birch and W. Kuyk (editors) — see [AntIV].

[**Bir-Sw63**] B. Birch and H. Swinnerton-Dyer, Notes on elliptic curves I, *J. reine angew. Math.* **212**(1963), 7–25.

[**Car-Ei56**] H. Cartan and S. Eilenberg, Homological algebra, Princeton Univ. Press, 1956.

[**Cas49**] J.W.S. Cassels, A note on the division values of $\wp(u)$, Proc. Camb. Phil. Soc. **45**(1949), 167–172.

[**Cas50**] ————, The rational solutions of the diophantine equation $Y^2 = X^3 - D$, *Acta Math.* **82**(1950), 243–273; Addenda and corrigenda, *ibid.* **84**(1951), 299.

[**Cas59**] ————, Intro. to the geometry of numbers, Springer Verlag,1959.

[**Cas66**] ————, Diophantine equations with special reference to elliptic curves, *J. London Math. Soc.* **41**(1966), 193–291.

[**Cas78**] ————, Rational quadratic forms, Academic Press, 1978.

[**Cas91**] ————, Lectures on elliptic curves, London Math. Soc., student texts 24, 1991.

[**Cas-Fr67**] J.W.S. Cassels and A. Frölich (editors), Algebraic number theory, Academic Press, 1967.

[**Cha95**] J.S. Chahal, Manin's proof of the Hasse inequality revisited, *Nieuw Arch. Wiskd.* **13**(1995), 219–232.

[**Coh93**] H. Cohen, A course in computational algebraic number theory, Springer, 1993.

[**Cohn80**] H. Cohn, Diophantine equations over $\mathbf{C}(t)$ and complex multiplication, Number theory Carbondale 1979, *Lecture notes in Mathematics* **751**, Springer Verlag, 1980.

[**Com-Na87**] S. Comalada and E. Nart, Courbes elliptiques avec bonne réduction partout,*C. R. Acad. Sci. Paris Sér. I Math.*, **305**(1987), 223–224.

[**Con66**] I. Connell, Abelian formal groups, Proc. A.M.S. **17**(1966), 958-9.

[**Con68**] ————, A natural transformation of the spec functor, *J. Algebra* **10**(1968), 69–91.

[**Con82**] ————, Modern algebra, North Holland, 1982.

[**Con92**] ————, Addendum to a paper of Harada and Lang, *J. Algebra* **145**(1992), 463–467.

[**Cox89**] D. Cox, Primes of the form $x^2 + ny^2$, John Wiley & Sons, 1989.

[**Cre92**] J.E. Cremona, Algorithms for modular elliptic curves, Cambridge Univ. Press, 1992; revised 2nd edition 1997.

[**Cre93**] ————, the above tables extended to conductors between 1000 and 5077, *anonymous ftp from* euclid.exeter.ac.uk, pub/cremona/. . . , 1993.

[**Dar96**] H. Darmon, Review of [Mer96], *Math. Reviews* **96i**(1996), #11057.

[**Des86**] A. Desboves, Résolution en nombres entiers et sous sa forme la plus générale, de l'équation cubique, homogène à trois inconnues, *Nouv. Ann. de la Math.* **5**(1886), 545–579.

[**Dic52**] L.E. Dickson, History of the theory of numbers, *in 3 volumes,* reprinted by Chelsea, 1952.

[**Dun66**] R.L. Duncan, Some inequalities for polynomials, *Amer. Math. Monthly* **73**(1966), 58–59.

[**Elk89**] N. Elkies, Supersingular primes for elliptic curves over number fields, *Compos. Math.* **72**(1989), 165–172.

[**Fle-Oe90**] M. Flexor and J. Oesterlé, Sur les points de torsion des courbes elliptiques, *Astérisque* **183**(1990), 25–36.

[**Fri22**] R. Fricke, Elliptische Funktionen, 2(?) vols., 1922.

[**Gan60**] F.R. Gantmacher, Matrix theory, reprinted by Chelsea, 1960.

[**Has33**] H. Hasse, Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F.K. Schmidtschen Kongruenzzetafunktionen in gewissen elliptischen Fällen, *Nach. Ges. Wiss. Göttingen, Math.-Phys. Kl.*(1933), 253–262.

[**Har-Wr54**] G.H. Hardy and E.M. Wright, An intro. to the theory of numbers, third edition (there may be later), Oxford, 1954.

[**Haz78**] M. Hazewinkel, Formal groups and applications, Academic Press, 1978.

[**Hei67**] W.J. Heinzer, Some properties of integral closure, *Proc. Amer. Math. Soc.*, **18**(1967), 749–753.

[**Hon68**] T. Honda, Formal groups and zeta functions, *Osaka J. Math.* **5**(1968), 199–213.

[**Hon70**] ———, On the theory of commutative formal groups, *J. Math. Soc. Japan* **22**(1970), 213–246.

[**How93**] E. Howe, On the group orders of elliptic curves over finite fields, *Compos. Math.* **85**(1993), 229–247.

[**Hur17**] A. Hurwitz, Über ternäre diophantische Gleichungen dritten Grades, *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* **62**(1917), 207–229; reproduced in Math. Werke, vol. II, 446-468, Birkhäuser, 1933.

[**Hus87**] D. Husemöller, Elliptic curves, Springer, 1987.

[**Ire-Ro82**] K. Ireland and M. Rosen, A classical introduction to modern number theory, Springer, 1982.

[**Jor13**] C. Jordan, Cours d'analyse de l'École Polytechnique, 3 vols., various editions, 1892–1913.

[**Kna92**] A.W. Knapp, Elliptic curves, Math. Notes 40, Princeton Univ. Press, 1992.

[**Kra89**] A. Kraus, Quelques remarques à propos des invariants c_4, c_6 et Δ d'une courbe elliptique, *Acta Arith.* **54**(1989), 75–80.

[**Lan58**] S. Lang, Intro. to algebraic geometry, Interscience, 1958.

[**Lan66**] ————, Intro. to transcendental numbers, Addison-Wesley, 1966.

[**Lan83**] ————, Fundamentals of diophantine geometry, Springer, 1983. *This supersedes* Diophantine Geometry, Interscience Publ., 1962.

[**Las82**] M. Laska, An algorithm for finding a minimal Weierstrass equation for an elliptic curve, *Math. Comp.* **38**(1982), 257–260.

[**Laz54**] M. Lazard, La non-existence des groupes de Lie formels non abéliens à un paramètre, *Comp. Rend. Acad. Sci. France* **239**(1954), 942–945.

[**Lut37**] E. Lutz, Sur l'équation $y^2 = x^3 = Ax - B$ dans les corps p-adiques, *J. reine angew. Math.* **177**(1937), 238–247.

[**Man69**] Yu. Manin, On cubic congruences to a prime modulus, *Izv. Akad. Nauk. SSSR* **20**(1956), AMS Transl. (2) **13**(1960), 1–7.

[**Maz72**] B. Mazur, Rational points of abelian varieties with values in towers of number fields, *Invent. Math.* **18**(1972), 183–266.

[**Mer96**] L. Merel, Bornes pour la torsion des courbes elliptiques sur les corps de nombres, *Invent. Math.* **124**(1996), 437–449.

[**Mil80**] J.S. Milne, Étale Cohomology, Princeton University Press, 1980.

[**Mol89**] R. Mollin(editor),

[**Mon92**] P. Monsky, Three constructions of rational points on $Y^2 = X^3 \pm NX$, *Math. Zeit.* **209**(1992), 445–462; Errata, *ibid.* **212**(1993), 141.

[**Mor22**] L.J. Mordell, On the rational solutions of the indeterminate equations of the third and fourth degrees, *Proc. Cambridge Philos. Soc.* **21**(1922), 179–192.

[**Mor66**] ————, The infinity of rational solutions of $y^2 = x^3 + k$, *J. London Math. Soc.* **41**(1966), 523–525.

[**Mor67**] ————, The diophantine equation $x^4 + my^4 = z^2$, *Quar. J. Math. Oxford (2)* **18**(1967), 1–6.

[**Mor69**] ————, Diophantine equations, Academic Press, 1969.

[**Nag25**] T. Nagell (spelt the German way: Nagel), Über die rationalen Punkte auf einigen kubischen Kurven, *Tôhoku Math. J.* **24**(1925), 48–53.

[**Nag28**] ————, Sur les propriétés arithmétiques des cubiques planes du premier genre, Acta Math. **52**(1928–9), 93–126.

[**Nag35**] ————, Solution de quelques problèmes dans la théorie arithmétique des cubiques planes du premier genre, *Skrifter Norske Videnskaps-Akademi i Oslo* No.1(1935), 1–25.

[**Nér52**] A. Néron, Problèmes arithmétiques et géométriques rattachés à la notion de rang d'une courbe algébrique dans un corps, *Bull. Math. France* **80**(1952), 101–166.

[**Ogg66**] A. Ogg, Abelian curves of 2-power conductor, *Proc. Camb. Phil. Soc.*, **62**(1966), 143–148.

[**Roq58**] P. Roquette, Einheiten und Divisorenklassen in endlich erzeugbar Körpern, *J. Deutsch. Math. Verein.*, **60**(1958), 1–27.

[**Ros95**] H.E. Rose, On a class of elliptic curves with rank at most two, *Math. Comp.* **64**(1995), 1251–1265.

[**Rüc87**] H.-G. Rück, A note on elliptic curves over finite fields, *Math. Comp.* **49**(1987), 301–304.

[**Sel51**] E.S. Selmer, The diophantine equation $ax^3 + by^3 + cz^3 = 0$, *Acta Math.* **85**(1951), 203–362; continued *ibid* **92**, 191–197.

[**Sel54**] ————, The exceptional points of a cubic curve which is symmetric in the homogeneous variables, *Math. Scand.* **2**(1954), 227–236.

[**Sel54a**] ————, A conjecture concerning rational points on cubic curves, *Math. Scand.* **2**(1954), 49–54.

[**Ser72**] J.-P. Serre, Propriétés galoisiennes des points d'ordre finis des courbes elliptiques, *Invent. Math.* **15**(1972), 259–331.

[**Ser89**] ————, Lectures on the Mordell-Weil theorem, Vieweg & Sohn, 1989.

[**Set78**] B. Setzer, Elliptic curves over complex quadratic fields, *Pacific J. Math.* **74**(1978), 235–250.

[**Sie88**] C.L. Siegel, Lectures on the geometry of numbers, Springer Verlag, 1988.

[**Sik95**] S. Siksek, Infinite descent on elliptic curves, *Rocky Mountain J. Math.* **25**(1995), 1501–1538.

[**Sil84**] J.H. Silverman, Weierstrass equations and the minimal discriminant of an elliptic curve, *Mathematika* **31**(1984), 245–251.

[**Sil86**] ————, The Arithmetic of Elliptic Curves, Springer, 1986.

[**Sil88**] ————, Computing heights on elliptic curves, *Math. Comp.* **51**(1988), 339–358.

[**Sil90**] ————, The difference between the Weil height and the canonical height on elliptic curves, *Math. Comp,* **55**(1990), 723–743.

[**Sil94**] ————, Advanced topics in the arithmetic of elliptic curves, *Graduate texts in Math.*, **151**, Springer Verlag, 1994.

[**Sil-Ta92**] ———— and J. Tate, Rational points on elliptic curves, Springer, 1992.

[**Str83**] R.J. Stroeker, Reduction of elliptic curves over imaginary quadratic number fields, *Pacific J. Math.* **108**(1983), 451–463.

[**Str-Top94**] R.J. Stroeker and J. Topp, On the equation $Y^2 = (X+p)(X^2+p^2)$, *Rocky Mountain J. of Math.*, **24**(1994), 1135–1161.

[**Tat61**] J. Tate, Rational points on elliptic curves, *mimeographed notes of lectures given at Haverford College*, 1961. (Much, but not all, of these lectures is contained in [Sil-Ta92].)

[**Was87**] L. Washington, Class numbers of the simplest cubic fields, *Math. Comp.* **48**(1987), 371–384.

[**Wat69**] W. Waterhouse, Abelian varieties over finite fields, *Ann. Sci. Ecole Normale Sup. (4),* **2**(1969), 521–560.

[**Web08**] H. Weber, Lehrbuch der Algebra, 3 vols., 3rd ed., reprinted by Chelsea.

[**Wei29**] A. Weil, L'Arithmétique sur les courbes algébriques, *Acta Math.* **52**(1929), 281–315.

[**Wei30**] ————, Sur un théorème de Mordell, *Bull. des Sci. Math.* **54**(1930),182–191.

[**Wei67**] ————, Basic number theory, Springer 1967.

[**Zag84**] D.B. Zagier, L-series of elliptic curves, the Birch-Swinnerton-Dyer conjecture, and the class number problem of Gauss, *Notices of the Amer. Math. Soc.* **31** (1984), 739–743.

[**Zar-Sa58**] O. Zariski and P. Samuel, Commutative algebra, vol. I, D. Van Nostrand, 1958.