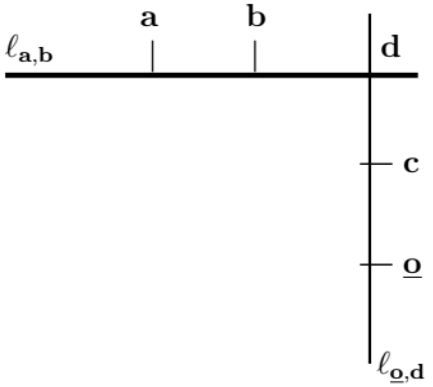
Elliptic Curves.

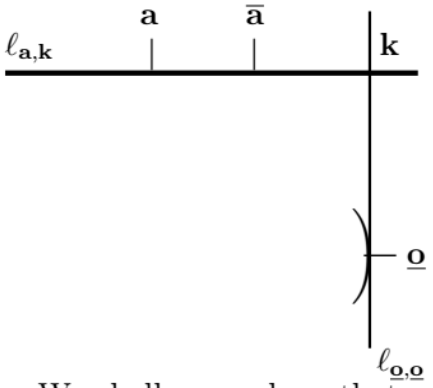
Section 1. The Group Law on an Elliptic Curve

Definition 1.1. An elliptic curve over a field K is (up to birational equivalence) a nonsingular projective cubic curve, defined over K, with a K-rational point.

Definition 1.2. Let $\mathcal{C} : F(X, Y, Z) = 0$ be an elliptic curve $/K$ [the notation $/K$ means 'defined over K'; that is, all of the coefficients of $\mathcal{C}$ are in the field K]. So, $\mathcal{C}$ is a nonsingular projective cubic curve, with a K-rational point, which we shall denote $\underline{\mathbf{o}}$. For any two points $\mathbf{a}, \mathbf{b}$ on $\mathcal{C}$, let $\ell_{\mathbf{a},\mathbf{b}}$ denote the line which meets $\mathcal{C}$ at $\mathbf{a}, \mathbf{b}$ [if $\mathbf{a}, \mathbf{b}$ are distinct then $\ell_{\mathbf{a},\mathbf{b}}$ is the unique line through $\mathbf{a}, \mathbf{b}$; if $\mathbf{a} = \mathbf{b}$ then $\ell_{\mathbf{a},\mathbf{b}}$ is the line tangent to $\mathcal{C}$ at $\mathbf{a} = \mathbf{b}$].



Let $\ell_{\mathbf{a},\mathbf{b}}$ denote the line which meets $\mathcal{C}$ at $\mathbf{a}, \mathbf{b}$.
Then $\ell_{\mathbf{a},\mathbf{b}}$ and $\mathcal{C}$ have 3 points of intersection (Bézout).
Let $\mathbf{d}$ be the third point of intersection between $\mathcal{C}$ and $\ell_{\mathbf{a},\mathbf{b}}$.
Now, let $\ell_{\underline{\mathbf{o}},\mathbf{d}}$ denote the line which meets $\mathcal{C}$ at $\underline{\mathbf{o}}$ and $\mathbf{d}$.
Let $\mathbf{c}$ be the third point of intersection between $\mathcal{C}$ and $\ell_{\underline{\mathbf{o}},\mathbf{d}}$.
Define $\mathbf{a} + \mathbf{b} = \mathbf{c}$.



Let $\ell_{\underline{\mathbf{o}},\underline{\mathbf{o}}}$ be the line tangent to $\mathcal{C}$ at $\underline{\mathbf{o}}$.
Let $\mathbf{k}$ be the third point of intersection between $\mathcal{C}$ and $\ell_{\underline{\mathbf{o}},\underline{\mathbf{o}}}$.
Now, let $\ell_{\mathbf{a},\mathbf{k}}$ be the line which meets $\mathcal{C}$ at $\mathbf{a}$ and $\mathbf{k}$.
Let $\overline{\mathbf{a}}$ be the third point of intersection between $\mathcal{C}$ and $\ell_{\mathbf{a},\mathbf{k}}$.
Define $-\mathbf{a}$ to be $\overline{\mathbf{a}}$.

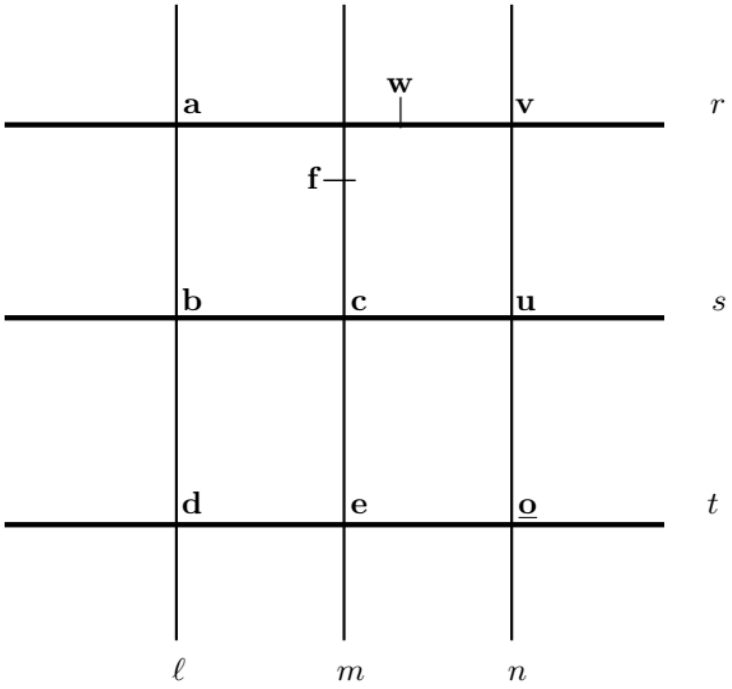
We shall soon show that $\mathbf{a} + \mathbf{b}$ is a commutative group law on the points on $\mathcal{C}$, with identity $\underline{\mathbf{o}}$ and the inverse of $\mathbf{a}$ given by $-\mathbf{a}$. First we need the following technical lemma.

Lemma 1.3. *Let $P_1, \ldots, P_8$ be such that no 4 points lie on a line and no 7 points lie on a conic. Then there exists a unique point P_9 which is a 9th point of intersection of any two cubics passing through $P_1, \ldots, P_8$.*

Optional Proof See 0.137.

Theorem 1.4. *Let $\mathcal{C}$ be an elliptic curve $/K$, with K-rational point $\underline{\mathbf{o}}$. Then $\mathbf{a}+\mathbf{b}$, as in Definition 1.2, gives a commutative group law on the points on $\mathcal{C}$, with identity $\underline{\mathbf{o}}$. The inverse of $\mathbf{a}$ is given by the point $-\mathbf{a}$, constructed in in Definition 1.2. Further, the K-rational points $\mathcal{C}(K)$ form a subgroup, called the Mordell-Weil group.*

Proof It is easy to show commutativity, the fact that $\underline{\mathbf{o}}$ is the identity, and the fact that $-\mathbf{a}$ is the inverse of $\mathbf{a}$. The only difficult problem is associativity. In order to prove associativity, consider the following diagram.



Here, r, s, t, ℓ, m, n are lines. On each line, the labelled points are the points of intersection between $\mathcal{C}$ and that line. From the construction of Definition 1.2:

$$\mathbf{a}+\mathbf{b}=\mathbf{e},$$

and so:

$$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\text{3rd point of intersection on } \ell_{\underline{\mathbf{o}},\mathbf{f}}.$$

Similarly:

$$\mathbf{b}+\mathbf{c}=\mathbf{v},$$

$$\mathbf{a}+(\mathbf{b}+\mathbf{c})=\text{3rd point of intersection on } \ell_{\underline{\mathbf{o}},\mathbf{w}}.$$

To show $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$, it is sufficient to show that $\mathbf{f}=\mathbf{w}$. Let $F_1=\ell mn$ and $F_2=rst$, both of which are cubic curves.

$\mathcal{C}$ and F_1 have 8 common points: $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{e},\mathbf{u},\mathbf{v},\underline{\mathbf{o}}$.

$\mathcal{C}$ and F_2 also have these 8 common points: $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{e},\mathbf{u},\mathbf{v},\underline{\mathbf{o}}$.

From Lemma 1.3, the 9th point of intersection of $\mathcal{C}$ and F_1 must be the same as the 9th point of intersection of $\mathcal{C}$ and F_2; that is, $\mathbf{f}=\mathbf{w}$, as required.

Hence, + is a commutative group law.

It remains to show that $\mathcal{C}(K)$ is a subgroup. We are given that $\underline{\mathbf{o}} \in \mathcal{C}(K)$. Let $\mathbf{a}, \mathbf{b} \in \mathcal{C}(K)$. It is sufficient to show that $\mathbf{a} + \mathbf{b} \in \mathcal{C}(K)$ and that $-\mathbf{a} \in \mathcal{C}(K)$.

Let $\mathbf{a} = (x_1, y_1)$ and $\mathbf{b} = (x_2, y_2)$, where $x_1, y_1, x_2, y_2 \in K$. Then the line through $\mathbf{a}, \mathbf{b}$ is (in affine form) $\ell_{\mathbf{a},\mathbf{b}} : y = \ell x + m$, where $\ell = \frac{y_1 - y_2}{x_1 - x_2} \in K$ and $m = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2} \in K$. Substitute $y = \ell x + m$ into the cubic equation for $\mathcal{C}$ to get; $\phi(x) = x^3 + c_2 x^2 + c_1 x + c_0 = 0$, defined over K. Let $\phi(x) = (x - x_1)(x - x_2)(x - x_3)$ be the factorisation of $\phi(x)$. Then x_1, x_2, x_3 are the 3 roots of ϕ and so $x_1 + x_2 + x_3 = -c_2$, giving: $x_3 = -c_2 - x_1 - x_2 \in K$ and $y_3 = \ell x_3 + m \in K$. The line $\ell_{\mathbf{a},\mathbf{b}}$ then meets $\mathcal{C}$ at $\mathbf{a}, \mathbf{b}, \mathbf{d} = (x_3, y_3) \in \mathcal{C}(K)$. The same argument shows that the line $\ell_{\underline{\mathbf{o}},d}$ through $\underline{\mathbf{o}}, \mathbf{d}$ has 3rd point of intersection $\mathbf{c}$ which is also in $\mathcal{C}(K)$. But $\mathbf{c} = \mathbf{a} + \mathbf{b}$ and so we have shown that $\mathbf{a} + \mathbf{b} \in \mathcal{C}(K)$. A similar argument shows that if $\mathbf{a} \in \mathcal{C}(K)$ then $-\mathbf{a} \in \mathcal{C}(K)$. Hence $\mathcal{C}(K)$ is a subgroup, as required. □

Aside: It is apparent that, in the above proof, we have dealt with the 'typical' case, where none of our points are repeated (for the proof of associativity), and none are at infinity (for the proof that $\mathcal{C}(K)$ is a subgroup, since the points were written in affine form). It is straightforward to check these special cases; we shall not bother to do so here.

Comment 1.5. When two nonsingular cubics $\mathcal{C}_1, \mathcal{C}_2$ are birationally equivalent over K (under $\phi : \mathcal{C}_1 \longrightarrow \mathcal{C}_2$), it can be shown that $\mathbf{a}, \mathbf{b}, \mathbf{c}$ on $\mathcal{C}_1$ are collinear iff $\phi(\mathbf{a}), \phi(\mathbf{b}), \phi(\mathbf{c})$ on $\mathcal{C}_2$ are collinear, and ϕ is an isomorphism between $\mathcal{C}_1(K)$ and $\mathcal{C}_2(K)$.

Comment 1.6. By an elliptic curve, we shall always mean a projective curve, but often write the equation in affine form. Note that, whichever way it is written, we are always referring to the projective curve. For example, if we say 'let $\mathcal{C} : y^2 = x^3 + 3$ be an elliptic curve', it should be understood that this is a shorthand notation for the corresponding projective curve $ZY^2 = X^3 + 3Z^3$.

Theorem 1.7. *Let K be a field satisfying char$(K) \neq 2, 3$ [recall – this means that $1 + 1 \neq 0$ and $1 + 1 + 1 \neq 0$]. Then any elliptic curve over K is birationally equivalent over K to a curve of the form $y^2 = x^3 + Ax + B$.*

When $K = \mathbb{Q}$, we can birationally transform any $y^2 =$ cubic in x to a curve of the form $y^2 = x^3 + Ax + B$, with $A, B \in \mathbb{Z}$, using only maps of the form $(x, y) \mapsto (ax + b, cy)$.

Comment 1.8. Let K be a field satisfying char$(K) \neq 2, 3$, and let $g(x)$ be a quartic polynomial over K with nonzero discriminant. It can be shown that any curve $\mathcal{D} : y^2 = g(x)$, with a K-rational point, is an elliptic curve, and is birationally equivalent over K to a curve of the form $y^2 = x^3 + Ax + B$ [see p.35 of Cassels].

Comment 1.9. We shall typically take our elliptic curves to have the form

$$\mathcal{E} : y^2 = x^3 + Ax + B, \text{ where } A, B \in K,$$

which should be regarded as shorthand for the projective curve $ZY^2 = X^3 + AXZ^2 + BZ^3$. Sometimes it will be convenient to include the x^2 term. Since $\mathcal{E}$ is nonsingular, we must have $\Delta = 4A^3 + 27B^2 \neq 0$, as was shown in Example 0.110. The notation $\Delta = 4A^3 + 27B^2$ is standard.

It is conventional to choose $\underline{\mathbf{o}} = (0, 1, 0)$, the point at infinity, as the identity [we shall always take $\underline{\mathbf{o}} = (0, 1, 0)$ unless otherwise stated]. Note that the line $Z = 0$ meets $\mathcal{E}$ at $\underline{\mathbf{o}}$ three times (such a point is called an *inflexion*). Given a point $\mathbf{a} = (X, Y, Z)$, if we take the line through $\mathbf{a}$ and $\underline{\mathbf{o}} = (0, 1, 0)$ then the third point of intersection is $(X, -Y, Z)$, which must then be $-\mathbf{a}$. In affine form: $-(x, y) = (x, -y)$.

This gives an easy rule for finding the inverse of a point, under the group law, namely: the inverse of $\mathbf{a}$ is its reflection in the x-axis.

So, for an elliptic curve $\mathcal{E}$ written in the form $y^2 =$ cubic in x, the points are $\underline{\mathbf{o}}$ (the point at infinity) and the affine points (x, y), and the group law has a simpler description:

Let $\mathbf{d} = (x_3, y_3)$ the 3rd point of intersection of $\mathcal{E}$ and $\ell_{\mathbf{a},\mathbf{b}}$.

Then $\mathbf{a} + \mathbf{b} = (x_3, -y_3)$, the reflection of $\mathbf{d}$ in the x-axis.

We illustrate the group law with the following computation (see also 0.143).

Example 1.10. Let $\mathcal{E} : y^2 = x^3 + 1$. Let us compute $\mathbf{a} + \mathbf{b}$, where $\mathbf{a} = (x_1, y_1) = (-1, 0)$ and $\mathbf{b} = (x_2, y_2) = (0, 1)$.

The line through $\mathbf{a}, \mathbf{b}$ is $\ell_{\mathbf{a},\mathbf{b}} : y = x + 1$. Substituting this into $\mathcal{E}$, we see that the x-coordinate of any point of intersection satisfies: $(x + 1)^2 = x^3 + 1$, and so:

$$x^3 - x^2 - 2x = 0. \qquad (*)$$

We are looking for (x_3, y_3), the 3rd point of intersection of $\mathcal{E}$ and $\ell_{\mathbf{a},\mathbf{b}}$. We first find x_3; note that x_1, x_2, x_3 must be the roots of $(*)$.

Method A (for finding x_3). Since the roots of $(*)$ are x_1, x_2, x_3, it follows that $x^3 - x^2 - 2x = (x - x_1)(x - x_2)(x - x_3)$; equating coefficients of x^2 gives that:

$$x_1 + x_2 + x_3 = -(\text{coefficient of } x^2 \text{ in } (*)) = -(-1) = 1,$$

so that $(-1) + 0 + x_3 = 1$, giving $x_3 = 2$.

Method B (for finding x_3). Factorise $(*)$ to give: $x(x+1)(x-2)$, whose roots are: $0, -1, 2$. Two of these are the already known $x_1 = -1, x_2 = 0$, and so x_3 must be the remaining root: $x_3 = 2$.

Having found x_3 (by either method), we use the equation of $\ell_{\mathbf{a},\mathbf{b}}$ to compute $y_3 = x_3 + 1 = 3$. In summary: $\mathcal{E}$ and $\ell_{\mathbf{a},\mathbf{b}}$ intersect at: $(-1, 0), (0, 1), (2, 3)$, and so $(-1, 0) + (0, 1) + (2, 3) = \underline{\mathbf{o}}$.

Finally, this gives: $(-1,0)+(0,1) = -(2,3) = (2,-3)$, using the rule that negation is given by reflection in the x-axis.

One can also obtain an explicit general formula for the group law.

Lemma 1.11. *Let $\mathcal{E}: y^2 = x^3 + Ax + B$, where $A, B \in K$, with (as usual) $\underline{\mathbf{o}}$ = the point at infinity. Let $(x_3, y_3) = (x_1, y_1) + (x_2, y_2)$.*

Case 1. *When $x_1 \neq x_2$ then:*

$$x_3 = \frac{x_1x_2^2 + x_1^2x_2 + A(x_1+x_2) + 2B - 2y_1y_2}{(x_1-x_2)^2}, \quad y_3 = -\ell x_3 - m,$$

$$\text{where: } \ell = \frac{y_1-y_2}{x_1-x_2}, \quad m = \frac{x_1y_2 - x_2y_1}{x_1-x_2}.$$

Case 2. *When $(x_1, y_1) = (x_2, y_2)$ then $(x_3, y_3) = (x_1, y_1) + (x_1, y_1)$ [which can be written as $2(x_1, y_1)$], and:*

$$x_3 = \frac{x_1^4 - 2Ax_1^2 - 8Bx_1 + A^2}{4y_1^2} = \frac{x_1^4 - 2Ax_1^2 - 8Bx_1 + A^2}{4(x_1^3 + Ax_1 + B)}, \quad y_3 = -\ell x_3 - m,$$

$$\text{where: } \ell = \frac{3x_1^2 + A}{2y_1}, \quad m = \frac{-x_1^3 + Ax_1 + 2B}{2y_1}.$$

Optional Proof See 0.144.

The above formulas give an alternative method for computing the group law, although in practice it often turns out to be easier to compute the group law from first principles, as in Example 1.10.

Comment 1.12. When $\Delta = 4A^3 + 27B^2 \neq 0$, all 3 roots of $x^3 + Ax + B$ are distinct, guaranteeing that $y^2 = x^3 + Ax + B$ has no singularities and is an elliptic curve.

When $\Delta = 0$, then this is no longer an elliptic curve and at least two roots of the cubic are repeated: $y^2 = (x-\alpha)^2(x-\beta)$. It is still the case that the set of nonsingular points on $\mathcal{E}$, denoted $\mathcal{E}_{ns}$, forms a group [see pp.39–41 of Cassels]. When $\beta \neq \alpha$ the singularity at $(\alpha, 0)$ is a node. When $\beta = \alpha$ the singularity is a cusp. In either case, the curve can be written: $\left(\frac{y}{x-\alpha}\right)^2 = x - \beta$, and so is birationally equivalent to the conic $w^2 = x - \beta$.

Definition 1.13. Let $\mathcal{E}$ be an elliptic curve and let P be a point on $\mathcal{E}$. For any positive integer m, let mP denote $P + \ldots + P$ [m times]. We say that P is an *m-torsion* point if $mP = \underline{\mathbf{o}}$. The m-torsion group of $\mathcal{E}$, denoted $\mathcal{E}[m]$, is the set of all m-torsion points. We also say that P has *order m* (or that P is a *point of order m*) if m is the smallest positive integer for which $mP = \underline{\mathbf{o}}$. When such m exists, P is a *torsion point* (P *has finite order*). If no such m exists, then P is a non-torsion point (P has *infinite order*). The group of all K-rational torsion points on $\mathcal{E}$ is denoted $\mathcal{E}_{\text{tors}}(K)$ [or sometimes $\mathcal{E}(K)_{\text{tors}}$].

Examples 1.14.
(a) Let $\mathcal{E} : y^2 = x^3 - x$, and let $P = (1, 0)$ so that $-P = (1, -0) = (1, 0) = P$, so that $2P = P + P = P - P = \underline{\mathbf{o}}$. But $1 \cdot P = P \neq \underline{\mathbf{o}}$, and so 2 is the smallest $m > 0$ such that $mP = \underline{\mathbf{o}}$. P has order 2 and $P \in \mathcal{E}_{\text{tors}}(\mathbb{Q})$.
(b) Let $\mathcal{E} : y^2 = x^3 + 1$, and let $P = (0, 1)$. First compute $P + P$. Using $2yy' = 3x^2$ at $(0, 1)$ gives $2 \cdot 1 \cdot y' = 3 \cdot 0^2$ and so the tangent line $\ell_{P,P}$ to $\mathcal{E}$ at P has slope 0 and equation of form $y = 0 \cdot x + m$. But the line goes through $(0, 1)$ and so $m = 1$ and the tangent line is $y = 1$. Substituting $y = 1$ into $y^2 = x^3 + 1$ gives $x^3 = 0$, with roots $0, 0, 0$. So, $\mathcal{E}$ meets $\ell_{P,P}$ at $(0, 1)$ with multiplicity 3, and $(0, 1) + (0, 1) + (0, 1) = \underline{\mathbf{o}}$. Hence: $(0, 1) + (0, 1) = -(0, 1) = (0, -1)$. In summary:

$1 \cdot (0, 1) = (0, 1), \quad 2 \cdot (0, 1) = (0, -1), \quad 3 \cdot (0, 1) = \underline{\mathbf{o}}.$

$(0, 1)$ has order 3 and $(0, 1) \in \mathcal{E}_{\text{tors}}(\mathbb{Q})$.

When $K = \mathbb{F}_p$, a finite field with p elements, there are of course only finitely many members of $\mathcal{E}(\mathbb{F}_p)$.

Aside: Each of the p possible x-coordinates $0, \ldots, p-1$ *has about a 50% chance of making* $x^3 + Ax + B$ *a square modulo p. When* $x^3 + Ax + B$ *is not a square, there are no corresponding y-coordinates. When* $x^3 + Ax + B$ *is a square, there are at most two corresponding y-coordinates. So, one might expect 'on average' about p affine points, that is, about* $p + 1$ *points, including the point at infinity.*

The following result gives a bound within which the number of points must lie.

Theorem 1.15. *(Hasse). Let* $\mathcal{E}$ *be an elliptic curve over* $\mathbb{F}_p$. *Let* $N_p = \#\mathcal{E}(\mathbb{F}_p)$ *where, as usual,* $\mathcal{E}(\mathbb{F}_p)$ *should be taken to including* $\underline{\mathbf{o}}$ [*so that* N_p *is the number of affine points* (x, y) *on* $\mathcal{E}$ *with* $x, y \in \mathbb{F}_p$, *plus* 1, *to include the point at infinity* $\underline{\mathbf{o}}$]. *Then:*

$$|N_p - (p + 1)| \leqslant 2\sqrt{p}, \text{ that is, } N_p \in [(p + 1) - 2\sqrt{p}, (p + 1) + 2\sqrt{p}].$$

Similarly, any curve $y^2 = Q(x)$, *where* $Q(x) = f_4x^4 + \ldots + f_0$ *has nonzero discriminant, has at least* $p - 1 - 2\sqrt{p}$ *affine points.*

Proof See p.118 of Cassels or p.131 of Silverman. □

Example 1.16. Let $\mathcal{E} : y^2 = x^3 + 4x + 1$, defined over $\mathbb{F}_{13}$. Then:

$$\#\mathcal{E}(\mathbb{F}_{13}) \geqslant 13 + 1 - 2\sqrt{13} > 13 + 1 - 2 \cdot 4 = 6, \text{ so that } \#\mathcal{E}(\mathbb{F}_{13}) \geqslant 7.$$

$$\#\mathcal{E}(\mathbb{F}_{13}) \leqslant 13 + 1 + 2\sqrt{13} < 13 + 1 + 2 \cdot 4 = 22, \text{ so that } \#\mathcal{E}(\mathbb{F}_{13}) \leqslant 21.$$

Note that at most 4 of the points on $\mathcal{E}(\mathbb{F}_{13})$ can be $\underline{\mathbf{o}}$ and points of the form $(x, 0)$, so there must exist at least 3 affine points $(x, y) \in \mathcal{E}(\mathbb{F}_{13})$ with $y \neq 0$.

Section 2. The p-adic Numbers $\mathbb{Q}_p$

For $\mathbb{Q}$, let $|\ |_\infty$ denote the standard absolute value [e.g. $|-5|_\infty = |5|_\infty = 5$]. Consider the sequence: $x_1 = 1.4, x_2 = 1.41, x_3 = 1.414, \ldots$, where x_n is the largest decimal to n decimal places satisfying $x_n^2 < 2$. Then $|x_m - x_n|_\infty \to 0$ as $m, n \to \infty$, so that the sequence is Cauchy in $\mathbb{Q}, |\ |_\infty$. The sequence x_n cannot be convergent, since if $x_n \to \alpha$ then clearly $\alpha^2 = 2$ and no such α exists in $\mathbb{Q}$. We say that $\mathbb{Q}, |\ |_\infty$ is *incomplete* (since not every Cauchy sequence is convergent) and the real numbers $\mathbb{R}$ give the *completion* of $\mathbb{Q}, |\ |_\infty$. The absolute value $|\ |_\infty$ is a special case of the following.

Definition 2.1. Let K be a field. A *valuation* on K is a function $|\ | : K \to \mathbb{R}$ satisfying:

(1) $|x| \geqslant 0$ for all $x \in K$, with equality if and only if $x = 0$.

(2) $|xy| = |x| \cdot |y|$ for all $x, y \in K$.

(3) $|x + y| \leqslant |x| + |y|$ for all $x, y \in K$ [the *triangle inequality*].

If a valuation also satisfies the stronger property:

(3)′ $|x + y| \leqslant \max(|x|, |y|)$, for all $x, y \in K$,

then we say that it is a *non-Archimedean valuation*; otherwise it is an *Archimedean valuation*.

For example, $\mathbb{Q}, |\ |_\infty$ (or $\mathbb{R}, |\ |_\infty$) is a valuation. It is Archimedean since, for example, $|1 + 1|_\infty \not\leqslant \max(|1|_\infty, |1|_\infty)$. We shall now introduce another valuation on $\mathbb{Q}$, which gives a different notion of size and distance.

Definition 2.2. Fix a prime p. Let $x = \frac{m}{n} \in \mathbb{Q}$. Write $\frac{m}{n} = p^r \frac{a}{b}$, where $p \not| \ a, p \not| \ b$. Then the *p-adic valuation* (or *p-adic absolute value* or *p-adic size*) is defined to be:

$|x|_p = |\frac{m}{n}|_p = p^{-r}$ [so, x is 'smaller' the higher the power of p dividing x].

We also define $|0|_p = 0$. For any $x, y \in \mathbb{Q}$, the *p-adic distance* between x and y is defined to be: $d_p(x, y) = |x - y|_p$. (Note that d_p is a metric)

Example 2.3. In $\mathbb{Q}, |\ |_3$, we have: $|\frac{4}{3}|_3 = |3^{-1}\frac{4}{1}|_3 = (3^{-(-1)}) = 3$, $|9|_3 = |3^2\frac{1}{1}|_3 = 3^{-2} = \frac{1}{9}$, and $|7|_3 = |3^0\frac{7}{1}|_3 = 3^{-0} = 1$.

Also, $d_3(-5, 3) = |-5 - 3|_3 = |-8|_3 = 1$, $d_3(-5, 19) = |-5 - 19|_3 = |-24|_3 = 3^{-1}$ and $d_3(\frac{1}{2}, \frac{1}{5}) = |\frac{3}{10}|_3 = 3^{-1}$.

For integers m, n, $m \not\equiv n \pmod 3 \iff d_3(m, n) = 1$, $m \equiv n \pmod 3 \iff d_3(m, n) \leqslant \frac{1}{3}$, $m \equiv n \pmod{3^2} \iff d_3(m, n) \leqslant \frac{1}{3^2}$, and so on. The integers m, n are 3-adically closer when they are congruent modulo a higher power of 3.

Lemma 2.4. *The function $|\ |_p$ of Definition 2.2 is a non-Archimedean valuation on $\mathbb{Q}$.*

Proof $(1), (2), (3)'$ are trivially true when x or $y = 0$. Let $x, y \in \mathbb{Q}$, $x, y \neq 0$, and write $x = p^r \frac{a}{b}, y = p^s \frac{c}{d}$, where $p \not| \ a, b, c, d$.

(1) $|x|_p = p^{-r} > 0$.

(2) $|xy|_p = |p^r \frac{a}{b} \ p^s \frac{c}{d}|_p = |p^{r+s} \ \frac{ac}{bd}|_p = p^{-(r+s)}$ [since $p \not| \ ac, bd$] $= p^{-r}p^{-s} = |x|_p|y|_p$.

(3)$'$ Wlog $r \leqslant s$, giving: $|x+y|_p = |p^r \frac{a}{b} + p^s \frac{c}{d}|_p = |p^r \left(\frac{a}{b} + p^{s-r} \frac{c}{d}\right)|_p = |p^r \frac{ad+p^{s-r}bc}{bd}|_p$
$= |p^r \frac{p^k \ell}{bd}|_p$ for some $k \geqslant 0$ and $\ell \in \mathbb{Z}$ with $p \not| \ \ell$ [since $ad + p^{s-r}bc \in \mathbb{Z}$]
$= p^{-(r+k)} \leqslant p^{-r} = |x|_p = \max(|x|_p, |y|_p)$. □

Comment 2.5. By induction, $|a_1 + \ldots + a_n|_p \leqslant \max(|a_1|_p, \ldots, |a_n|_p)$. It is also easy to show that $|x|_p \neq |y|_p \implies |x+y|_p = \max(|x|_p, |y|_p)$. Furthermore, if $|a_k|_p > |a_i|_p$ for all i, $1 \leqslant i \leqslant n, i \neq k$, then $|a_1 + \ldots + a_n|_p = \max(|a_1|_p, \ldots, |a_n|_p) = |a_k|_p$.

Definition 2.6. Let $K, |\ |$ be a field with valuation. For $a_n, \ell \in K$, we say that the sequence a_n *converges* to ℓ [denoted $a_n \to \ell$] in $K, |\ |$ when $|a_n - \ell| \to 0$ in $\mathbb{R}, |\ |_\infty$ as $n \to \infty$. That is: for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that, $|a_n - \ell| < \epsilon$ for all $n > N$. Given a sequence $a_n \in K$, if there exists $\ell \in K$ such that $a_n \to \ell$ in $K, |\ |$ then we say that a_n *converges* in $K, |\ |$, or that it is *convergent* in $K, |\ |$. It is *Cauchy* if $|a_m - a_n| \to 0$ in $\mathbb{R}, |\ |_\infty$ as $m, n \to \infty$. That is: for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that, $|a_m - a_n| < \epsilon$ for all $m, n > N$.

We say that $K, |\ |$ is *complete* if every Cauchy sequence is convergent.

Examples 2.7.

(a) Let $a_n = 6^n$. Then $|a_n - 0|_3 = |6^n|_3 = 3^{-n} \to 0$ as $n \to \infty$. So $a_n \to 0$ in $\mathbb{Q}, |\ |_3$.

(b) Let $a_1 = 1, \ a_2 = 11, \ a_3 = 111, \ldots$ so that $9a_n = 999\ldots9$ [n times] and $9a_n + 1 = 10^n$. Then $|9a_n - (-1)|_5 = |10^n|_5 = 5^{-n} \to 0$, giving $9a_n \to -1$ in $\mathbb{Q}, |\ |_5$. It follows that $a_n \to -\frac{1}{9}$ in $\mathbb{Q}, |\ |_5$.

(c) Let $x_0 = a_0 = 3$. Then $a_0^2 = 9 \equiv 2 (\text{mod } 7)$, and $|x_0^2 - 2|_7 = |a_0^2 - 2|_7 = |7|_7 = 7^{-1} < 1$. We want to find $a_1 \in \{0, \ldots, 6\}$ such that $(a_0 + a_1 7)^2 \equiv 2 \ (\text{mod } 7^2)$.

This is satisfied $\iff a_0^2 + 2a_0a_1 7 + a_1^2 7^2 \equiv 2 \ (\text{mod } 7^2)$

$\iff 6a_1 7 \equiv 2 - 9 = -7 \ (\text{mod } 7^2) \iff 6a_1 \equiv -1 \ (\text{mod } 7) \iff a_1 \equiv 1 \ (\text{mod } 7)$,

so we can take $a_1 = 1$. Let $x_1 = a_0 + a_1 7 = 3 + 1 \times 7 = 10$. Then $x_1^2 = 100 \equiv 2 \ (\text{mod } 7^2)$ and $|x_1^2 - 2|_7 = 7^{-2}$.

Aside: note how the solvability of the last congruence is affected by $|2a_0|_7 = |f'(a_0)|_7$, where $f(x) = x^2 - 2$.

When we similarly solve for $a_2 \in \{0, \ldots, 6\}$ such that $(a_0 + a_1 7 + a_2 7^2)^2 \equiv 2 \pmod{7^3}$ we find that $a_2 = 2$, giving $x_2 = a_0 + a_1 7 + a_2 7^2 = 3 + 7 + 98 = 108$. Check: $x_2^2 \equiv 2 \pmod{7^3}$ and $|x_2^2 - 2|_7 \leqslant 7^{-3}$.

We can inductively find $x_n = a_0 + a_1 7 + \ldots + a_n 7^n$ such that $x_n^2 \equiv 2 \pmod{7^{n+1}}$, that is, $|x_n^2 - 2|_7 \leqslant 7^{-(n+1)}$. Hence $x_n^2 \to 2$ in $\mathbb{Q}, |\ |_7$.

Intuitively, $(3 + 1 \cdot 7 + 2 \cdot 7^2 + \ldots)^2 = 2$ in $|\ |_7$. The sequence x_n is easily seen to be Cauchy in $\mathbb{Q}, |\ |_7$. The sequence is not convergent since if $x_n \to \alpha$ in $\mathbb{Q}, |\ |_7$ then $\alpha^2 = 2$, which is impossible for $\alpha \in \mathbb{Q}$.

(d) Again, let $a_0 = 3$, but now define $a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$, for $n \geqslant 0$, where $f(x) = x^2 - 2$ [the Newton-Raphson formula]. Then:

$a_0 = 3,\ a_1 = 3 - \frac{3^2-2}{2 \cdot 3} = \frac{11}{6},\ a_2 = \frac{11}{6} - \frac{(\frac{11}{6})^2 - 2}{2\frac{11}{6}} = \frac{193}{132}$, and so on.

Check that: $|a_0^2 - 2|_7 = |3^2 - 2|_7 \leqslant 7^{-1}$, $|a_1^2 - 2|_7 = |(\frac{11}{6})^2 - 2|_7 = |\frac{49}{36}|_7 \leqslant 7^{-2}$, and that a_n satisfies the same properties as x_n of Example (c), namely: $|a_n^2 - 2|_7 \leqslant 7^{-(n+1)}$ so that $a_n^2 \to 2$ in $\mathbb{Q}, |\ |_7$, again forcing a_n to be Cauchy but not convergent.

The last two examples show that $\mathbb{Q}$ is incomplete with respect to the valuation $|\ |_7$, and indeed $\mathbb{Q}$ is incomplete with respect to any $|\ |_p$. We now define an extension of $\mathbb{Q}$ which performs the same role with respect to $|\ |_p$ that $\mathbb{R}$ performs with respect to $|\ |_\infty$.

Definition 2.8. The set of *p-adic numbers* $\mathbb{Q}_p$ is the completion of $\mathbb{Q}$ with respect to the valuation $|\ |_p$, and is the smallest field containing $\mathbb{Q}$ which is complete with respect to $|\ |_p$. For any $\alpha, \beta \in \mathbb{Q}_p$, we say that $\alpha \equiv \beta \pmod{p^n} \iff |\alpha - \beta|_p \leqslant p^{-n}$ ['α is congruent to β modulo p^n']. A member of $\mathbb{Q}_p$ (a *p-adic number*) x can be written in following form (the *p-adic expansion* of x):

$$x = \sum_{n=N}^{\infty} a_n p^n, \text{ where } N \in \mathbb{Z}, a_N \neq 0 \text{ and each } a_n \in \{0, \ldots, p-1\},$$

in which case $|x|_p = p^{-N}$, and the a_n are the *digits* of x. We normally use the shorthand notation $a_N \ldots a_0, a_1 a_2 \ldots$ to represent the above sum. Note that $x \in \mathbb{Q}$ exactly when the digits are eventually periodic.

Examples 2.9.

(a) $w = 4 \cdot 5^{-2} + 1 \cdot 5^{-1} + 4 \cdot 5^0 + 1 \cdot 5^1 + 4 \cdot 5^2 + \ldots \in \mathbb{Q}_5$ and $|w|_5 = 5^2$. This can be denoted $414, \overline{14}$.

(b) $\alpha = 3 \cdot 7^0 + 1 \cdot 7^1 + 2 \cdot 7^2 + \ldots \in \mathbb{Q}_7$ from Example 2.7(c) satisfies $\alpha^2 = 2$.

On the other hand, there is no $\beta \in \mathbb{Q}_7$ such that $\beta^2 = 3$ since any such β would satisfy $|\beta|_7^2 = |\beta^2|_7 = |3|_7 = 1$ and so would have 7-adic expansion $\beta = b_0 + b_1 7 + b_2 7^2 + \ldots$ and would

satisfy $(b_0 + b_1 7 + b_2 7^2 + \ldots)^2 = 3$. This would give: $b_0^2 \equiv 3 \pmod 7$, which is impossible, since 3 is not a quadratic residue mod 7 [none of $0^2, 1^2, 2^2, 3^2, 4^2, 5^2, 6^2$ are $\equiv 3 \pmod 7$].
(c) In $\mathbb{Q}_5$: $27 = 2 + 5^2 = 2 \cdot 5^0 + 0 \cdot 5^1 + 1 \cdot 5^2 = 2,01$ [the 5-adic expansion of 27].
(d) Let us find the 5-adic expansion of $-1/4$. We have $|-1/4|_5 = 1$ so that the 5-adic expansion of $-1/4$ must be of the form $\alpha = a_0 + a_1 5 + a_2 5^2 + \ldots$, each $a_i \in \{0, 1, 2, 3, 4\}$ and $a_0 \neq 0$. This satisfies $-1 = 4(a_0 + a_1 5 + a_2 5^2 + \ldots)$ which gives $-1 \equiv 4a_0 \pmod 5$ and so $a_0 = 1$. Then $-1 = 4(1 + a_1 5 + a_2 5^2 + \ldots)$ gives $-5 \equiv 4a_1 5 \pmod{5^2}$, giving $-1 \equiv 4a_1 \pmod 5$, and so $a_1 = 1$. Similarly, we find that $a_2 = 1, a_3 = 1, \ldots$ and we suspect that $-1/4 = 1, \overline{1}$.

Let $\alpha = 1, \overline{1}$. Then $\alpha - 1 = 0, \overline{1} = 5\alpha$, so that $4\alpha = -1$, giving $\alpha = -1/4$, proving that we have the correct 5-adic expansion.

Comment 2.10. The field $\mathbb{Q}$ is often referred to as a *global field* and its completions with respect to valuations, namely $\mathbb{R}$ and $\mathbb{Q}_p$, for any prime p, are its *local fields* (or *localisations*). An equation defined over $\mathbb{Q}$ which has points in $\mathbb{R}$ and every $\mathbb{Q}_p$, but not in $\mathbb{Q}$, is said to *violate the Hasse Principle.*

Definition 2.11. Let K be a field with a non-Archimedean valuation $|\ |$. We say that $x \in K$ is an *integer* (with respect to the valuation) when $|x| \leqslant 1$, and $R = \{x \in K : |x| \leqslant 1\}$ is the *ring of integers* (or *valuation ring*) of K. The set $\mathcal{M} = \{x \in K : |x| < 1\}$ is the *maximal ideal*, and $k = R/\mathcal{M}$ is the *residue field* [also called the *field of digits*]. The *valuation group* is the set $G_K = \{|x| : x \in K^*\}$ under multiplication. We say that the valuation is *discrete* if there exists $\delta > 0$ such that $1 - \delta < |x| < 1 + \delta \implies |x| = 1$. When the valuation is discrete, there exists an element $\mathfrak{p} \in \mathcal{M}$ such that $\mathcal{M} = \mathfrak{p}R$; we say that such an element is a *prime element* for the valuation.

The ring of integers for $\mathbb{Q}_p$ is often denoted $\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leqslant 1\}$. The valuation group $G_{\mathbb{Q}_p} = \{p^r : r \in \mathbb{Z}\} = \{\ldots, p^{-2}, p^{-1}, p^0, p^1, p^2, \ldots\}$, so that $\mathbb{Q}_p$ is discrete, and we can take p as a prime element (or indeed any element with valuation p^{-1}). The maximal ideal is $\mathcal{M} = p\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leqslant p^{-1}\}$ and the residue field $\mathbb{Z}_p/p\mathbb{Z}_p$ is isomorphic to $\mathbb{F}_p$, the finite field with p elements.

The following result show how, in some respects, analysis is simpler for non-Archimedean valuations.

Theorem 2.12. *Let K be a field, complete with respect to a non-Archimedean valuation $|\ |$, and let x_n be a sequence in K. Then: $x_n \to 0$ in $K \iff \sum x_n$ is convergent in K.*

Proof Let $S_N = \sum_{n=1}^{N} x_n$.

$\Rightarrow$: Assume that $x_n \to 0$ in K. Then:

$|S_N - S_M| = |x_{M+1} + \ldots + x_N| \leqslant \max\big(|x_{M+1}|, \ldots, |x_N|\big) \to 0$ as $M, N \to \infty$.

S_N is Cauchy and so convergent (since K is complete), giving that $\sum x_n$ is convergent.

$\Leftarrow$: Assume that $\sum x_n$ is convergent, that is, $S_N \to \ell$ for some $\ell \in K$. Then:

$|x_n - 0| = |x_n| = |S_n - S_{n-1}| = |S_n - \ell + \ell - S_{n-1}| \leqslant |S_n - \ell| + |S_{n-1} - \ell| \to 0$ as $n \to \infty$,

so that $x_n \to 0$ in $K, |\ |$. □

For example, $\sum n!$ converges in any $\mathbb{Q}_p$, since $|n!|_p \to 0$ [it is unknown whether $\sum n! \in \mathbb{Q}$].

The above result applies to $\mathbb{Q}_p$ (since it is non-Archimedean), but not to $\mathbb{R}$ (where, for example, $x_n = \frac{1}{n}$ is a standard counterexample).

Comment 2.13. It is easy to see that, the rules for finite sums in Comment 2.5 and apply to infinite series, namely, when $\sum a_n$ converges, $|\sum a_n| \leqslant \max|a_n|$. Furthermore, if there exists a_k such that $|a_k| > |a_i|$ for all $i \neq k$, then $|\sum a_n| = |a_k|$; in particular, it is then impossible for $\sum a_n = 0$.

Aside: Recall Example 2.7(d), where $x_0 = 3$, and $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, where $f(x) = x^2 - 2$, defined a sequence, which is Cauchy (but not convergent) in $\mathbb{Q}, |\ |_7$, and which is convergent in $\mathbb{Q}_7$ to a root of $f(x)$. The following describes when an initial approximation a_0 gives a solution to $f(x)$.

Theorem 2.14. *(Hensel's Lemma). Let K be a field, complete with respect to a non-Archimedean valuation $|\ |$, with valuation ring $R = \{x \in K : |x| \leqslant 1\}$.*

Let $f(x) \in R[x]$ and let $a_0 \in R$ satisfy: $\quad |f(a_0)| < |f'(a_0)|^2. \qquad (*)$

Then there exists a unique $a \in R$ such that $f(a) = 0$ and $|a - a_0| \leqslant |f(a_0)|/|f'(a_0)|$.

Proof Define $f_j(x)$ by: $\quad f(x + y) = f_0(x) + f_1(x)y + f_2(x)y^2 + \ldots,$

so that $f_0(x) = f(x), f_1(x) = f'(x)$. Define $b_0 = -f(a_0)/f'(a_0)$. By $(*)$, $|b_0| < 1$.

Define $a_1 = a_0 + b_0 = a_0 - f(a_0)/f'(a_0)$. Then:

$|f'(a_1) - f'(a_0)| = |f'(a_0 + b_0) - f'(a_0)| = |(\text{poly in } a_0)b_0 + (\text{poly in } a_0)b_0^2 + \ldots|$

$\qquad \leqslant |b_0| < |f'(a_0)| \quad (\text{by } (*)),$

so that $|f'(a_1)| = |f'(a_0)|$.

Also, $|f(a_1)| = |f(a_0 + b_0)| = |f_0(a_0) + f_1(a_0)b_0 + f_2(a_0)b_0^2 + \ldots|$

$= |f_2(a_0)b_0^2 + \ldots|$ [since $f_0(a_0) + f_1(a_0)b_0 = 0$]

$\leqslant \max_{j \geqslant 2}|f_j(a_0)||b_0|^j \leqslant |b_0|^2 = \frac{|f(a_0)|^2}{|f'(a_0)|^2} = \rho|f(a_0)| < |f(a_0)|$, where $\rho = \frac{|f(a_0)|}{|f'(a_0)|^2} < 1$.

Summarising: $|f'(a_1)| = |f'(a_0)|$ and $|f(a_1)| \leqslant \rho|f(a_0)| < |f(a_0)|$, where $\rho = \frac{|f(a_0)|}{|f'(a_0)|^2} < 1$.

For all n, given $a_n \in R$, define $b_n = -f(a_n)/f'(a_n)$ and $a_{n+1} = a_n + b_n = a_n - f(a_n)/f'(a_n)$. Assume, as induction hypothesis, that:

$|f'(a_n)| = \ldots = |f'(a_1)| = |f'(a_0)|$ and $|f(a_n)| \leqslant \rho|f(a_{n-1})| \leqslant \ldots \leqslant \rho^n|f(a_0)|$. (1)

Then, as above: $|f'(a_{n+1})| = \ldots = |f'(a_1)| = |f'(a_0)|$.

Then $|f(a_{n+1})| \leqslant |b_n|^2$ [justified as for the case $n = 0$ above]

$= \frac{|f(a_n)|^2}{|f'(a_n)|^2} = \frac{|f(a_n)|^2}{|f'(a_0)|^2}$ [by (1), the induction hypothesis]

$\leqslant \frac{|f(a_0)|}{|f'(a_0)|^2}|f(a_n)|$ [since $|f(a_n)| \leqslant |f(a_0)|$ by (1), the induction hypothesis]

$= \rho|f(a_n)| \leqslant \rho^{n+1}|f(a_0)|$ [by (1), the induction hypothesis].

By induction, $\forall n$, $|f'(a_n)| = |f'(a_0)|$ and $|f(a_n)| \leqslant \rho^n|f(a_0)|$ which $\to 0$ as $n \to \infty$. (2)

Now, $|b_n| = |f(a_n)|/|f'(a_n)| = |f(a_n)|/|f'(a_0)| \to 0$, so by Theorem 2.12,

$a_n = a_0 + b_0 + b_1 + \ldots + b_n$ converges to a, say.

By continuity of polynomials, $f(a) = \lim f(a_n) = 0$ [by (2)]. Furthermore:

$|a - a_0| = |\sum b_n| \leqslant \max|b_n| = \max\frac{|f(a_n)|}{|f'(a_n)|} = \max\frac{|f(a_n)|}{|f'(a_0)|} = \frac{|f(a_0)|}{|f'(a_0)|}$ [by (2)], as required.

For uniqueness, imagine $\hat{a} \neq a$ also satisfied $f(\hat{a}) = 0$ and $|\hat{a} - a_0| \leqslant |f(a_0)|/|f'(a_0)|$. Let $\hat{b} = \hat{a} - a \neq 0$.

Then $0 = f(\hat{a}) - f(a) = f(a + \hat{b}) - f(a) = \hat{b}f_1(a) + \hat{b}^2 f_2(a) + \ldots$ (3)

But $|\hat{b}| = |\hat{a} - a_0 + a_0 - a| \leqslant \max(|\hat{a} - a_0|, |a - a_0|) \leqslant |f(a_0)|/|f'(a_0)|$

$< |f'(a_0)|$ [by (*)] $= |f_1(a_0)| = |f_1(a)|$ [by (2) and continuity of $|f'(x)|$].

This gives $|\hat{b}^j f_j(a)| \leqslant |\hat{b}^j| \leqslant |\hat{b}^2| < |\hat{b}f_1(a)|$ (since $|\hat{b}| \neq 0$ & $|\hat{b}| < |f_1(a)|$) for $j \geqslant 2$, so that the leading term of the sum in (3) has valuation strictly greater than the valuations of the other terms, which is inconsistent with the sum being 0. Hence a is unique. □

Example 2.15. Let $f(x) = x^3 - 7$ and $a_0 = 3$. Then $|f(a_0)|_5 = |3^3 - 7|_5 = 5^{-1}$ and $|f'(a_0)|_5 = |3 \cdot 3^2|_5 = 1$. So $|f(a_0)|_5 < |f'(a_0)|_5^2$ and by Hensel's Lemma there exists $a \in \mathbb{Z}_5$ such that $f(a) = 0$, that is: $a^3 = 7$.

Corollary 2.16. *Let $\alpha \in \mathbb{Q}_p$ with $|\alpha|_p = 1$. When $p \neq 2$, α is a square in $\mathbb{Q}_p$ iff it is a square modulo p. When $p = 2$, α is a square in $\mathbb{Q}_p$ iff $\alpha \equiv 1$ (mod 8).*

Example 2.17. $23 \in \left(\mathbb{Q}_7^*\right)^2$ since $|23|_7 = 1$ and $23 \equiv 2 \equiv 3^2 \pmod 7$. However, $24 \notin \left(\mathbb{Q}_7^*\right)^2$ since $|24|_7 = 1$ and $24 \equiv 3 \pmod 7$, which is not a quadratic residue mod 7.

The corollary does not apply to decide the status of 14, but in fact we can see that $14 \notin \left(\mathbb{Q}_7^*\right)^2$, since if $14 = \gamma^2$ for some $\gamma \in \mathbb{Q}_7$ then $|\gamma|_7^2 = |\gamma^2|_7 = |14|_7 = 7^{-1}$, contradicting the fact that $|\gamma|_7 = 7^r$ for some $r \in \mathbb{Z}$.