Spring-School RWTH Aachen 2008
Mathematik mit SAGE

Wir wollen uns in den Umgang mit SAGE einarbeiten, indem wir, in die nachstehend aufgeführten Gruppen aufgeteilt, das jeweils beschriebene Projekt bearbeiten. Hierzu ist es notwendig, dass die Teilnehmer sich soweit als möglich schon vorab mit den für das Projekt notwendigen mathematischen Begriffen und der eigentlichen Problemstellung vertraut machen. Sie finden zu Ihrem Projekt jeweils eine Liste der Literatur, mit der Sie sich vor Beginn der Spring School auseinandersetzen sollten. Wir haben versucht, diese Liste möglichst kurz zu halten; dennoch werden Sie sicherlich je nach Vorbildung noch zusätzlich einige Begriffe selbständig erarbeiten müssen. Bei Ihrer Projektbeschreibung finden Sie eine Liste derjenigen mathematischen Begriffe, die Sie zu Beginn der Veranstaltung in jedem Fall nachgelesen haben sollten.

Die zweite Schwierigkeit, die es zu meistern gilt, ist das Erlernen von Python und SAGE. Zwar werden wir uns beim ersten Treffen hiermit beschäftigen, jedoch wird die 90-minütige Einführung natürlich nicht wirklich ausreichen, um Sie zu einem Experten zu machen. Daher ist es wichtig, dass Sie auch schon vorab etwas in der Dokumentation zu Python [Python] und zu SAGE [SAGE] querlesen. Falls Sie Python zur Verfügung haben (oder sich zutrauen, es selbstständig zu installieren), so sollten Sie damit auch schon etwas experimentieren.

Links

[Python] http://www.python.org/doc/current/tut/tut.html
[SAGE] http://www.sagemath.org/documentation.html

Projekt I: Kegelschnitte

Teilnehmer

Jan HACKFELD, Margarete TENHAAK, Philipp TENHAAK

Projektbeschreibung

Sei K ein Körper und F eine symmetrische $3x3$-Matrix über K mit $\det(F) \neq 0$, sei

$$CS := \left\{ [x : y : z] : (x, y, z) F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \right\}$$

der zugeordnete projektive Kegelschnitt über dem Körper K. Wir nehmen an, dass CS nicht leer ist. Sei N ein Punkt auf CS und g eine fest vorgegeben Gerade, die N nicht enthält, und sei CS_0 die Menge, die entsteht, wenn man aus CS die Schnittpunkte von g mit CS entfernt. Wir definieren die *Summe* zweier Punkte A und B von CS_0 folgendermassen: Sei $g_{A,B}$ die Gerade durch A und B (die Tangente an CS durch A, falls $A = B$), sei S der Schnittpunkt von $g_{A,B}$ mit g, sei $g_{N,S}$ die Gerade durch N und S, und sei schließlich C der zweite Schnittpunkt von $g_{N,S}$ mit CS (und $C = N$, falls $g_{N,S}$ Tangente an CS ist). Wir setzen $A+B := C$. Hierdurch wird tatsächlich eine Gruppenstruktur auf CS_0 erklärt.

Wir wollen diese Gruppen in SAGE implementieren, indem wir etwa zwei Klassen `CSGroup` und `CSGroupElement` entwerfen.

Sofern Zeit bleibt, können wir ggf. noch einen Algorithmus implementieren, der zu vorgegebenem Kegelschnitt über einem Primkörper prüft, ob er einen Punkt enthält, und falls ja, einen solchen berechnet (Anwendung des Satzes von Legendre).

Vorbereitung

Lesen Sie Seite 1 bis 20 in [1], und versuchen Sie den etwas abstrakten Text zu verstehen. Überfliegen Sie anschließend im gleichen Buch die einschlägigen Stellen zu Kegelschnitten (Conics) — Sie werden möglicherweise vieles davon nicht vollständig verstehen, ein Blick wird dennoch nützlich sein. Blättern Sie durch den Artikel [2]: lassen Sie sich nicht entmutigen, das meiste Vokubaluar wir Ihnen unbekannt sein. Wenn wir währed der Veranstaltung gemeinsam in [2] etwas ansehen, wird es Ihnen helfen, sich dort zumindest schon einmal

grob orientiert zu haben. Überlegen Sie, welche Berechnungen zur Lösung der Projektaufgabe notwendig sind.

Falls Ihnen Samuel zu abstrakt ist, so koennen Sie vielleicht in der Bibliothek [3] ausleihen; das letzte Kapitel (ab der 2. Auflage) betrifft die projektive Geometrie.

Begriffe

Projektive Ebene über einem beliebigen Körper, homogene Koordinaten, ebene algebraische Kurve, Kegelschnitt (conic section), Formeln aus der analytischen Geometrie für den Schnittpunkt zweier projektiver Geraden, für die projektive Gerade durch zwei Punkte.

Literatur

1 Pierre Samuel, Projective Geometry. Springer 1988

2 Franz Lemmermeyer, Conics — A Poor man's Elliptic Curves, preprint 2003

3 M. Koecher, A. Krieg: Ebene Geometrie, Springer-Verlag 1993

Projekt II: Untergruppen von $\mathrm{SL}(2,\mathbb{Z})$

Teilnehmer

Hatice BOYLAN, Judith KREUZER, Dominic STEFFEN GEHRE

Projektbeschreibung

Die Modulgruppe $\Gamma = \mathrm{SL}(2,\mathbb{Z})$ ist endlich erzeugt. Als Erzeugende kann man etwa die Matrizen $T := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ und $S := \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ nehmen. Nach einem allgemeinen (und leicht einzusehenden Satz, vgl. [1], Satz 4.1) besitzt die Gruppe Γ für jede natürliche Zahl n nur endlich viele Untergruppen mit Index n. Diese enthält man im Wesentlichen, indem man alle Gruppenhomomorphismen $\Gamma \to S_n$ auflistet, wo n die symmetrischen Gruppe von n Elementen bedeutet. Hierzu wiederum benötigt man den Satz, dass Γ isomorph zum Quotienten der freien Gruppe mit zwei Erzeugenden s und t nach dem von den Relationen $s^2 = (st)^3$ und $s^4 = 1$ erzeugten Normalteiler ist.

Wir wollen eine Funktion `findSubgroups( n)` entwerfen, die alle Untergruppen vom Index n in Γ ausgibt, und eine Klasse `SubroupOfModularGroup`, deren Instanzen mittels eines Gruppenhomomorphismus $\Gamma \to S_n$ initialisiert werden. Methoden dieser Klasse sind etwa: `isCongruenceSubgroup()`, `genus()`, `nEllipticFixedPoints()` etc.

Vorbereitung

Lesen Sie die ersten Abschnitte in [2], die sich mit $\Gamma = \mathrm{SL}(2,\mathbb{Z})$ beschäftigen. Schauen Sie sich insbesondere die Sätze 1.2.4, 1.2.5 an. Sehen Sie sich den Satz 4.1 in [1] an; der Beweis dieses Satzes ist der Schlüssel zur Lösung der gestellten Aufgabe.

Begriffe

$\Gamma := \mathrm{SL}(2,\mathbb{Z})$, Präsentation einer Gruppe mittels Erzeugender und Relationen, Präsentation von Γ (vgl. [2], Theorem 1.2.5, nach diesem Satz ist Γ isomorph zur freien Gruppe in den Erzeugenden v und v, dividiert durch die Relationen $v^2 = (vu)^3$, $v^4 = 1$), Operation von Γ auf der oberen komplexen Halbebene, Klassifikation der Fixpunkte unter dieser Operation, Untergrup-

pen von Γ, Kongruenzuntergruppen, $\Gamma_0(N)$, $\Gamma(N)$, Geschlecht, Index und Spitzen einer Untergruppe von Γ.

Literatur

1 R.C. Lyndon und P.E. Schupp, Combinatorial Group Theory, Springer 1977

2 Rankin, Modular Forms. Cambridge University Press, 1977

3 Hsu, Identifying Congruence Subgroups of the Modular Group, Proc. AMS 124 (1996), 1351–1359

Projekt IIIa: Ringe von Modulformen

Teilnehmer

Anna PIPPICH, Anna POSINGIES, Daniel JACOBS

Projektbeschreibung

Wird mündlich vereinbart.

Vorbereitung

Begriffe

Literatur

Projekt IIIb: Ringe von Modulformen

Teilnehmer

Marc ENSENBACH, Michael HENTSCHEL, Martin RAUM

Projektbeschreibung

Wird mündlich vereinbart.

Vorbereitung

Begriffe

Literatur

Projekt IV: Höhen algebraischer Zahlen

Teilnehmer

Till DIECKMANN, Elisabeth PETERNELL, Cornelia WIRTZ

Projektbeschreibung

Die Komplexität einer algebraischen Zahl wird in der diophantischen Analysis durch ihre Höhe $H(\alpha)$ gemessen. Die Höhe ist immer ≥ 1, und es ist $H(\alpha) = 1$ genau dann, wenn α eine Einheitswurzel ist. Ein Satz von Zhang besagt, *Eine algebraische Zahl kann nicht gleichzeitig nahe bei* 0 *und bei 1 sein*, genauer: Sind α, β zwei von 0, 1 und $\frac{1+\sqrt{-3}}{2}$ verschiedene algebraische Zahlen und gilt $\alpha + \beta = 1$, so folgt

$$H(\alpha)H(\beta) \geq \sqrt{\frac{1+\sqrt{5}}{2}},$$

mit Gleichheit genau dann, wenn α oder β eine primitive 10-te Einheitswurzel ist. In [1] wird dieser Satz auf algebraische Lösungen α, β einer beliebigen über $\mathbb{Q}$ definierten ebenen Kurve verallgemeinert: Zu jeder solchen Kurve gibt es eine Konstante $C > 1$, sodass für jeden (algebraischen) Punkt (α, β) auf dieser Kurve (bis auf endlich viele Ausnahmen) $H(\alpha)H(\beta) \geq C$ gilt. Die Konstante kann sogar scharf bestimmt werden und wird von endlich vielen Lösungen angenommen.

Wir wollen eine Funktion `findConstant( C)` entwerfen, die zu gegebener ebener Kurve C über $\mathbb{Q}$ die optimale Konstante C für die oben geschilderte Abschätzung ausgibt. Ggf. werden wir hierzu vorab noch eine Klasse `PlaneAffineAlgebraicCurve` entwerfen. Anschliessed werden wir etwas experimentieren und z.B. die Konstanten C für Kegelschnitte studieren. Vielleicht gibt es hier etwas Neues zu entdecken.

Vorbereitung

Erarbeiten Sie sich Theorem 2.1 ([1], S. 32) und die Details des Beweises. Lesen Sie ggf. vorab den Beweis zum Satz von Zhang ([1], S. 15), um die Idee des Beweises besser zu verstehen. Versuchen Sie die gestellte Aufgabe gedanklich in Teilprojekte zu zerlegen.

Begriffe

Algebraische Zahl, affine ebene algebraische Kurve über $\mathbb{Q}$ (vgl. [1], S. 31], die Höhe $H(\alpha)$ einer algebraischen Zahl (vgl. [1], S. 10).

Literatur

1 N-P. Skoruppa, Heights, Lecture Notes, Siegen 2003

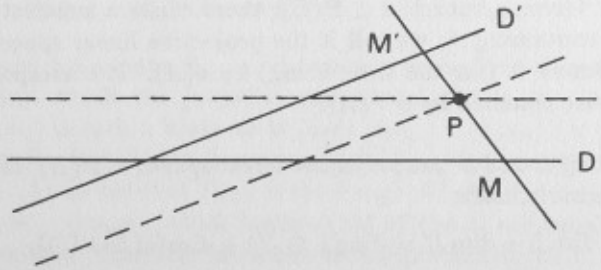
Pierre Samuel, Projective Geometry. Springer 1988

CHAPTER 1

Projective Spaces

1.1. Projective Spaces and Projective Bases

Consider, in the plane, two non-parallel lines D and D', and a point P not contained in either line. To each point M of D associate the point $M' = p(M)$ where the line PM intersects D':



Notice that $p(M)$ is not defined when PM is parallel to D'; on the other hand, the point A' where D' intersects the parallel to D containing P is not in the image of p. There's something "missing" in D and D'; the right thing to work with seems to be the set of *projecting lines*, or lines containing the center P of projection. This motivates the following definition:

Definition 1. Given a vector space E over a field K, the *projective space* associated with E is the set $\mathbf{P}(E)$ of (vector) lines in E.

In this book the term "field" will include skew fields as well as commutative ones, except where we indicate otherwise (see index for a list of such sections). If K is a skew field we assume for concreteness that E is a left vector space over K.

One can also see $\mathbf{P}(E)$ as the quotient of the set $E \setminus 0$ of non-zero vectors modulo the equivalence relation "$x \sim y$ if and only if $y = ax$ for some $a \in K$" (naturally, $a \neq 0$). Thus we have a canonical map $p : E \setminus 0 \to \mathbf{P}(E)$ that associates to each vector x the vector line Kx it spans.

Definition 2. The *dimension* of $\mathbf{P}(E)$ is the integer $\dim E - 1$, which we denote by $\dim \mathbf{P}(E)$.

The projective space $\mathbf{P}(K^{n+1})$ is denoted by $\mathbf{P}_n(K)$; its dimension is n. Projective spaces of dimension one and two are called projective lines and planes, respectively.

Notice that $\mathbf{P}(0)$ is empty; by definition 2, its dimension is -1. A zero-dimensional projective space reduces to a point.

Definition 3. A *(projective) linear subvariety*, or *linear subspace*, of $\mathbf{P}(E)$ is the image $L = p(V)$ of a vector subspace V of E.

This definition embodies an abuse of notation: to be precise we should write $L = p(V \setminus 0)$.

Notice that a projective linear space $L = p(V)$ is the projective space $\mathbf{P}(V)$ associated with V.

An intersection of projective linear spaces is a (possibly empty) projective linear space. Given a subset $A \subset \mathbf{P}(E)$, there exists a smallest projective linear space containing A; we call it the projective linear space *generated* by A, and denote it (for the time being) by $v(A)$. It corresponds to the vector subspace spanned by $p^{-1}(A)$.

Theorem 1. *If L and L' are projective linear spaces in $\mathbf{P}(E)$, the following dimension formula holds:*

$$\dim L + \dim L' = \dim(L \cap L') + \dim\big(v(L \cup L')\big).$$

Proof. This is a direct translation, via definition 2, of the well-known formula for vector subspaces:

$$\dim V + \dim V' = \dim(V \cap V') + \dim(V + V'). \qquad \square$$

Corollary. *If $\dim L + \dim L' \geq \dim \mathbf{P}(E)$, the intersection $L \cap L'$ is non-empty.*

Proof. In fact, theorem 1 gives $\dim(L \cap L') \geq 0$, and the only empty projective linear space has dimension -1. $\square$

We say that a subset $A \subset \mathbf{P}(E)$ is *projectively independent* if it is the image under p of a linearly independent subset of E.

Homogeneous coordinates

We assume from now on that E is finite-dimensional.

Given a basis $(e_0, e_1, \ldots, e_n)$ for E, we can associate to each point $A \in \mathbf{P}(E)$ certain $(n+1)$-tuples of elements of K, namely, the coordinates of the vectors $x \in E$ such that $A = p(x)$. By definition, these $(n+1)$-tuples are all non-zero (that is, they have at least one non-zero component) and proportional to one another: if $(x_0, x_1, \ldots, x_n)$ is one $(n+1)$-tuple, all others will be of the form $(ax_0, ax_1, \ldots, ax_n)$, where $a \in K$ is non-zero. The set of such $(n+1)$-tuples is called the *homogeneous class* of $A \in \mathbf{P}(E)$, and each representative of this class is a set of *homogeneous coordinates* for A. The mapping thus defined from $\mathbf{P}(E)$ into the set of projective classes is called a *projective coordinate system*.

Projective coordinate systems can be characterized intrinsically in terms of $\mathbf{P}(E)$:

Theorem 2. *Let K be commutative.*

(a) *A projective coordinate system on an n-dimensional projective space $\mathbf{P}(E)$ is uniquely determined by the $n+2$ points with homogeneous coordinates $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 1)$, $(1, 1, \ldots, 1)$. Any $n+1$ of these points form a projectively independent set.*

(b) *Conversely, for each $(n+2)$-tuple $(P_0, P_1, \ldots, P_{n+1})$ of points in $\mathbf{P}(E)$ all of whose $(n+1)$-subtuples are projectively independent, there exists a projective coordinate system assigning the coordinates $(1, 0, \ldots, 0)$ to P_0, $\ldots$, $(0, 0, \ldots, 1)$ to P_n and $(1, 1, \ldots, 1)$ to P_{n+1}.*

Proof. The $n+1$ points $P_0, \ldots, P_n$ are not enough to determine the basis of E from which the projetive coordinate system derives, because if $(e_0, e_1, \ldots, e_n)$ is such a basis, so is $(a_0 e_0, a_1 e_1, \ldots, a_n e_n)$ for any non-zero $a_0, \ldots, a_n \in K$. But if both bases assign to P_{n+1} the homogeneous coordinates $(1, \ldots, 1)$ we see that P_{n+1} is the image of both $e_0 + e_1 + \cdots + e_n$ and $a_0 e_0 + a_1 e_1 + \cdots + a_n e_n$, which implies that all the a_j are equal to the same non-zero scalar a. Thus the two bases are proportional, $(e_0, e_1, \ldots, e_n)$ and $(ae_0, ae_1, \ldots, ae_n)$.

Now if $M \in \mathbf{P}(E)$ comes from a point $x \in E$ whose coordinates in the first basis are $(x_0, x_1, \ldots, x_n)$, the coordinates of x in the second basis will be $(x_0 a^{-1}, x_1 a^{-1}, \ldots, x_n a^{-1})$: the two sets of coordinates are proportional (that is, left-proportional) because K is commutative.

To prove part (b), lift $P_0, \ldots, P_n$ to any basis $(e_0, \ldots, e_n)$ of E, and consider the coordinates $(b_0, \ldots, b_n)$ of a vector $u \in p^{-1}(P_{n+1})$ in this basis. All the b_j are different from zero, so we just change our basis to $(b_0 e_0, \ldots, b_n e_n)$. $\square$

Part (b), and the last assertion in part (a), hold even if K is a skew field.

Definition 4. An $(n+2)$-tuple $(P_0, P_1, \ldots, P_{n+1})$ of points in $\mathbf{P}(E)$ is called a *(projective) frame* (or *projective base*) of $\mathbf{P}(E)$ if, for some projective coordinate system, the homogeneous coordinates of $P_0, \ldots, P_n$ are $(1,0,\ldots,0)$, $(0,1,\ldots,0)$, $\ldots$, $(0,0,\ldots,1)$, respectively, and those of P_{n+1} are $(1,\ldots,1)$. The points $P_0, \ldots, P_n$ are called the *vertices*, and P_{n+1} the *unit point*, of the frame.

Corollary. *An $(n+2)$-tuple $(P_0, P_1, \ldots, P_{n+1})$ of points in an n-dimensional projective space is a projective frame if and only if any $n+1$ points among them are projectively independent.* □

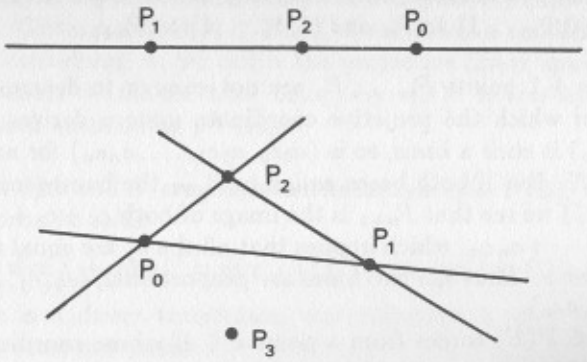
It amounts to the same to say that no $(n-1)$-dimensional projective linear space, or *hyperplane*, contains $n+1$ of these points.

This corollary holds even if K is a skew field.

Examples

(1) A frame for a projective line is formed by any three distinct points ("pairwise distinct", as purists would have it).

(2) In a projective plane, a frame is formed by four points, three of which form a non-degenerate triangle and the fourth of which does not belong to any of the sides of the triangle. In this way no three points are collinear.



Homogeneous coordinates can be used to write equations for projective linear spaces of $\mathbf{P}(E)$. Namely, given a basis $(e_0, \ldots, e_n)$ for the vector space E, a hyperplane H has equation

$$x_0 b_0 + x_1 b_1 + \cdots + x_n b_n = 0, \qquad \text{with } b_j \in K \text{ not all zero,} \tag{1}$$

which expresses the condition that the point with coordinates $(x_0, \ldots, x_n)$ be on H. The same equation (1) also expresses the condition that a point of $\mathbf{P}(E)$ with homogeneous coordinates $(x_0, \ldots, x_n)$ lies in the projective hyperplane $p(H)$; notice that any other set $(ax_0, \ldots, ax_n)$ of homogeneous coordinates for this point also satisfies (1).

As an intersection of hyperplanes, a projective linear space is defined by a system of homogeneous equations of the form (1). More precisely, if a projective linear space L has *codimension* d, that is, if its dimension is $n-d$, the space can be defined by a system of d equations whose left-hand sides are linearly independent linear forms.

Notice that in (1), the coefficients are written to the right of the variables. In fact, if f is a linear form having H as its kernel, we have $f(x_0 e_0 + \cdots + x_n e_n) = x_0 f(e_0) + \cdots + x_n f(e_n) = 0$.

If K is commutative, one defines an *algebraic subset* of E to be any subset given by a system of polynomial equations

$$F_j(x_0, x_1, \ldots, x_n) = 0 \qquad \text{for } j = 1, \ldots, q, \tag{2}$$

in some fixed basis of E. A change of basis alters these equations, but not their property of being polynomial, nor their degrees.

In translating this to the projective case, it's best to assume that the polynomials F_j are homogeneous. Then a system of equations of the form (2), if satisfied by one set of homogeneous coordinates of a point of $\mathbf{P}(E)$ in a given projective frame, is satisfied by the whole homogeneous class of the point. The equations are said to define an *algebraic subset* of $\mathbf{P}(E)$.

Cardinality over finite fields

Let K be the field $\mathbf{F}_q$ with q elements. If $\mathbf{P}(E)$ has dimension n, its characterization as a quotient space of $E \setminus 0$ immediately shows that

$$\#\mathbf{P}(E) = \frac{q^{n+1}-1}{q-1} = q^n + q^{n-1} + \cdots + q + 1. \tag{3}$$

Thus a projective line over $\mathbf{F}_q$ has $q+1$ points (at least three, since $q \geq 2$), and a projective plane q^2+q+1 points.

The number of bases of E is $(q^{n+1}-1)(q^{n+1}-q)\cdots(q^{n+1}-q^n)$, since we can start by choosing any non-zero vector, then any vector not proportional to the first, and so on. Since a projective frame is determined, up to a non-zero scalar factor, by a basis of E (theorem 2), we conclude that the number of frames of $\mathbf{P}(E)$ is

$$(q^{n+1}-1)(q^{n+1}-q)\cdots(q^{n+1}-q^{n-1})q^n. \tag{4}$$

For lines and planes, respectively, the number of frames is $q(q^2-1) = q(q-1)(q+1)$ and $q^2(q^3-1)(q^3-q) = q^3(q-1)^2(q+1)(q^2+q+1)$.

$\mathbf{P}(E)$ has as many d-dimensional projective linear spaces as E has $(d+1)$-dimensional vector subspaces. The number of such subspaces is the number

of sets of $d+1$ linearly independent vectors in E, divided by the number of such sets as span the same subspace. This shows that the number of d-dimensional projective linear spaces of $\mathbf{P}(E)$ is

$$\frac{(q^{n+1}-1)(q^{n+1}-q)\cdots(q^{n+1}-q^d)}{(q^{d+1}-1)(q^{d+1}-q)\cdots(q^{d+1}-q^d)}. \tag{5}$$

For q large this number is asymptotically $q^{(d+1)(n-d)}$.

In particular, the number of lines in $\mathbf{P}(E)$ is

$$\frac{(q^{n+1}-1)(q^n-1)}{(q-1)^2(q+1)}.$$

1.2. Projective Transformations and the Projective Group

Let u be a linear map from a vector space E into a vector space F. Since u preserves vector lines, it defines a map between the quotient spaces $\mathbf{P}(E)$ into $\mathbf{P}(F)$, as long as non-zero vectors are mapped into non-zero vectors, that is, u is one-to-one. The map $\mathbf{P}(u) : \mathbf{P}(E) \to \mathbf{P}(F)$ thus obtained is called a *projective map*, and a *projective transformation* if it is bijective, that is, if $\dim \mathbf{P}(E) = \dim \mathbf{P}(F)$. Projective transformations are sometimes called homographies.

When u is not one-to-one we obtain a map defined on the complement of $p(\ker(u))$.

Given another one-to-one linear map v from F into a third vector space G, we can write

$$\mathbf{P}(v \circ u) = \mathbf{P}(v) \circ \mathbf{P}(u); \tag{6}$$

we also clearly have $\mathbf{P}(\mathrm{Id}_E) = \mathrm{Id}_{\mathbf{P}(E)}$.

Theorem 3. *Let E and F be vector spaces, with $\dim E \geq 2$, and let $Z = \{a \in K \mid ab = ba \text{ for all } b \in K\}$ be the center of K. Two one-to-one linear maps u and u' from E into F satisfy $\mathbf{P}(u) = \mathbf{P}(u')$ if and only if there exists a scalar $a \in Z$ such that $u'(x) = au(x)$ for every $x \in E$.*

The cases $\dim E = 0, 1$ are left to the reader.

Proof. The condition is obviously sufficient, since $a \in Z$ implies that $u \mapsto au$ is linear. Conversely, if $\mathbf{P}(u) = \mathbf{P}(u')$, there exists, for every non-zero $x \in E$, some scalar $a(x)$ such that $u'(x) = a(x)u(x)$. Taking x and y linearly independent and expressing $u'(x+y)$ in two different ways we find $a(x) = a(x+y) = a(y)$. This implies that $a(x) = a(y)$ for every x and y,

since we can find z proportional to neither x nor y. There remains to show that $a = a(x)$ is central.

For any $b \in K$ and non-zero $x \in E$, we have $u'(bx) = au(bx) = abu(x)$ and $u'(bx) = bu'(x) = bau(x)$. Since $u(x) \neq 0$, this implies that $ab = ba$, showing that $a \in Z$. □

Corollary. *A one-to-one linear map of a vector space that transforms each vector into a multiple of itself is of the form $u \to au$, where a is a central scalar.* □

For K commutative this condition can also be stated in terms of eigenspaces. A map of the form $u \to au$ is called a *homothety*.

It follows from (6) that the projective transformations of $\mathbf{P}(E)$ into itself form a group, called the *projective group* of $\mathbf{P}(E)$ and denoted by $\mathrm{PGL}(E)$. Theorem 3 can be rephrased to say that if $\dim \mathbf{P}(E) \geq 2$ we have $\mathrm{PGL}(E) = \mathrm{GL}(E)/Z^*$, where $\mathrm{GL}(E)$ is the linear group of E and Z the center of K.

Notice that the fixed points of a projective transformation $\mathbf{P}(u)$ are the images of the (non-zero) eigenvectors of u.

Assume K commutative and fix a projective frame for $\mathbf{P}(E)$ (or, equivalently, fix a basis of E up to a scalar factor). A projective transformation of $\mathbf{P}(E)$ can be expressed in this basis by a class of proportional non-singular matrices, whose entries b_{ij} are defined by the condition that

$$ay_j = b_{j0}x_0 + b_{j1}x_1 + \cdots + b_{jn}x_n \qquad \text{for } j = 0, 1, \ldots, n, \tag{8}$$

where $a \in K^*$ is arbitrary, $(x_0, \ldots, x_n)$ are the homogeneous coordinates of an arbitrary point in E and $(y_0, \ldots, y_n)$ the homogeneous coordinates of its image.

Theorem 4. *Let $\mathbf{P}(E)$ and $\mathbf{P}(E')$ be projective spaces of same dimension n over a commutative field K, with projective frames $(P_0, \ldots, P_n, P_{n+1})$ and $(P'_0, \ldots, P'_n, P'_{n+1})$, respectively. There exists a unique projective transformation $h : \mathbf{P}(E) \to \mathbf{P}(E')$ such that $h(P_i) = P'_i$ for all $i = 0, 1, \ldots, n, n+1$.*

Proof. Lift $(P_0, \ldots, P_n)$ to a basis $(e_0, \ldots, e_n)$ of E such that $p(e_0 + \cdots + e_n) = P_{n+1}$ (theorem 2), and lift $(P'_0, \ldots, P'_n)$ to $(e'_0, \ldots, e'_n)$. If h exists and is of the form $h = \mathbf{P}(u)$, each $u(e_i)$, for $i = 0, \ldots, n$, must be of the form $a_i e'_i$, where a_i is a non-zero scalar. Since $h(P_{n+1}) = P'_{n+1}$, the vector $u(e_0 + \cdots + e_n)$ can be written $b(e'_0 + \cdots + e'_n)$. Thus all the a_i are equal to b; this determines u up to a multiplicative scalar, and $h = \mathbf{P}(u)$ uniquely (theorem 3). The existence of u is obvious: define u by $u(e_i) = e'_i$ for $i = 0, 1, \ldots, n$. □

If K is skew, uniqueness fails. For example, take a and b in K such that $ab \neq ba$. If u is the linear map that takes the canonical basis $(e, f) = ((1,0),(0,1))$ of K^2

into (ae, af), it is easily checked that $h = \mathrm{P}(u)$ leaves invariant the points of the "canonical" frame of $\mathrm{P}(K^2)$ (those with homogeneous coordinates $(1,0)$, $(0,1)$ and $(1,1)$). But $\mathrm{P}(u)$ cannot be the identity, otherwise $u(e+bf) = ae + baf$ would be of the form $c(e+bf)$, which would imply $c = a$ and $ab = ba$.

Remark. Theorem 4 shows that, for K commutative, the number of elements of $\mathrm{PGL}(E)$ is equal to the number of frames of $\mathrm{P}(E)$. In particular, if K is finite, this number is given by formula (4).

1.3. Projective and Affine Spaces

Recap on affine spaces

Recall that an *affine space* is a set E on which the additive group of a vector space (denoted by $v(E)$ or $\vec{E}$) acts simply transitively. One often says that the elements of E are *points* and those of $v(E)$ are vectors, or translations of E; and $v(E)$ itself is called the vector space *underlying* E. The image of a point a under the translation t is generally denoted by $t+a$, whence the formulas

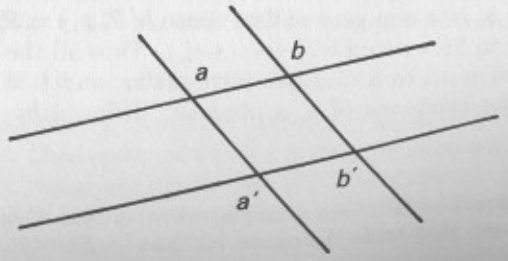
$$s + (t + a) = (s + t) + a,$$
$$0 + a = a,$$

which simply translate the fact that a group is acting on a set. The unique translation that takes a point a into a point b is denoted by $b - a$, or $\overrightarrow{ab}$. In this notation we have *Chasles's formula*

$$(c - b) + (b - a) = c - a. \tag{9}$$

The commutativity of the group $v(E)$ is equivalent to the *parallelogram rule*

$$b - a = b' - a' \quad \text{if and only if} \quad a - a' = b - b'. \tag{10}$$



The choice of a point $a \in E$ allows one to identify E with its underlying vector space $v(E)$: each point $m \in E$ gets associated with the vector $m - a \in v(E)$. Although this choice of an origin, or *vectorialization*, is by no means intrinsic, one often performs it in order to make calculations easier.

An *affine frame* of E is made up of a point a_0 and a basis $(e_1, \ldots, e_n)$ of $v(E)$. The coordinates of a point m in this frame are those of the vector $m - a_0$ in the given basis. It amounts to the same to give the $n+1$ points $a_0, a_1 = e_1 + a_0, \ldots, a_n = e_n + a_0$.

An *affine linear subspace* (or *subvariety*) L is a subset of E that is either empty or of the form $L = V + a$, where V is a vector subspace of $v(E)$ and a is a point in E. Since $V + a = V' + a'$ if and only if $V = V'$ and $a' - a \in V$, the vector subspace V is uniquely determined by L; it is called the *direction* of L. Two affine subspaces are called *parallel* if they have the same direction. When we choose an origin for E, affine subspaces (other than the empty one) are simply translates of vector subspaces of $v(E)$. Every intersection of affine subspaces is one, so we have the notion of the affine subspace generated by a subset $A \subset E$.

Given points $m_1, \ldots, m_q \in E$ and scalars $a_1, \ldots, a_q \in K$ such that $a_1 + \cdots + a_n = 1$, we define the *barycenter* of the m_i with weights a_i as the unique point g such that

$$g - p = a_1(m_1 - p) + \cdots + a_q(m_q - p)$$

for every $p \in E$.

The operation of taking barycenters is "associative": a barycenter of barycenters of points m_i is a barycenter of the m_i. It can be shown that the set of barycenters of a set of points m_i is just the affine subspace generated by the m_i. In particular, if a subset $S \subset E$ is invariant under the operation of taking barycenters of sets of points, S is an affine subspace.

Remark. For $K \neq \mathbf{F}_2$ a subset invariant under the operation of taking barycenters of *two points* is an affine subspace. But over $\mathbf{F}_2$ the barycenters of two points are just the two points, so all subsets are invariant under this operation.

In an affine frame, with coordinates denoted by $(x_1, \ldots, x_n)$, an affine subspace is defined by a system of linear equations

$$x_1 a_{j1} + \cdots + x_n a_{jn} = b_j \qquad \text{for } j = 1, \ldots, q,$$

where the a_{ji} and the b_j are scalars.

One can choose the linear forms so that their left-hand sides are linearly independent; then the dimension of the affine subspace is $n - q$. (The dimension of an affine subspace is the dimension of its direction.)

This is only true about non-empty affine subspaces. The empty affine subspace can be defined by the equations $x_1 = 0$ and $x_1 = 1$, for example.

Example: the complement of a hyperplane in a projective space

Theorem 4. *Let P be an n-dimensional projective space and $H \subset P$ a hyperplane. Denote by T the set containing the identity and the projective transformations of P that leave invariant exactly those points of P that belong to H. Then T is a group isomorphic to the additive group of an n-dimensional vector space, and it acts simply transitively on $P \setminus H$.*

Proof. Write $P = \mathbf{P}(V_1)$ and $H = \mathbf{P}(H_1)$, where V_1 is a vector space over K and $H_1 \subset V_1$ a hyperplane. Choose $z \in V_1$ such that V_1 is the direct sum of H_1 and Kz, and consider $u \in \mathrm{GL}(V_1)$ such that $\mathbf{P}(u) \in T$ and $\mathbf{P}(u) \neq \mathrm{Id}$. Since $\mathbf{P}(u)$ leaves invariant all points in H, there exists $a \in K$ such that $u(x) = ax$ for every $x \in H_1$. On the other hand, we can write $u(z) = h + bz$, with $h \in H_1$ and $b \in K$.

A fixed point of $\mathbf{P}(u)$ comes from an eigenvector of u; if we write such a vector in the form $x + cz$, with $x \in H_1$ and $c \in K$, the condition is that $u(x+cz) = d(x+cz)$ for some $d \in K$, that is, that $ax + c(h+bz) = dx + dcz$. This is equivalent to the system

$$\text{(S)} \qquad \begin{aligned} (a-d)x + ch &= 0, \\ cb &= dc. \end{aligned}$$

Thus $\mathbf{P}(u)$ has a fixed point outside H if and only if there exists a solution (c, d, x) of (S) with $c \neq 0$. If $h = 0$, there exists such a solution with $c = 1$, $d = b$ and $x = 0$; the assumption $\mathbf{P}(u) \in T$ then requires $\mathbf{P}(u) = 1$, whence $a = b$. If $h \neq 0$, on the other hand, (S) and $c \neq 0$ together imply $d = cbc^{-1}$ and $(a - cbc^{-1})x = -ch$; this has a solution unless $a - cbc^{-1}$ vanishes for every non-zero c, that is, unless a and b are equal and belong to the center of K.

Thus if $\mathbf{P}(u) \in T$ we can normalize u so as to make $a = b = 1$; then u is the identity on H_1 and $u(z)$ is of the form $u(z) = h_u + z$, where $h_u \in H_1$ is uniquely determined by $\mathbf{P}(u)$. Furthermore, it's easy to see that $h_{u \circ v} = h_u + h_v$ if $\mathbf{P}(u), \mathbf{P}(v) \in T$. Hence the group structure on T. That T acts simply transitively on $P \setminus H$ follows from the fact that the translations h_u act simply transitively on $z + H_1$. □

Embeddings of an affine space in a projective space

Let E be an affine space over K, with underlying vector space $v(E)$, and $a \in E$ a point. The *projective closure* $\hat{E}$ of E is defined as

$$\hat{E} = \mathbf{P}\big(v(E) \times K\big). \tag{11}$$

We also define an injection $j_a : E \to \hat{E}$ by

$$j_a(m) = p(m - a, 1) \qquad \text{for } m \in E. \tag{12}$$

This injection is determined, up to a translation, by the choice of a. The image $j_a(E)$ is the complement of the hyperplane $\mathbf{P}\big(v(E) \times 0\big)$ in $\hat{E}$.

Let E and F be affine spaces over K. A map $f : E \to F$ is called an *affine map* if there exists a linear map $v(f) : v(E) \to v(F)$ such that $v(f)(m - m') = f(m) - f(m')$ for every $m, m' \in E$. If $f : E \to F$ is affine and one-to-one, it defines a projective map

$$\hat{f} = \mathbf{P}\big(v(f) \times \mathrm{Id}_K\big) : \hat{E} \to \hat{F}. \tag{13}$$

This map extends f in the sense that

$$\hat{f}\big(j_a(m)\big) = j_{f(a)}\big(f(m)\big) \qquad \text{for every } m \in E; \tag{14}$$

this is because $\hat{f}\big(j_a(m)\big)$ is the canonical image in $\hat{F}$ of $\big(f(m) - f(a), 1\big)$.

With the obvious notations, we have

$$\hat{u} \circ \hat{v} = \widehat{u \circ v}, \tag{15}$$

so there is a canonical injection from the affine group of E into the projective group of $\hat{E}$. Its image consists of the projective transformations that leave the *hyperplane at infinity* $\mathbf{P}\big(v(E) \times 0\big)$ globally invariant.

Affine coordinates and homogeneous coordinates

Thus every affine space E can be seen as the complement $P \setminus H$ of a hyperplane in a projective space. This hyperplane and all sets lying in it are said to be *at infinity*. Notice that two affine subspaces are parallel if and only if they have the same points at infinity (more rigorously, we should say that their projective completions have the same points at infinity, but we won't be so sticky).

Conversely, if we pick a hyperplane H in a projective space P and concentrate on the affine structure of $P \setminus H$, we'll often say that H is the hyperplane at infinity, or that H has been *sent to infinity*. We can then choose a projective coordinate system $(x_0, x_1, \ldots, x_n)$ on P in such a way that H is the hyperplane of equation $x_0 = 0$. If $m \in P$ is a point not on H, the n-tuple $(x_0^{-1}x_1, \ldots, x_0^{-1}x_n)$ gives the affine coordinates of m. If a projective linear space L of P which is not at infinity is given by the system of (homogeneous) equations

$$x_0 a_{j0} + x_1 a_{j1} + \cdots + x_n a_{jn} = 0 \qquad \text{for } j = 1, \ldots, q,$$

its intersection with $P \setminus H$ is the affine subspace defined by the (affine) equations $a_{j0} + y_1 a_{j1} + \cdots + y_n a_{jn} = 0$. If K is commutative and L is an algebraic subset of P, not contained in H, and defined by the homogeneous polynomial equations

$$F_j(x_0, \ldots, x_n) = 0 \qquad \text{for } j = 1, \ldots, q,$$

the intersection $L \cap (P \setminus H)$ is defined by the equations $F_j(1, y_1, \ldots, y_n) = 0$. In sum, to pass from projective to affine equations, just take $x_0 = 1$.

When L is at infinity, the system of affine equations obtained by this procedure is "impossible", that is, it has no solutions, even over the algebraic closure of K.

Conversely, let E be an affine space with a fixed affine frame, and denote by $(y_1, \ldots, y_n)$ the coordinates of a point $m \in E$. Embed E in $\hat{E} = P$ using the injection j_a associated with the origin a of the chosen frame. By (11) and (12), $(1, y_1, \ldots, y_n)$ is a set of homogeneous coordinates for $j_a(m)$, in the corresponding projective coordinate system of $K \times v(E)$. If L is an affine subspace of E defined by the equations

$$y_1 a_{j1} + \cdots + y_n a_{jn} = b_j \qquad \text{for } j = 1, \ldots, q,$$

the projective closure $\hat{L}$ of L is given by the equations $x_1 a_{j1} + \cdots + x_n a_{jn} = x_0 b_j$, and we have $L = (P \setminus H) \cap \hat{L}$, where H is the hyperplane at infinity.

Now assume that K is commutative and consider an algebraic subset A of the affine space E, defined by a single polynomial equation $F(y_1, \ldots, y_n) = 0$. Denote by d the total degree of F, and form the homogeneous polynomial F_h of degree d associated with F:

$$F_h(x_0, x_1, \ldots, x_n) = x_0^d F(x_1/x_0, \ldots, x_n/x_0). \tag{15}$$

The algebraic subset of P defined by the homogeneous equation $F_h(x_0, x_1, \ldots, x_n) = 0$ is called the *projective closure* of A, and is denoted by $\hat{A}$. Since $F_h(1, y_1, \ldots, y_n) = F(y_1, \ldots, y_n)$, the set A is the intersection of $\hat{A}$ with $P \setminus H$. The points of $\hat{A} \setminus A$ are called points at infinity of A; they make up an algebraic subset of H.

In this discussion we have limited ourselves to hypersurfaces, or algebraic sets defined by a single equation (hypersurfaces are called curves or surfaces if $n = 2$ or 3, respectively). The dimension of such objects, whether affine or projective, is $n - 1$, by any reasonable definition.

In treating lower-dimensional algebraic subsets of E, defined by several polynomial equations, it's not enough to homogenize the defining equations; one must also homogenize all the polynomials in the ideal generated by them.

For example, consider in $\mathbf{C}^3$ the circle C defined by the equations $x^2 + y^2 + z^2 - 1 = 0$ and $x^2 + y^2 + z^2 - 2x = 0$. The corresponding homogeneous equations are $x^2 + y^2 + z^2 - t^2 = 0$ and $x^2 + y^2 + z^2 - 2xt = 0$ (where the homogenizing variable is written t instead of x_0). The algebraic set defined by these two equations is the union of C with a curve at infinity, of equation $x^2 + y^2 + z^2 = t = 0$, which is called an umbilic. But the actual projective closure of C is smaller than that: it has only two points at infinity, where it intersects the umbilic. The reason is that the polynomial $2x - 1$, for example, is in the ideal generated by $x^2 + y^2 + z^2 - 1$ and $x^2 + y^2 + z^2 - 2x$, so by definition points in $\hat{C}$ must be zeros of the homogenized polynomial $2x - t = 0$. In this case we can get around the problem of extra points at infinity by replacing one of the two equations of spheres that define C by the equation $2x - 1 = 0$ of their radical plane. There are cases, however, where no such replacement is possible.

Given a projective space P and a system of projective coordinates for it, say $(x_0, x_1, \ldots, x_n)$, the hyperplanes H_i of equation $x_i = 0$, for $i = 0, \ldots, n$, have empty intersection, which means that P is the union of the $n + 1$ affine spaces $P \setminus H_i$. The affine coordinates in $P \setminus H_i$ of a point whose homogeneous coordinates are $(x_0, x_1, \ldots, x_n)$ are given by

$$(x_i^{-1} x_0, \ldots, x_i^{-1} x_{i-1}, x_i^{-1} x_{i+1}, \ldots, x_i^{-1} x_n).$$

Consider a point in $P \setminus (H_0 \cup H_i)$, and let its affine coordinates in $P \setminus H_0$ be $(y_1, \ldots, y_n)$; by assumption, $y_i \neq 0$. The homogeneous coordinates of this point are $(1, y_1, \ldots, y_n)$, so its affine coordinates in $P \setminus H_i$ are

$$(y_i^{-1}, y_i^{-1} y_1, \ldots, y_i^{-1} y_{i-1}, y_i^{-1} y_{i+1}, \ldots, y_i^{-1} y_n). \tag{16}$$

In practice we allow ourselves some abuses in notation. For example, if we start from the affine curve C defined by $x^3 + xy + 1 = 0$ and denote by z the homogenizing variable, the projective closure $\hat{C}$ of C is given by $x^3 - xyz + z^3 = 0$; in order to study the point $(x, y, z) = (0, 1, 0)$, the only point at infinity of the closure, we can make $y = 1$, obtaining the equation $x^3 + z^3 - xz = 0$ for the "affine piece" of $\hat{C}$ that lies in the affine space $y \neq 0$. Obviously the letters x, y, z don't have the same meaning in the three equations.

Simple and multiple points

Here we assume that K is commutative and infinite. Let V be an affine hypersurface with equation $F(y_1, \ldots, y_n) = 0$. We will write the polynomial F in the form

$$F(Y) = F_0 + F_1(Y) + \cdots + F_d(Y),$$

where $Y = (Y_1, \ldots, Y_n)$, F_j is homogeneous of degree j and $F_d \neq 0$. The integer d is called the *degree* of F. Let D be the affine line defined by $y_i = a_i + b_i t$, where $i = 1, \ldots, n$ and $t \in K$ is a parameter; D goes through the point $A = (a_1, \ldots, a_n)$. The parameter values at the intersections of D with V are the roots of the equation

$$F(a_1 + b_1 t, \ldots, a_n + b_n t) = 0. \tag{18}$$

This equation is identically satisfied if and only if V contains D, since we assumed K infinite; from now on we exclude this case. Otherwise (18) has degree at most d, so D has at most d distinct common points $P_1, \ldots, P_r$ with V. Let $t_1, \ldots, t_r$ be their parameters. The multiplicity m_j of the root t_j of (18) depends only on the point P_j; it remains the same if we change frames or if we change the parameter along D (by an affine transformation). This number m_j is called the *intersection multiplicity* of V and D at P_j.

If the point $A = (a_1, \ldots, a_n)$ is on V, (18) has a root at $t = 0$. This root is simple if and only if the coefficient of t in (18) is non-zero. By Taylor's formula, this coefficient is $F_1'(a) b_1 + \cdots + F_n'(a) b_n$, where $F_i'(a)$ is the i-th partial derivative of F at $(a_1, \ldots, a_n)$. If at least one partial derivative at A is non-zero, we say that A is a *simple point* of V. Then the root $t = 0$ is simple unless the vector $(b_1, \ldots, b_n)$ satisfies $F_1'(a) b_1 + \cdots + F_n'(a) b_n = 0$;

this condition amounts to saying that the tip $(y_1, \dots, y_n)$ of the vector is on the hyperplane

$$F_1'(a)(y_1 - a_1) + \cdots + F_n'(a)(y_n - a_n) = 0, \tag{19}$$

called the *tangent hyperplane* to V at A. The lines of this hyperplane that go through A and whose intersection multiplicity with V at A is at least two are said to be *tangent* to V at A.

A point $A = (a_1, \dots, a_n)$ of V such that $F_1'(a) = \cdots = F_n'(a) = 0$ is said to be *multiple* (or *singular*); such points form an algebraic subset of V, defined by $n+1$ equations. To study such a point more closely, we make it the origin; then $F_0 = F_1 = 0$ in (17). Let m be the smallest integer such that the homogeneous polynomial F_m is non-zero; m is called the *multiplicity* of A on V. Equation (18) becomes

$$F_m(b_1, \dots, b_m)t^m + \cdots + F_d(b_1, \dots, b_m)t^d = 0; \tag{20}$$

thus the intersection multiplicity of V and D at A is m unless D lies in the *tangent cone* of equation $F_m(y_1, \dots, y_m) = 0$.

Assume now that all the roots of (18) are in K (for example, if K is algebraically closed). If (18) has maximal degree, namely d, we can say that D and V have d common points, where each point P_j is counted with its intersection multiplicity m_j. But if (18) has degree less than d, because its highest coefficient $F_d(b_1, \dots, b_n)$ vanishes, it's no longer true that V and D have d common points. Where are the other points gone? To infinity, of course. Indeed, the relation $F_d(b_1, \dots, b_n) = 0$ implies that the point at infinity of D is in V (or rather, in $\hat{V}$); we say then that the direction of D is an *asymptotic direction* of V.

Examples. The intersection of the plane curve $x^4 - y^4 - xy = 0$ with the line $y = bx$ is determined by the equation $(1 - b^4)x^4 - bx^2 = 0$. This equation has $x = 0$ as a double root, that is, the origin is a double point. The degree drops to 2 for $b = 1, -1, i, -i$, which are the slopes of the asymptotic directions of the curves.

The surface $x^2 + y^3 + z^5 = 0$ has only one multiple point in affine space, the origin: if the partial derivatives $2x$, $3y^2$, $5z^4$ all vanish we have $x = y = z = 0$ in characteristic $\neq 2, 3, 5$, and in characteristic 2, 3 or 5 two of the coordinates are zero, hence so is the third because of the equation of the surface. Its asymptotic directions, those along which the highest-degree term vanishes, are the directions contained in the plane $z = 0$. The intersection $z = t = 0$ of this plane with the plane at infinity is the part at infinity of the projective closure of the surface. All points in this intersection are singular (in the closure); to see this, one can observe that the degree of (18) drops by 2 in all directions such that $z = 0$, or else write the equation of the surface in the affine patches $x \neq 0$ and $y \neq 0$, say, and take partial derivatives (the equation in the patch $y \neq 0$, for example, is $t^2 + x^2t^3 + z^5 = 0$).

Thus we're led to consider, in a projective space P, the intersection of a hypersurface V of homogeneous equation $G(x_0, \dots, x_n) = 0$ with a line D of parametric equation

$$x_i = c_i u + d_i v \qquad \text{for } i = 0, \dots, n,$$

where the "parameter" (u, v) in K^2 is to be understood projectively, that is, $(u, v) \neq (0, 0)$, and two proportional pairs parametrize the same point. The intersection of V and D is governed by the equation

$$G(c_0u + d_0v, \dots, c_nu + d_nv) = 0. \tag{21}$$

This is a homogeneous equation of degree $d = \deg G$ in u and v. Replacing K, if necessary, by an algebraic extension, we can write the left-hand side of (21) as a product of linear factors in (u, v), as follows: factor out u^k, for $k \geq 0$ maximal, then solve the equation obtained by making $u = 1$; each root e_j of this equation yields a factor $v - e_j u$ of (21). Counting each factor with its exponent, we obtain d solutions (u, v), each of which can be put in the form $(0, 1)$ or $(1, e_j)$. Thus we obtain exactly d points common to V and D. In the old literature this is expressed by saying that V and D have d common points, "real or imaginary" (replace K by its algebraic closure), "distinct or not" (count multiplicities), "at finite distance or at infinity" (replace the affine hypersurface by its closure).

Thus we see where the "disappearing" intersection points go when equation (18) drops from degree d to degree $d - k$. The idea is to take G above to be the homogeneous polynomial associated with F, and the c_i and d_i $(i = 1, \dots, n)$ to describe the same line D whose affine representation is $y_i = a_i + b_i t$: writing $x_0 = u$ and $x_i = a_i u + b_i v$, for example, we get $c_0 = 1$, $d_0 = 0$, $c_i = a_i$, $d_i = b_i$. Then (21) becomes

$$G(u, a_1u + b_1v, \dots, a_nu + b_nv) = 0. \tag{22}$$

Upon setting $v = tu$ this equation becomes $u^n G(1, a_1 + b_1t, \dots, a_n + b_nt) = 0$, which reduces to (18) if $u \neq 0$. The $d - k$ roots $t_j \in K$ of (18) account for the $d - k$ factors $v - t_j u$ in the left-hand side of (22); but there are also k factors u, corresponding to the point at infinity of D, counted k times. Notice that the intersection multiplicities are the same in the affine and the projective cases.

Finally, let's spell out the projective version of the notions of simple points and tangent hyperplanes. Assume that a point A, with homogeneous coordinates $(c_0, \dots, c_n)$, lies on V, so that the coefficient of u^d in the left-hand side of (21) is zero. D and V intersect at A with multiplicity one if and only if v is a simple factor in the left-hand side of (21), if and only if the coefficient of $u^{d-1}v$ is non-zero. By Taylor's formula, this coefficient is $d_0G_0'(c) + \cdots + d_nG_n'(c)$, where $G_i'(c)$ is the i-th partial derivative of G evaluated at $(c_0, \dots, c_n)$.

Now A is a simple point of V if any line intersecting V at A does so with multiplicity one; by the previous paragraph, this happens if and only if at

least one of the $G_i'(c)$ is non-zero. Thus the multiple points of V are defined by the $n+2$ equations $G(x) = G_i'(x) = 0$. If A is a simple point of V the *tangents* to V at A (that is, the lines whose intersection with V at A has multiplicity at least 2) are characterized by belonging to the hyperplane of equation

$$x_0 G_0'(x) + x_1 G_1'(x) + \cdots + x_n G_n'(x) = 0, \tag{23}$$

the *tangent hyperplane* to V at A.

Remark. By *Euler's formula* $d_0 G_0'(c) + \cdots + d_n G_n'(c) = dG(x)$, a point where all the partial derivatives of G vanish is on V if d is not a multiple of the characteristic of K.

If $G(x_0, \ldots, x_n)$ is obtained by homogenizing $F(y_1, \ldots, y_n)$, it is easy to see that $G_i'(1, y_1, \ldots, y_n) = F_i'(y_1, \ldots, y_n)$ for $i = 1, \ldots, n$, whence

$$G_0'(1, y_1, \ldots, y_n) = dF(y) - y_1 F_1'(y) - \cdots - y_n F_n'(y).$$

Applying this formula to a simple point $A = (1, a_1, \ldots, a_n)$ of V we see that, since $F(a) = 0$, equation (23) reduces to the affine equation for the tangent hyperplane (19).

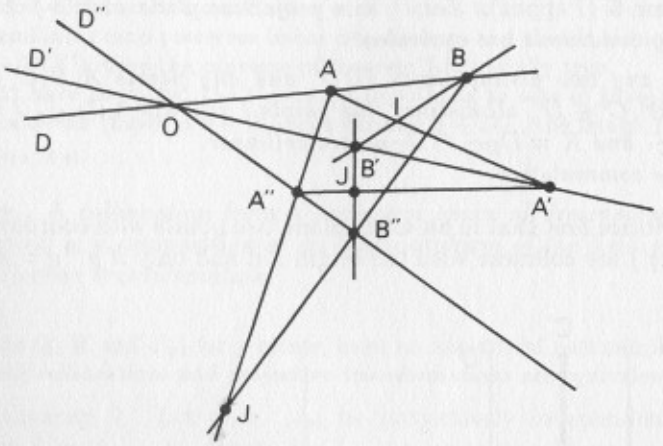
Three important theorems.

We will often use phrases borrowed from elementary geometry, such as "draw the line passing through two points", "collinear points", "concurrent lines", "coplanar lines", and so on. The line passing through two (distinct) points of an affine or projective space will be denoted by D_{ab} or simply ab.

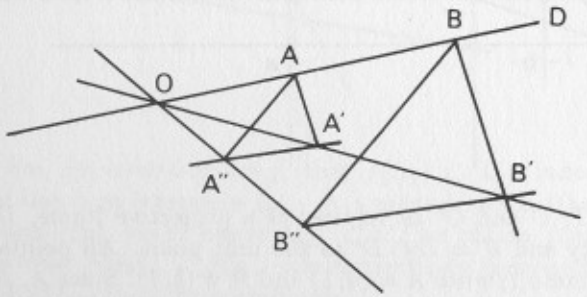
Theorem 5 (Desargues). *In a projective space P, let D, D' and D'' be distinct lines having a common point O. If $A, B \in D$, $A', B' \in D'$ and $A'', B'' \in D''$ are points distinct from one another and from O, the three intersection points $I = D_{AA'} \cap D_{BB'}$, $J = D_{AA''} \cap D_{BB''}$ and $K = D_{A'A''} \cap D_{B'B''}$ are collinear.*

Proof. These intersection points are well-defined: D and D', for example, lie on the same plane, so $D_{AA'}$ and $D_{BB'}$ also line on that plane; the two being distinct we can apply theorem 1 (section 1). To show collinearity, start with the case when the three lines D, D' and D'' are not coplanar. Then they generate a three-dimensional projective linear space, which contains the planes $AA'A''$ and $BB'B''$. Again by theorem 1, these two planes must have a line in common, which contains I, J and K.

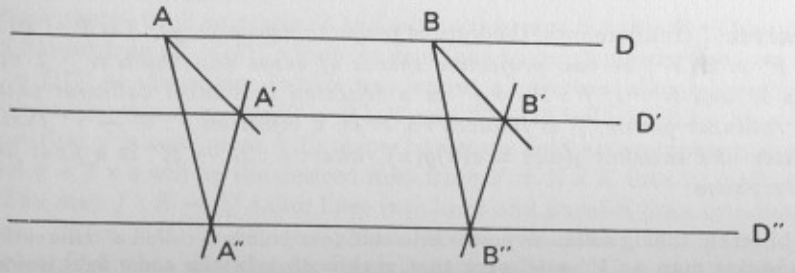
The case when D, D' and D'' are coplanar follows by projection, but we will give a direct proof. In the plane of the three lines, let the line at infinity be D_{IJ}, and assume first that $O \notin D_{IJ}$. Let the origin be O. Looking at $A, B, \ldots, B''$ as vectors, we can find scalars $a, a', a'' \in K$ such that $B = aA$, $B' = a'A'$ and $B'' = a''A''$. Since I is at infinity, AA' and BB' are parallel, so there exists $c \in K$ such that $B' - B = c(A' - A)$, that



is, $a'A' - aA = cA' - cA$; this implies $a' = a = c$ because A and A' are linerarly independent. Similarly $a = a''$. But then $B'' - B' = a(A'' - A')$, which shows that $D_{A'A''}$ and $D_{B'B''}$ are parallel, that is, their intersection K is on the line at infinity D_{IJ}.



Finally, if $O \in D_{IJ}$, all three lines D, D' and D'' are parallel and $ABB'A'$ and $ABB''A''$ are parallelograms. The translation $B - A$ takes A' to B' and A'' to B'', so $D_{A'A''}$ and $D_{B'B''}$ are parallel, and again K is at infinity. □

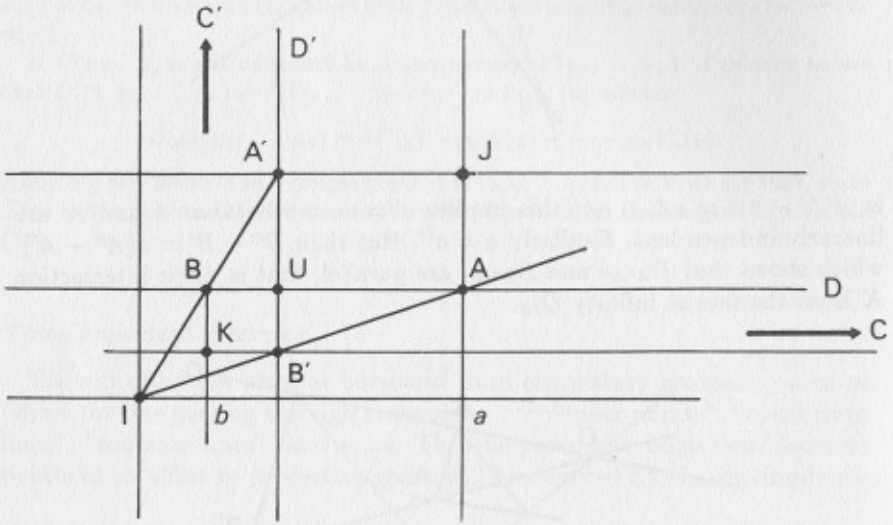


Theorem 6 (Pappus). *Let P be a projective plane over a field K. The following conditions are equivalent:*

(1) *For any two distinct lines D, D' and any points $A, B, C \in D$ and $A', B', C' \in D'$, all distinct, the points $I = D_{AB'} \cap D_{BA'}$, $J = D_{CA'} \cap D_{AC'}$ and $K = D_{BC'} \cap D_{CB'}$ are collinear.*

(2) *K is commutative.*

Proof. Notice first that in an affine plane two points with coordinates (p, q) and (p', q') are collinear with the origin I if and only if $p^{-1}q = p'^{-1}q'$.



Now take I, C and C' as vertices of a projective frame, $D_{CC'}$ as the line at infinity and $U = D \cap D'$ as the unit point. All points in D have second coordinate 1; write $A = (a, 1)$ and $B = (b, 1)$. Since A, B' and I are collinear and B' has first coordinate 1, we have $B' = (1, a^{-1})$. Similarly, the coordinates of A' are $(1, b^{-1})$. Thus J has coordinates (a, b^{-1}) and K has coordinates (b, a^{-1}); they are collinear with the origin if and only if $a^{-1}b^{-1} = b^{-1}a^{-1}$, if and only if $ab = ba$. Since $a, b \neq 0$ are arbitrary I, J, K are always collinear if and only if K is commutative. □

Theorem 7 (fundamental theorem of projective geometry). *Let $P = \mathbf{P}(V)$ and $P' = \mathbf{P}(V')$ be two projective spaces of same dimension $n \geq 2$ over fields K and K'. If $f : P \to P'$ is a bijection that takes collinear points into collinear points, f is induced on P by a bijection $g : V \to V'$ that is additive and satisfies $g(ax) = s(a)g(x)$, where $s : K \to K'$ is a fixed field isomorphism.*

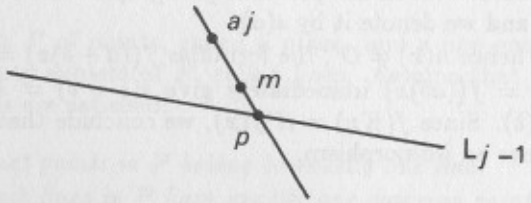
A bijection taking collinear points into collinear points is called a *collineation*. An additive map $g : V \to V'$ such that $g(x) = s(a)g(x)$ for some field isomorphism $s : K \to K'$ is called *semilinear* (with respect to to s). It is clear that a bijective semilinear map preserves linear dependence, and that it induces a map $f : \mathbf{P}(V) \to \mathbf{P}(V')$; thus the converse of theorem 7 is trivially true.

We could have restricted the theorem's hypothesis to sets of three collinear points, because for (fixed) $a, b \in A$ and a variable $x \in D_{ab}$, the image $f(x)$ is on the line $D_{f(a),f(b)}$.

Corollary. *A collineation from a projective space of dimension at least two into itself is a composition of an automorphism of the field of scalars with a projective transformation.* □

The fields $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{F}_p$, for p prime, have no non-trivial automorphisms, so for such field collineations and projective transformations are equivalent.

Proof of theorem 7. Let $a_0, \ldots, a_n$ be projectively independent points in P. For $j = 0, \ldots, n$, denote by L_j the projective linear space generated by $a_0, \ldots, a_j$ and by L'_j the projective linear space generated by $f(a_0), \ldots, f(a_j)$.

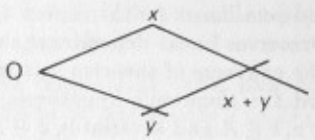


First we see, by induction on j, that $f(L_j) \subset L'_j$: indeed, for every $m \in L_j$, the line $a_j m$ intersects L_{j-1} at a point p; since $f(m)$ is collinear with $f(a_j)$ and $f(p)$, which lies in $L'_{j-1} \subset L'_j$, we obtain $f(m) \in L'_j$. This also implies that $f(a_0), \ldots, f(a_n)$ are projectively independent, because $P' = f(P) = f(L_n) \subset L'_n$ by surjectivity.

On the other hand we have $f(L_j) \supset L'_j$, because f is surjective, and points $m \in P \setminus L_j$ are mapped outside L'_j (complete the set $(a_0, \ldots, a_j, m)$ into a set of $n+1$ projectively independent points; by the previous paragraph $f(a_0), \ldots, f(a_j), f(m)$ will be projectively independent). Thus we conclude that $f(L_j) = L'_j$.

Now choose in P an origin O and a hyperplane at infinity H. Then $f(H)$ is a hyperplane H' of P', which we also take to be at infinity; and we take $O' = f(O)$ as the origin. In this way we get a bijection, also denoted by f, between the vector spaces $E = P \setminus H$ and $E' = P' \setminus H'$; we'll be done if we show that f is semilinear with respect to some field automorphism s, since then $g = f \times s$ will be the desired map from $V = E \times K$ into $V' = E' \times K'$.

The map $f : E \to E'$ takes lines into lines and parallel lines into parallel lines. Since $f(O) = O'$, the parallelogram rule shows that $f(x + y) = f(x) + f(y)$ when x and y are linearly independent.



Otherwise we have $y \in Kx$ and, since we assumed $\dim E \geq 2$, we may we take a point $z \notin Kx$, which will be linearly independent of x, y and $x+y$. In addition, $y+z$ is linearly independent of x. In this case, too, the additivity of f is verified, because

$$f(x+y+z) = f\big((x+y)+z\big) = f(x+y) + f(z) = f\big(x+(y+z)\big)$$
$$= f(x) + f(y+z) = f(x) + f(y) + f(z).$$

For $a \in K$ and $x \neq O$, the points O, x and ax are collinear, hence so are O, $f(x)$ and $f(ax)$. Thus there exists $s(a,x) \in K'$ such that $f(ax) = s(a,x)f(x)$. If x and y are linearly independent, so are $f(x)$ and $f(y)$, and we see, by calculating $f\big(a(x+y)\big)$ in two ways, that $s(a,x) = s(a,y) = s(a,x+y)$; this can be checked for x and y linearly dependent as well, using an auxiliary vector z as in the previous paragraph. Thus $s(a,x)$ does not depend on x, and we denote it by $s(a)$.

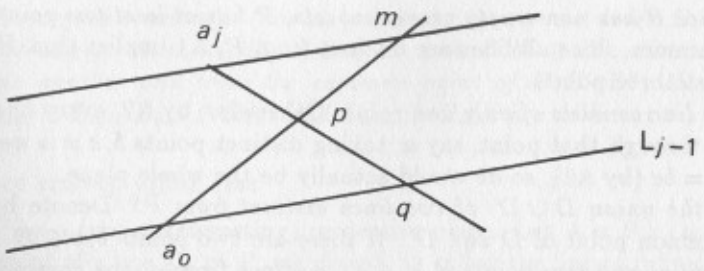
For $x \neq O$, hence $h(x) \neq O'$, the formulas $f\big((a+b)x\big) = f(ax) + f(bx)$ and $f\big(a(bx)\big) = f\big((ab)x\big)$ immediately give $s(a+b) = s(a) + s(b)$ and $s(ab) = s(a)s(b)$. Since $f(Kx) = K'f(x)$, we conclude that $s : K \to K'$ is surjective, hence an isomorphism. □

Remark on the affine analogue of theorem 7.

Let $f : A \to A'$ be a bijection taking triples of collinear points into triples of collinear points, where A and A' are affine spaces of same dimension $n \geq 2$, over K and K', respectively. When $K = \mathbf{F}_2$, this assumption is vacuous, because lines have only two elements. However, if $K \neq \mathbf{F}_2$, one can show that f is semilinear (as a map from the vectorialization of A at an arbitrary point O to the vectorialization of A at $O' = f(O)$).

The proof is very similar to that of theorem 7. One takes affinely independent points $a_0, \ldots, a_n \in A$, and, denoting by L_j and L'_j the affine subspaces generated by $a_0, \ldots, a_j$ and $f(a_0), \ldots, f(a_j)$, respectively, one shows by induction over j that $f(L_j) \subset L'_j$. For $m \in L_j$, there is no difficulty if the line ma_j intersects L_{j-1}. If not, ma_j is parallel to some direction in L_{j-1}; one then takes an auxiliary point $p \in D_{ma_0}$, so the line a_jp is not parallel to L_{j-1}. Setting $q = a_jp \cap L_{j-1}$, one concludes from the collinearity of a_j, p and q that $f(p) \in L'_j$, and from the collinearity of m, p and a_0 that $f(m) \in L'_j$.

There follows from the surjectivity of f, as in theorem 7, that $f(L_j) = L'_j$ for every j, and that f takes lines into lines and parallel lines into parallel lines. An application of the parallelogram rule and the same calculations as in theorem 7 complete the proof.



1.4. Axiomatic Presentation of Projective and Affine Planes

Incidence axioms: projective case

The fundamental theorem of projective geometry (theorem 7) hints that it is possible to reconstruct projective geometry from the notion of collinearity. That's just what we're going to do, axiomatically, in the case of the plane.

Consider a set P of points, called a *plane*, and a non-empty family of proper, non-empty subsets of P, called *lines*. Assume that the following *incidence axioms* are satisfied:

(A1) *Two distinct points in P belong to exactly one line.*
(A2) *Two distinct lines in P have exactly one common point.*

Remarks

(1) Notice the symmetry of the two assertions, which can be rephrased to say that "two points determine a unique line" and "two lines determine a unique point". We will come back to this topic in section 5, when we discuss duality. Notice also that A1 by itself already implies that two distinct lines have at most one common point, and similarly for A2.
(2) An axiomatic definition of n-dimensional projective spaces would involve $n-1$ families of non-empty, proper subsets of P, the j-dimensional projective linear subspaces of P for $j = 1, \ldots, n-1$, satisfying the following conditions: any $j+1$ points not contained in an $(j-1)$-dimensional projective linear space determine a unique j-dimensional projective linear space; any intersection of projective linear spaces is one; and the dimension of the intersection of two projective linear spaces is given by the formula in theorem 1, the notion of the projective linear space generated by a set making sense by the previous condition. This is all quite easy to write down explicitly in the case $n = 3$.

The following are immediate consequences of axioms A1 and A2. (We denote the line going through a and b by D_{ab} or ab.)

Franz Lemmermeyer, Conics — A Poor man's Elliptic Curves, preprint 2003

CONICS - A POOR MAN'S ELLIPTIC CURVES

FRANZ LEMMERMEYER

Contents

Introduction

The aim of this article is to show that the arithmetic of Pell conics admits a description which is completely analogous to that of elliptic curves: there is a theory of 2-descent with associated Selmer and Tate-Shafarevich groups, and there should be an analog of the conjecture of Birch and Swinnerton-Dyer. For the history and a theory of the first 2-descent, see [6, 7, 8]. The idea that unit groups of number fields and the group of rational points on elliptic curves are analogous is not new; see e.g. [1, 2, 5, 14] for some popularizations of this point of view. It is our goal here to show that, for the case of the unit group of real quadratic number fields, this analogy can be made much more precise.

1. The Group Law on Pell Conics and Elliptic Curves

Let $F \in \mathbb{Z}[X,Y]$ a polynomial. If $\deg F = 2$, the affine curve of genus 0 defined by $F = 0$ is called a conic. Let d be a squarefree integer $\ne 1$ and define

$$\Delta = \begin{cases} d & \text{if } d \equiv 1 \bmod 4, \\ 4d & \text{if } d \equiv 2,3 \bmod 4. \end{cases}$$

Then the curves $\mathcal{C} : X^2 - \Delta Y^2 = 4$ are called Pell conics; they are irreducible, nonsingular affine curves with a distinguished integral point $N = (2,0)$.

If $\deg F = 3$, the projective curve E described by F has genus 1 if it is nonsingular; if in addition it has a rational point, then E is called an elliptic curve defined over $\mathbb{Q}$. Elliptic curves given by a Weierstraß equation $Y^2 = X^3 + aX + b$ are irreducible, nonsingular projective curves with a distinguished integral point $\mathcal{O} = [0:1:0]$ at infinity.

Both types of curves have a long history: Pythagorean triples correspond to rational points on the Pell conic $X^4 + 4Y^2 = 4$, solutions of the Pell equations have been studied by the Greeks, the Indians, and the contemporaries of Fermat, such as Brouncker and Wallis. Problems leading to elliptic curves occur in the books of Diophantus and were studied by Bachet, Fermat, de Jonquières, Euler, Cauchy, Lucas, and Sylvester before Poincaré laid down his program for studying diophantine equations given by curves according to their genus.

1.1. Group Law on Conics.

The group law on non-degenerate conics C defined over a field F is quite simple: fix any rational point N on C; for finding the sum of two rational points $A, B \in C(F)$, draw the line through N parallel to AB, and denote its second point of intersection with C by $A + B$. In the special case of Pell conics, the resulting formulas can be simplified to

Proposition 1. *Consider the conic $\mathcal{C} : Y^2 - \Delta X^2 = 4$, and put $N = (2,0)$. Then the group law on $\mathcal{C}$ with neutral element N is given by*

$$(r,s) + (t,u) = \Big(\frac{rt + \Delta su}{2}, \frac{ru + st}{2}\Big).$$

It is now easily checked that the map sending points $(r,s) \in \mathcal{C}(\mathbb{Z})$ to the unit $\frac{r+s\sqrt{d}}{2}$ with norm 1 in the quadratic number field $K = \mathbb{Q}(\sqrt{\Delta})$ is a group homomorphism. Observe that the associativity of the geometric group law is equivalent to a special case of Pascal's theorem, which in turn is a very special case of Bezout's Theorem.

1.2. **Group Law on Elliptic curves.** Given an elliptic curve $E : y^2 = x^3 + ax + b$ defined over an algebraically closed field K, we define an addition law on E by demanding that $A + B + C = 0$ for points $A, B, C \in E(K)$ if and only if A, B, C are collinear. Since vertical lines intersect E only in two affine points, we have to regard E as a projective curve; then vertical lines intersect E in two affine points as well as in the point at infinity. Associativity follows geometrically from a special case of Bezout's Theorem.

2. The Group Structure

Let us now compare the known results about the group structure of Pell conics over the most common rings and fields. Generally, we will study conics in the affine plane over integral domains, and elliptic curves in the projective plane over fields.

2.1. **Finite Fields.** Let $\mathcal{C} : x^2 - \Delta y^2 = 4$ be a Pell conic defined over a finite field $\mathbb{F}_q$ with $q = p^f$ elements, and assume that $\mathcal{C}$ is smooth, i.e. that $p \nmid \Delta$. Then

$$\mathcal{C}(\mathbb{F}_q) \simeq \mathbb{Z}/m\mathbb{Z}, \quad \text{where } m = q - \left(\frac{\Delta}{p}\right)^f.$$

If Δ is a square mod p and p is odd, this is immediately clear since there is an affine isomorphism between $\mathcal{C}$ and the hyperbolas $X^2 - Y^2 = 1$ and $XY = 1$; in particular, one has $\mathcal{C}(\mathbb{F}_q) \simeq \mathbb{F}_q^\times = \mathrm{GL}_1(\mathbb{F}_q)$ in this case.

On the elliptic curve side, we know that

$$E(\mathbb{F}_q) \simeq \mathbb{Z}/n_1\mathbb{Z} \oplus \mathbb{Z}/n_2\mathbb{Z}, n_2 \mid n_1,$$

Moreover, we have $\#E(\mathbb{F}_p) = (p+1) - a_p$, where $|a_p| \leq 2\sqrt{p}$ by Hasse's theorem.

2.2. **p-adic Numbers.** If p is an odd prime not dividing Δ, then

$$\mathcal{C}(\mathbb{Z}_p) \simeq \begin{cases} \mathbb{Z}/(p-1) \oplus \mathbb{Z}_p & \text{if } (\frac{\Delta}{p}) = +1, \\ \mathbb{Z}/(p+1) \oplus \mathbb{Z}_p & \text{if } (\frac{\Delta}{p}) = -1, \\ \mathbb{Z}/2 \oplus \mathbb{Z}_p & \text{if } p \mid \Delta \neq -3, \\ \mathbb{Z}/6 \oplus \mathbb{Z}_p & \text{if } p = 3,\ \Delta = -3. \end{cases}$$

Reduction modulo p^k then yields

$$\mathcal{C}(\mathbb{Z}/p^k) \simeq \begin{cases} \mathbb{Z}/(p-1) \oplus \mathbb{Z}/p^{k-1} & \text{if } (\frac{\Delta}{p}) = +1, \\ \mathbb{Z}/(p+1) \oplus \mathbb{Z}/p^{k-1} & \text{if } (\frac{\Delta}{p}) = -1, \\ \mathbb{Z}/2 \oplus \mathbb{Z}/p^k & \text{if } p \mid \Delta \neq -3, \\ \mathbb{Z}/6 \oplus \mathbb{Z}/3^{k-1} & \text{if } p = 3, \Delta = -3. \end{cases}$$

For elliptic curves $E/\mathbb{Q}_p$ we have a reduction map sending $\mathbb{Q}_p$-rational points to points defined over $\mathbb{F}_p$. The group $E_{ns}(\mathbb{F}_p)$ is the set of all nonsingular points of E over $\mathbb{F}_p$. The subgroups $E_i(\mathbb{Q}_p)$ $(i = 0, 1)$ of $E(\mathbb{Q}_p)$ are defined as the inverse images of $E_{ns}(\mathbb{F}_p)$ and of the point of infinity of $E(\mathbb{F}_p)$ under the reduction map. These groups sit inside the exact sequence

$$0 \longrightarrow E_1(\mathbb{Q}_p) \longrightarrow E_0(\mathbb{Q}_p) \longrightarrow E_{ns}(\mathbb{F}_p) \longrightarrow 0.$$

The structure of $E_{ns}(\mathbb{F}_p)$ is known: if $E/\mathbb{F}_p$ is nonsingular, it was discussed in Subsection 2.1; if $E/\mathbb{F}_p$ is singular, then $E_{ns}(\mathbb{F}_p)$ is isomorphic to $\mathcal{C}(\mathbb{F}_p)$ for a certain conic $\mathcal{C}$, and we say that E has additive, multiplicative or split multiplicative

reduction if the conic is a parabola ($\mathcal{C}(\mathbb{F}_p) \simeq \mathbb{F}_p$), a hyperbola ($\mathcal{C}(\mathbb{F}_p) \simeq \mathbb{F}_p^\times$), or an ellipse ($\mathcal{C}(\mathbb{F}_p) \simeq \mathbb{F}_{p^2}[1]$, the group of elements with norm 1 in $\mathbb{F}_{p^2}$).

We also know hat $E_1(\mathbb{Q}_p) \simeq \mathbb{Z}_p$ and that the quotient group $E(\mathbb{Q}_p)/E_0(\mathbb{Q}_p)$ is finite. Its order c_p is called the Tamagawa number for the prime p, and clearly $c_p = 1$ we have for all primes $p \nmid \Delta$. More exactly it can be shown (albeit with some difficulty) that $c_p \le 4$ if E has additive reduction, and that c_p is the exact power of p dividing Δ otherwise.

2.3. **Integral and Rational Points.** Now let us compare the structure of the groups of rational points: for elliptic curves, we have the famous theorem of Mordell-Weil that $E(\mathbb{Q}) \simeq E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$, where $E(\mathbb{Q})_{\text{tors}}$ is the finite group of points of finite order, and r is the Mordell-Weil rank. For conics, on the other hand, we have two possibilities: either $C(\mathbb{Q}) = \varnothing$ (for example if $C : x^2 + y^2 = 3$) or $C(\mathbb{Q})$ is infinite, and in fact not finitely generated (see Tan [12]). The analogy can be saved, however, by looking at integers instead of rational numbers: if K is a number field with ring of S-integers $\mathcal{O}_S$, then

$$C(\mathcal{O}_S) \simeq C(\mathcal{O}_S)_{\text{tors}} \oplus \mathbb{Z}^r \qquad\qquad E(K) \simeq E(K)_{\text{tors}} \oplus \mathbb{Z}^r$$

where $r \ge 0$ is called the Mordell-Weil rank. Shastri [10] computed the rank r for the unit circle over number fields K and $S = \varnothing$.

Note that the group of integral points on the hyperbola $XY = 1$ is isomorphic to $R^\times = \mathrm{GL}_1(R)$. Number theoretic algorithms working with the multiplicative group of $R = \mathbb{Z}/p\mathbb{Z}$ in general have an analog for conics, as we will see in the next section.

3. Applications

3.1. **Primality Tests.** The classical primality test due to Lucas is the following:

Proposition 2. *An odd integer n is prime if and only if there exists an integer a satisfying the following two conditions:*

i) $a^{n-1} \equiv 1 \bmod n$;
ii) $a^{(n-1)/r} \not\equiv 1 \bmod n$ *for every prime* $r \mid (n-1)$.

This primality test is based on the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^\times$, that is, on the group $\mathcal{H}(\mathbb{Z}/n\mathbb{Z})$ of $\mathbb{Z}/n\mathbb{Z}$-rational points on the hyperbola $\mathcal{H} : XY = 1$. Something similar works for any Pell conic:

Proposition 3. *Let $n \ge 5$ be an odd integer and $\mathcal{C} : X^2 - \Delta Y^2 = 4$ a nondegenerate Pell conic defined over $\mathbb{Z}/n\mathbb{Z}$ with neutral element $N = (2, 0)$, and assume that $(\Delta/n) = -1$. Then n is a prime if and only if there exists a point $P \in \mathcal{C}(\mathbb{Z}/n\mathbb{Z})$ such that*

i) $(n+1)P = N$;
ii) $\frac{n+1}{r} P \ne N$ *for any prime r dividing $n+1$.*

Of course, for both tests there are 'Proth-versions' in which only a part of $N \pm 1$ needs to be factored.

The following special case of Proposition 3 is well known: if $n = 2^p - 1$ is a Mersenne number, then $n \equiv 7 \bmod 12$ for $p \ge 3$, hence $(3/n) = -1$; if we choose the Pell conic $\mathcal{C} : X^2 - 12Y^2 = 4$ and and $P = (4, 1)$, then the test above is nothing

but the Lucas-Lehmer test. We remark in passing that Gross [3] has come up with a primality test for Mersenne numbers based on elliptic curves.

3.2. **Factorization Methods.** The factorization method based on elliptic curves is very well known. Can we replace the elliptic curve by conics? Yes we can, and what we get is the $p-1$-factorization method for integers N if we consider the conic $\mathcal{H}: xy = 1$, and some $p \pm 1$-factorization method for general Pell conics. The details are easy to work out for anyone familiar with Pollard's $p-1$-method.

4. 2-Descent

Consider the Pell conic $\mathcal{C}: X^2 - \Delta Y^2 = 4$. Define a map $\alpha : \mathcal{C}(\mathbb{Q}) \longrightarrow \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ by

$$\alpha(x,y) = \begin{cases} (x+2)\mathbb{Q}^{\times 2} & \text{if } x \neq -2, \\ -\Delta\mathbb{Q}^{\times 2} & \text{if } x = -2. \end{cases}$$

If $P = (x,y) \in \mathcal{C}(\mathbb{Z})$ with $x > 0$, then P gives rise to an integral point on the descendant $\mathcal{T}_a(\mathcal{C}) : aX^2 - bY2 = 4$, where a is a positive squarefree integer determined by $\alpha(P) = a\mathbb{Q}^{\times 2}$, and $ab = \Delta$. Conversely, any integral point on some $\mathcal{T}_a(\mathcal{C})$ gives rise to an integral point with positive x-coordinate on the Pell conic $\mathcal{C}$.

It can be shown that α is a group homomorphism, and that we have an exact sequence

$$0 \longrightarrow 2\mathcal{C}(\mathbb{Z}) \longrightarrow \mathcal{C}(\mathbb{Z}) \xrightarrow{\alpha} \mathbb{Q}^\times/\mathbb{Q}^{\times 2}.$$

Moreover, we have $\#\mathrm{im}\,\alpha = 2^r$, where r is the Mordell-Weil-rank of $C(\mathbb{Z})$, and the elements of $\mathrm{im}\,\alpha$ are represented by the first descendants $\mathcal{T}_a$ with $\mathcal{T}_a(\mathbb{Z}) \neq \varnothing$. Thus computing the Mordell-Weil rank is equivalent to counting the number of first descendants $\mathcal{T}_a$ with an integral point (see [8]).

The situation is completely analogous for elliptic curves $E: Y^2 = X(X^2+aX+b)$ with a rational point $(0,0)$ of order 2, except that here we also have to consider the 2-isogenous curve $\widehat{E}: Y^2 = X(X^2 + \widehat{a}\,X + \widehat{b})$, where $\widehat{a} = -2a$ and $\widehat{b} = a^2 - 4b$. We have two Weil maps $\alpha : E(\mathbb{Q}) \longrightarrow \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ and $\widehat{\alpha} : \widehat{E}(\mathbb{Q}) \longrightarrow \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$, and the Mordell-Weil rank is given by Tate's formula $2^{r+2} = \#\mathrm{im}\,\alpha \cdot \#\mathrm{im}\,\widehat{\alpha}$. For more information on the descent via 2-isogenies we refer to Silverman & Tate [11].

4.1. **Selmer and Tate-Shafarevich Group.** The subset of descendants $\mathcal{T}_a : ar^2 - bs^2 = 4$ with a rational point form a subgroup $\mathrm{Sel}_2(C)$ of $\mathbb{Q}^\times/\mathbb{Q}^{\times 2}$ called the 2-Selmer group of C. Next we define the Tate-Shafarevich group $\mathbf{Ш}_2(C)$ by the exact sequence

$$1 \longrightarrow \mathrm{im}\,\alpha \longrightarrow \mathrm{Sel}_2(C) \longrightarrow \mathbf{Ш}_2(C) \longrightarrow 1.$$

In [8] we have shown that the 2-part of the Tate-Shafarevich group of the Pell conic $\mathcal{C}: X^2 - \Delta Y^2 = 4$ is $\mathbf{Ш}_2(\mathbb{Z}) \simeq \mathrm{Cl}^+(k)^2[2]$.

For a cohomological definition of Selmer and Tate-Shafarevich groups, we need to interpret conics as principal homogeneous spaces. Every conic $X^2 - \Delta Y2 = 4c$ is a principal homogeneous space for $\mathcal{C}(\mathbb{Q})$; this is to say that the map

$$\mu : \mathcal{D}(\mathbb{Z}) \times \mathcal{C}(\mathbb{Z}) \longrightarrow \mathcal{D}(\mathbb{Z}) : \mu((u,v),(x,y)) = (\tfrac{ux+\Delta vy}{2}, \tfrac{vx+uy}{2}).$$

has the following properties:

(1) $\mu(p,N) = p$ for all $p \in \mathcal{D}(\overline{\mathbb{Q}})$, where $N = (2,0)$ is the neutral element of $\mathcal{C}$.
(2) $\mu(\mu(p,P),Q) = \mu(p,P+Q)$ for all $p \in \mathcal{D}(\overline{\mathbb{Q}})$ and all $P,Q \in \mathcal{C}(\overline{\mathbb{Q}})$.

(3) For all $p, q \in \mathcal{D}(\mathbb{Q})$ there is a unique $P \in \mathcal{C}(\mathbb{Q})$ with $\mu(p, P) = q$.

Here $\overline{\mathbb{Q}}$ denotes the algebraic closure of $\mathbb{Q}$.

Note, however, that only those $\mathcal{D}$ with $c \mid \Delta$ are principal homogeneous space for $\mathcal{C}(\mathbb{Z})$, i.e., satisfy the property that for all $p, q \in \mathcal{D}(\mathbb{Z})$ there is a $P \in \mathcal{C}(\mathbb{Z})$ with $\mu(p, P) = q$. Also observe that the conics $\mathcal{D}$ with $c \mid \Delta$ can be written in the form $aX^2 - bY^2 = 4$ with $ab = \Delta$, that is, these are exactly the first descendants.

4.2. **Heights.** For a rational number $q = \frac{m}{n}$ in lowest terms, define its height $H(q) = \log \max\{|m|, |n|\}$; note that $H(0) = 0$ and $H(q) \ge 0$ for all $q \in \mathbb{Q}$. For rational points $P = (x, y) \in C(\mathbb{Q})$ on a conic $C : X^2 - \Delta Y^2 = 4$ put $H(P) = H(x)$.

Define the canonical height $\widehat{h}(P)$ by

$$\widehat{h}(P) = \lim_{n\to\infty} \frac{H(2^n P)}{2^n}.$$

The canonical height $\widehat{h}$ on the Pell conic $\mathcal{C} : X^2 - \Delta Y^2 = 4$ has all the suspected properties (and more):

(1) $|\widehat{h}(P) - H(P)| < \log 4$;
(2) $\widehat{h}(T) = 0$ if and only if $T \in \mathcal{C}(\mathbb{Q})_{\text{tors}}$;
(3) $\widehat{h}(mP) = m\widehat{h}(P)$ for all integers $m \ge 1$;
(4) $\widehat{h}(P + Q) \le \widehat{h}(P) + \widehat{h}(Q)$;
(5) the square of the canonical height satisfies the parallelogram equality

$$\widehat{h}(P - Q)^2 + \widehat{h}(P + Q)^2 = 2\widehat{h}(P)^2 + 2\widehat{h}(Q)^2$$

for all $P, Q \in \mathcal{C}(\mathbb{Q})$.

In addition, there are explicit formulas for the canonical height. It is an easy exercise to show that every rational point on a Pell conic has the form $P = (x, y)$ with $x = \frac{r}{n}$, $y = \frac{s}{n}$, and $(r, n) = (s, n) = 1$. In this case we have

$$\widehat{h}(P) = \begin{cases} \log \frac{|r|+|s|\sqrt{\Delta}}{2} & \text{if } \Delta > 0, \\ \log |n| & \text{if } \Delta < 0. \end{cases}$$

The finiteness of $\mathcal{C}(\mathbb{Z}_S)/2\mathcal{C}(\mathbb{Z}_S)$ and the existence of a height function implies the theorem of Mordell-Weil.

5. Analytic Methods

5.1. **Zeta Functions.** Both for conics and elliptic curves over $\mathbb{Q}$ there is an analytic method that sometimes provides us with a generator for the group of integral or rational points on the curve. Before we can describe this method, we have to talk about zeta functions of curves.

Take a conic C or an elliptic curve E defined over the finite field $\mathbb{F}_p$; let N_r denote the cardinalities of the groups of $\mathbb{F}_{p^r}$-rational points on C and E respectively, where we count solutions in the affine plane for C and in the projective plane for E. Then

$$Z_p(T) = \exp\Big(\sum_{r=1}^{\infty} N_r \frac{T^r}{r}\Big)$$

is called the zeta function of C or E over $\mathbb{F}_p$.

For the parabola $C : y = x^2$, we clearly have $C(\mathbb{F}_q) \simeq \mathbb{F}_q$, hence $N_r = p^r$, and we find

$$Z_p(T) = \exp\Big(\sum_{r=1}^{\infty} p^r \frac{T^r}{r}\Big) = \exp(-\log(1-pT)) = \frac{1}{1-pT}.$$

For the conic $X^2 - \Delta Y^2 = 4$ we find after a little calculation

$$Z_p(T) = \frac{1}{(1-pT)(1-\chi(p)T)},$$

where χ is the Dirichlet character defined by $\chi(p) = (\Delta/p)$. The substitution $T = p^{-s}$ turns this into

$$\zeta_p(s;\mathcal{C}) = \frac{1}{(1-p^{1-s})(1-\chi(p)p^{-s})}.$$

For nonsingular elliptic curves over $\mathbb{F}_p$ we similarly get

$$Z_p(T) = \frac{P(T)}{(1-T)(1-pT)},$$

where $P(T) = qT^2 - a_pT + 1$ and a_p is defined by $\#E(\mathbb{F}_p) = p + 1 - a_p$.

5.2. **L-Functions for Conics.** Now we take the zeta function for each p and multiply them together to get a global zeta function. The first factor $1/(1-p^{1-s})$ gives us the product

$$\prod_{p\,\text{odd prime}} \frac{1}{1-p^{1-s}} = \zeta(s-1)(1-2^{1-s}),$$

that is, essentially the Riemann zeta function.

The other factor, on the other hand, is more interesting:

$$L(s,\chi) = \prod_p \frac{1}{1-\chi(p)p^{-s}}$$

is a Dirichlet L-function for the quadratic character $\chi = (\Delta/\cdot)$. This function converges on the right half plane $\operatorname{Re} s > 1$ and can be extended to a holomorphic function on the complex plane.

Now the nice thing discovered by Dirichlet (in his proof that every arithmetic progression $ax+b$ with $(a,b) = 1$ contains infinitely many primes) is that, for every nontrivial (quadratic) character χ, $L(s,\chi)$ has a nonzero value at $s = 1$. In fact, he was able to compute this value:

$$L(1,\chi) = \begin{cases} h \cdot \frac{2\pi}{w\sqrt{|\Delta|}} & \text{if } \Delta < 0, \\ h \cdot \frac{2\log\varepsilon}{\sqrt{\Delta}} & \text{if } \Delta > 0 \end{cases}$$

where $\chi(p) = (\Delta/p)$, and where w, Δ, h and $\varepsilon > 1$ are the number of roots of unity, the discriminant, the class number and the fundamental unit of $\mathbb{Q}(\sqrt{\Delta}\,)$.

The upshot is this: if $\Delta > 0$, the group $\mathcal{C}(\mathbb{Z})$ has rank 1; by using only local information (numbers of $\mathbb{F}_{p^r}$-rational points on $\mathcal{C}$) we have constructed a function whose value at 1 gives, up to well understood constants, a power of a generator of $\mathcal{C}(\mathbb{Z})$, namely the h-th power of the fundamental unit.

The functional equation of Dirichlet's L-function allows us to rewrite Dirichlet's formula as

$$\lim_{s\to 0} s^{-r}L(s,\chi) = \frac{2hR}{w},$$

where $r=0$ and $R=1$ for $\Delta<0$, and $r=1$ and $R=\log\varepsilon$ for $\Delta>0$.

Observe that the evaluation of the L-funtion (which was defined using purely local data) at $s=0$ yields a generator of the free part of the group $\mathcal{C}(\mathbb{Z})$ (which is a global object)!

5.3. **L-Functions for Elliptic Curves.** The really amazing thing is that exactly the same thing works for elliptic curves of rank 1: by counting the number N_r of $\mathbb{F}_{p^r}$-rational points on E, we get a zeta function $Z_p(T)$ that can be shown to have the form

$$Z_p(T) = \frac{P(T)}{(1-T)(1-pT)}$$

for some polynomial $P(T)\in\mathbb{Z}[T]$ of degree 2 (if p does not divide the discriminant of E). In fact, if $p\nmid E$ we have $P(T)=1-a_pt+pt^2$, where $a_p=p+1-\#E(\mathbb{F}_p)$.

Put $L_p(s)=1/P(p^{-s})$ and define the L-function

$$L(s,E)=\prod_p L_p(s).$$

Hasse conjectured that this L-function can be extended analytically to the whole complex plane; moreover, there exists an $N\in\mathbb{N}$ such that

$$\Lambda(s,E) = N^{s/2}(2\pi)^{-s}\Gamma(s)L(s,E)$$

satisfies the functional equation $\Lambda(s-2,E)=\pm\Lambda(s,E)$ for some choice of signs. For curves with complex multiplication, this was proved by Deuring; the general conjecture is a consequence of the now proved Taniyama-Shimura conjecture.

6. Birch–Swinnerton-Dyer

6.1. **Birch and Swinnerton-Dyer for Elliptic Curves.** The conjecture of Birch and Swinnerton-Dyer for elliptic curves predicts that $L(s,E)$ has a zero of order r at $s=1$, where r is the rank of the Mordell-Weil group. More exactly, it is believed that

$$\lim_{s\to 1}\ (s-1)^rL(s;E) = \frac{\Omega\cdot\#\text{Ш}(E/\mathbb{Q})\cdot R(E/\mathbb{Q})\cdot\prod c_p}{(\#E(\mathbb{Q})_{\text{tors}})^2},$$

where r is the Mordell-Weil rank of $E(\mathbb{Q})$, $\Omega=c_\infty$ the real period, $\text{Ш}(E/\mathbb{Q})$ the Tate-Shafarevich group, $R(E/\mathbb{Q})$ the regulator of E (some matrix whose entries are canonical heights of basis elements of the free part of $E(\mathbb{Q})$), c_p the Tamagawa number for the prime p (trivial for all primes not dividing the discriminant), and $E(\mathbb{Q})_{\text{tors}}$ the torsion group of E.

6.2. **Birch and Swinnerton-Dyer for Conics.** We now want to interpret Dirichlet's class number formula in a similar way. Let $k=\mathbb{Q}(\sqrt{\Delta}\,)$ denote the quadratic number field associated to the Pell conic $\mathcal{C}: X^2-\Delta Y^2=4$. Then we conjecture that there is a cohomological definition of the Tate-Shafarevich group $\text{Ш}(\mathcal{C})$ whose 2-torsion coincides with the group $\text{Ш}_2(\mathcal{C})$ defined above, and that we have

$$\text{Ш}(\mathcal{C}) \simeq \mathrm{Cl}^+(k)^2.$$

If we (preliminarily) define the Tamagawa numbers by

$$c_p = \begin{cases} 2 & \text{if } p \mid \Delta, \\ 1 & \text{otherwise,} \end{cases}$$

then Gauss's genus theory implies that

$$\prod c_p = 2(\mathrm{Cl}^+(k) : \mathrm{Cl}^+(k)^2).$$

Thus if we put $\Omega = \frac{1}{2}$, then $\Omega \cdot \#\mathbf{Ш}(\mathcal{C}) \cdot \prod c_p = h^+$ equals the class number of k in the strict sense, hence is equal to $2^u \cdot h$, where $u = 1$ if $N\varepsilon = +1$, and $u = 0$ otherwise.

If $\Delta > 0$, let $\eta > 1$ denote a generator of the free part of $\mathcal{C}(\mathbb{Z})$; then the regulator of $\mathcal{C}$ equals $\widehat{h}(\eta) = \log \eta$. Now we find $R(\mathcal{C}) = 2^{1-u}R$, hence $\Omega \cdot \#\mathbf{Ш}(\mathcal{C}) \cdot R(\mathcal{C}) \cdot \prod c_p = h^+ \log \eta = 2hR$; this also holds for $\Delta < 0$ if we put $R = 1$.

Finally, $\mathcal{C}(\mathbb{Z})_{\mathrm{tors}}$ is the group of roots of unity contained in k, and we find

$$\frac{2hR}{w} = \frac{\Omega \cdot \#\mathbf{Ш}(\mathcal{C}) \cdot R(\mathcal{C}) \cdot \prod c_p}{\#\mathcal{C}(\mathbb{Z})_{\mathrm{tors}}}$$

in (almost) perfect analogy to the Birch–Swinnerton-Dyer conjecture for elliptic curves.

In fact, the analogy would be even closer if we would replace $\#\mathcal{C}(\mathbb{Z})_{\mathrm{tors}}$ by $(\#\mathcal{C}(\mathbb{Z})_{\mathrm{tors}})^2$ and adjust the formulas for c_2 and c_3 for the two Pell conics with nontrivial torsion; this would also allow us to put $\Omega = 1$.

7. Summary

The analogy between Pell conics and elliptic curves is summarized in the following table:

	GL_1	Pell conics	elliptic curves
group structure on	affine line	affine plane	projective plane
defined over	rings	rings	fields
group elements	S-units	S-integral points	rational points
group structure	$\mathbb{Z}/2 \oplus \mathbb{Z}^{\#S}$	$C(\mathbb{Z}_S)_{\mathrm{tors}} \oplus \mathbb{Z}^r$	$E(\mathbb{Q})_{\mathrm{tors}} \oplus \mathbb{Z}^r$
associativity	clear	Pascal's Theorem	Bezout's Theorem
factorization alg.	$p - 1$	$p \pm 1$	ECM
primality tests	Lucas-Proth	Lucas-Lehmer	ECPP
Ш	1	$\mathrm{Cl}^+(k)^2$	?
L-series	$\mathbb{Z}$	quadratic field	modular form

Moreover, cyclotomic fields are for Pell conics what modular curves are for elliptic curves, and cyclotomic units correspond to Heegner points. The analog of Heegner's Lemma (if a curve of genus 1 of the form $Y^2 = f_4(X)$, where f_4 is a quartic polynomial with rational coefficients, has a K-rational point for some number field K of odd degree, then the curve has a rational point; cf. [4]) is due to Nagell [9], who proved the same result with f_4 replaced by a quadratic polynomial f_2.

8. Questions

Although the arithmetic of conics is generally regarded as being almost trivial, there are a lot of questions that are still open. The main problem is a good definition of the Tamagawa numbers in the case of conics, a cohomological description of the Selmer and Tate-Shafarevich groups, and the proof of $\mathbf{Ш}(\mathcal{C}) \simeq \mathrm{Cl}^+(k)^2$.

The next problem is the analytic construction of generators of $\mathcal{C}(\mathbb{Z}_S)$ if $S \neq \varnothing$. This suggests looking at the Stark conjectures, which predict that we can construct certain units (actually S-units) in number fields. It seems, however, that we cannot hope to find "independent" elements (see [13]).

On a simpler level there's the question whether iterated 2-descents on Pell conics provide an algorithm for computing the fundamental unit that is faster than current methods. And how does 3-descent on Pell conics work?

We can also think of generalizing the approach described here: the groups GL_1 and the Pell conics are special norm tori in the theory of algebraic groups, and there's the question of how much of the above carries over to the more general situation. The norm-1 tori associated to pure cubic fields can be described geometrically as cubic surfaces $\mathcal{S}$; do the groups of integral points on $\mathcal{S}$ admit a geometric group law? It is known that the groups of rational points on cubic surfaces coming from norm forms satisfy the Hasse principle; is there a connection between the 3-class groups of these fields and the Tate-Shafarevich groups on $\mathcal{S}$ defined as above as the obstruction to lifting the Hasse principle from rational to integral points?

On the elliptic curve side, there are a few questions suggested by the analogy worked out in this article. For example, is there a natural group whose order equals $\#\mathbf{Ш}(E) \cdot \prod c_p$? Recall that $\exp(\widehat{h}(P))$ is algebraic for rational points on Pell conics; are there meromorphic functions F such that $F(\widehat{h}(P))$ is algebraic for rational points P on elliptic curves, at least for curves with complex multiplication?

Acknowledgments

This article owes a lot to work done while I was at the University of Seoul in August 2002; I would like to thank Soun-Hi Kwon for the invitation and the hospitality.

References

[1] H. Darmon, *Wiles' theorem and the arithmetic of elliptic curves*, in: Modular Forms and Fermat's Last Theorem, G. Cornell et al. (eds.), Springer Verlag 1997, 549–569; cf. p. 2

[2] H. Darmon, C. Levesque, *Sommes infinies, équations diophantiennes et le dernier théorème de Fermat*, Gazette des Sciences Mathématiques du Québec, Vol. XVIII, Avril 1996; cf. p. 2

[3] B. Gross, *An elliptic curve test for Mersenne primes*, preprint 2003; cf. p. 5

[4] K. Heegner, *Diophantische Analysis und Modulfunktionen*, Math. Z. **56** (1952), 227–253; cf. p. 9

[5] F. Lemmermeyer, *Kreise und Quadrate modulo p*, Math. Sem. Ber. **47** (2000), 51–73; cf. p. 2

[6] F. Lemmermeyer, *Higher Descent on Pell Conics. I. From Legendre to Selmer*, preprint 2003; cf. p. 2

[7] F. Lemmermeyer, *Higher Descent on Pell Conics. II. Two Centuries of Missed Opportunities*, preprint 2003; cf. p. 2
[8] F. Lemmermeyer, *Higher Descent on Pell Conics. III. The First* 2*-Descent*, preprint 2003; cf. p. 2, 5
[9] T. Nagell, *Un théorème arithmétique sur les coniques*, Arkiv f. Mat. **2** (1952), 247–250 9
[10] P. Shastri, *Integral Points on the Unit Circle*, J. Number Theory **91** (2001), 67–70; cf. p. 4
[11] J. Silverman, J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag 1992; cf. p. 5
[12] Lin Tan, *The group of rational points on the unit circle*, Math. Mag. **69** (1996), no. 3, 163–171; cf. p. 4
[13] B. Tangedal, *A question of Stark*, Pac. J. Math. **180** (1997), 187–199; cf. p. 10
[14] D. Zagier, *The Birch-Swinnerton-Dyer conjecture from a naive point of view*, Arithmetic algebraic geometry (Texel, 1989), 377–389, Progr. Math., **89** 1991; cf. p. 2

R.C. Lyndon und P.E. Schupp, Combinatorial Group Theory, Springer 1977

the set $\phi_1, \phi_2, \phi_3, \ldots$ of all homomorphisms of G into finite groups. We can thus enumerate all images $\phi_i(w)$ where w is a word on the generators of G. If some $\phi_i(w) \neq 1$, then $w \neq 1$, and we put w on the list of words not equal to 1. Since G is residually finite, if w is any word of G not equal to the identity, there exists some ϕ_i with $\phi_i(w) \neq 1$. Thus we list all words not equal to 1 in G. This concludes the proof of the theorem. □

V. Dyson (1974) and S. Meskin (1974) have exhibited finitely generated, recursively presented, residually finite groups with unsolvable word problem.

We next turn to a theorem of Marshall Hall (1949).

Theorem 4.7. *Let G be a finitely generated group. Then the number of subgroups of G having any fixed finite index n is finite. If H is a subgroup of finite index in G, then H contains a subgroup K characteristic in G with finite index in G.*

□ Let n be a positive integer. For each subgroup H of index n, choose a complete set $c_1, \ldots, c_n$ of representatives of the right cosets of H in G with $c_1 = 1$. Now G permutes the cosets Hc_i by multiplication on the right. This induces a homomorphism ψ_H from G into the symmetric group, S_n, of permutations of $\{1, \ldots, n\}$ as follows. For $g \in G$, $\psi_H(g)$ is the permutation which sends i to j if $Hc_ig = Hc_j$. Since $Hc_1 = H$, $\psi_H(g)$ fixes the number 1 if and only if $g \in H$. If H and L are distinct subgroups of index n, there is an element g in one subgroup but not the other. Thus $\psi_H(g) \neq \psi_L(g)$ and ψ_H and ψ_L are distinct. Since G is finitely generated, there are only finitely many homomorphisms from G into S_n, and the number of subgroups of index n is thus finite.

If H is a subgroup of finite index n in G, let $H_1, \ldots, H_m$ be all the distinct subgroups of index n. Let $K = \bigcap_{i=1}^{m} H_i$. Then K is of finite index since it is the intersection of finitely many subgroups of finite index. Let α be any automorphism of G. Since the image $\alpha(H_i)$ of each H_i is again a subgroup of index n, α permutes the H_i. Thus

$$\alpha(K) = \bigcap_{i=1}^{m} \alpha(H_i) = K$$

and K is characteristic in G. □

If G is any group, $\mathrm{Aut}(G)$ will denote the group of all automorphisms of G.

The next theorem is due to G. Baumslag (1963).

Theorem 4.8. *If G is a finitely generated residually finite group, then* $\mathrm{Aut}(G)$ *is also residually finite*

□ Let $A = \mathrm{Aut}(G)$, and let $1 \neq \alpha \in A$. Then there is an element $c \in G$ such that $\alpha(c)c^{-1} = c^* \neq 1$. Since G is residually finite, there is a subgroup H of finite index in G with $c^* \notin H$. By the previous theorem, H contains a characteristic subgroup K of finite index in G. Since K is characteristic, we can define a homomorphism $\psi: A \to \mathrm{Aut}(G/K)$ by

$$\psi(\beta)[Kg] = K\beta(g).$$

Now $\mathrm{Aut}(G/K)$ is finite and $\psi(\alpha) \neq 1$. □

Since free groups are residually finite, the theorem shows that the automor-

Rankin, Modular Forms. Cambridge University Press, 1977

1: Groups of matrices and bilinear mappings

1.1. Notation. Modular functions and forms will first be defined in chapter 4. In this chapter we study the groups on which these functions are defined. We write:

$\mathbb{C}$ for the set of all (finite) complex numbers, with the usual topology.

$\mathbb{R}$ for the set of all (finite) real numbers.

$\mathbb{Q}$ for the set of all rational numbers.

$\mathbb{Z}$ for the set of all (rational) integers.

$\mathbb{Z}^+$ for the set of all positive integers.

$\mathbb{H} = \{z : z \in \mathbb{C}, \operatorname{Im} z > 0\}$, the *upper half-plane.*

$\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, *the extended complex plane.* This is the one-point compactification of $\mathbb{C}$. In the topology on $\bar{\mathbb{C}}$, a set A is open if either (i) A is an open subset of $\mathbb{C}$, or (ii) $\infty \in A$ and $\bar{\mathbb{C}} - A$ is compact in $\mathbb{C}$. With this topology, $\bar{\mathbb{C}}$ is homeomorphic to the two-sphere, i.e. to the Riemann sphere $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$.

$\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, *the one-point compactification of* $\mathbb{R}$.

$\bar{\mathbb{H}} = \mathbb{H} \cup \bar{\mathbb{R}}$.

$\mathbb{P} = \mathbb{Q} \cup \{\infty\}$.

$\mathbb{H}' = \mathbb{H} \cup \mathbb{P}$.

Other subsets of $\bar{\mathbb{C}}$ will be defined later.

We write throughout

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad S = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \tag{1.1.1}$$

for $a, b, c, d, \alpha, \beta, \gamma, \delta \in \mathbb{C}$ and put

$$|T| = \det T = ad - bc.$$

Let

$$\Theta = \{T : a, b, c, d \in \mathbb{C}, |T| = 1\}, \tag{1.1.2}$$

and

$$\Omega = \{T : a, b, c, d \in \mathbb{R}, |T| = 1\}. \tag{1.1.3}$$

Then Θ is a group under matrix multiplication with the identity element

$$I=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \tag{1.1.4}$$

and Θ contains Ω as a subgroup. In group theory the groups Θ and Ω are referred to as the *special linear* groups $\mathrm{SL}(2,\mathbb{C})$ and $\mathrm{SL}(2,\mathbb{R})$, respectively. It is easily verified that, for each $T\in\Theta$,

$$T^{-1}=\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}. \tag{1.1.5}$$

The subgroup consisting of I and

$$-I=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix} \tag{1.1.6}$$

is denoted by Λ; it is the centre of Θ and Ω.

For each $T\in\Theta$ we write

$$\operatorname{tr} T=a+d=2\cos\theta_T,$$

where θ_T is any complex number for which this equation is valid. Then it is easily shown by induction that, for any $q\in\mathbb{Z}^+$,

$$T^q=F_qT-F_{q-1}I, \tag{1.1.7}$$

where

$$F_q=\frac{\sin q\theta_T}{\sin\theta_T}. \tag{1.1.8}$$

This is a polynomial of degree $q-1$ in $\cos\theta_T$ and so is defined even when $\sin\theta_T=0$. We note also that

$$\operatorname{tr}(T^q)=2\cos q\theta_T. \tag{1.1.9}$$

Further, if $\theta_T=\pi k/q$ for some integer k, then

$$T^q=(-1)^kI. \tag{1.1.10}$$

With each $T\in\Theta$ we associate a bilinear† mapping, which we also call T, defined on $\bar{\mathbb{C}}$ by

$$w=T(z)=\frac{az+b}{cz+d}\quad(z\in\bar{\mathbb{C}}),$$

† Other terms used are *linear*, *linear fractional* and *Möbius*. No confusion should arise with the use of the word *bilinear* in multilinear algebra.

and we also, for brevity, write Tz in place of $T(z)$. For example,

$$T(-d/c)=\infty \quad\text{and}\quad T\infty=a/c.$$

These relations hold even when $c=0$, for then a and d are non-zero, so that $-d/c$ and a/c both mean ∞; this is a consequence of the rules

$$z+\infty=\infty+z=\infty,\quad \frac{z}{\infty}=0\quad(z\in\mathbb{C}),$$

$$z\infty=\infty z=\infty,\quad \frac{z}{0}=\infty\quad(z\in\bar{\mathbb{C}}-\{0\}).$$

The mapping T is a bijective mapping of $\bar{\mathbb{C}}$ onto itself, the inverse mapping being given by

$$z=T^{-1}(w)=\frac{dw-b}{-cw+a}.$$

if Γ is any subgroup of Θ, the mappings T defined by matrices $T\in\Gamma$ form a group $\hat{\Gamma}$ under composition as group operation; i.e., if $S\in\Gamma$, $T\in\Gamma$, then $(ST)(z)$ means $S\{T(z)\}$. This is easily checked. For, by the definitions of S and T,

$$ST=\begin{bmatrix}\alpha a+\beta c & \alpha b+\beta d\\ \gamma a+\delta c & \gamma b+\delta d\end{bmatrix}, \tag{1.1.11}$$

while

$$S\{T(z)\}=\frac{\alpha\dfrac{az+b}{cz+d}+\beta}{\gamma\dfrac{az+b}{cz+d}+\delta}=\frac{(\alpha a+\beta c)z+(\alpha b+\beta d)}{(\gamma a+\delta c)z+(\gamma b+\delta d)}.$$

The group $\hat{\Gamma}$ is called the *inhomogeneous group* associated with Γ, which is called a *homogeneous group*.

Let ϕ denote the mapping:

$$\phi\colon \text{matrix } T\mapsto\text{bilinear mapping } T.$$

Then the above remarks show that ϕ is a homomorphism of Θ onto $\hat{\Theta}$. Let Γ be a subgroup of Θ, so that ϕ is a homomorphism of Γ onto $\hat{\Gamma}$. The subgroups $\hat{\Gamma}$ that we consider will usually act not on the whole of $\bar{\mathbb{C}}$, but on some subset $\mathbb{D}$. We suppose that $\mathbb{D}$ is a subset of $\bar{\mathbb{C}}$ such that $\Gamma\mathbb{D}=\mathbb{D}$, i.e. $T\mathbb{D}=\mathbb{D}$ for all $T\in\hat{\Gamma}$. We suppose further

that $\mathbb{D}$ contains more than two points; usually $\mathbb{D}$ will be $\bar{\mathbb{C}}$, $\mathbb{C}$, $\mathbb{H}$ or $\mathbb{H}'$.

The identity mapping in $\hat{\Gamma}$ is $w = z$ $(z \in \mathbb{D})$, and we have

$$T(z) = \frac{az+b}{cz+d} = z \quad \text{for all } z \in \mathbb{D}$$

if and only if $az + b = cz^2 + dz$ for all $z \in \mathbb{D}$. Since $\mathbb{D}$ contains more than two points, this gives

$$b = c = a - d = 0;$$

i.e. $a = d = \pm 1$, $b = c = 0$, so that $T = \pm I$. Thus the kernel of ϕ is Λ if $-I \in \Gamma$ and is I if $-I \notin \Gamma$. Hence we have

$$\hat{\Gamma} \cong \Gamma/\Lambda \quad (\text{if } -I \in \Gamma), \quad \hat{\Gamma} \cong \Gamma \quad (\text{if } -I \notin \Gamma). \tag{1.1.12}$$

In particular, $\hat{\Theta} \cong \Theta/\Lambda$ and $\hat{\Omega} \cong \Omega/\Lambda$. The groups $\hat{\Theta}$ and $\hat{\Omega}$ are referred to in group theory as the *linear fractional* groups LF(2, $\mathbb{C}$) and LF(2, $\mathbb{R}$), respectively.

When $-I \notin \Gamma$ we can adjoin $-I$ to Γ, obtaining an overgroup $\bar{\Gamma}(= \Gamma\Lambda = \Lambda\Gamma)$ of Γ having Γ as a subgroup of index 2, and then

$$\hat{\Gamma} \cong \Gamma \cong \bar{\Gamma}/\Lambda. \tag{1.1.13}$$

By writing Γ for the homogeneous group and $\hat{\Gamma}$ for the associated inhomogeneous group we indicate that we regard the latter as being determined by the former. This point of view is especially convenient when we are concerned with algebraic properties of groups and, in particular, with multiplier systems. On the other hand, a different point of view can be taken when analytic properties are under discussion. For we shall be concerned with classes of functions f defined on $\mathbb{H}$ for which the quotient $f(Tz)/f(z)$ is the same for each member of the class, when T belongs to a certain given group of bilinear transformations. Since this quotient takes the same value for $-T$ as for T, it is often convenient to assume that $-I$ belongs to the associated matrix group. Accordingly, if we start with a group $\hat{\Gamma}$ of bilinear transformations, we may define the associated homogeneous group Γ to consist of all matrices T such that the associated bilinear transformation belongs to $\hat{\Gamma}$; it then follows that $-T \in \Gamma$ whenever $T \in \Gamma$. It is easily checked that Γ is in fact a group. (Authors who adopt this analytic point of view commonly write Γ for the inhomogeneous group and $\bar{\Gamma}$ for the homogeneous group.)

We now introduce a notation that we shall find useful when dealing with coset representatives. Let S be a set closed under an operation, which we call multiplication, and let A and B be subsets of S. Then we write, as is customary,

$$AB = \{x \in S: x = ab, a \in A, b \in B\}.$$

An element $x \in AB$ may be expressible in more than one way as a product ab for $a \in A$, $b \in B$. If, however, each $x \in AB$ is expressible in exactly one way as a product ab for $a \in A$, $b \in B$, we write

$$AB = A \cdot B.$$

It is easily verified that $A \cdot (B \cdot C) = (A \cdot B) \cdot C$, provided that one side is defined, in which case the other is also.

If A is a finite set, we write $|A|$ for the number of its elements.

Now suppose that $\Gamma_2 \subseteq \Gamma_1 \subseteq \Theta$, Γ_1 and Γ_2 being subgroups of Θ. Then the statements

$$\Gamma_1 = \Gamma_2 \cdot \mathscr{R}, \quad \Gamma_1 = \mathscr{L} \cdot \Gamma_2$$

are equivalent to the statements that $\mathscr{R}$ is a set of right coset representatives of Γ_1 modulo Γ_2, and that $\mathscr{L}$ is a set of left coset representatives, respectively. We call $\mathscr{R}$ a *right transversal*, and $\mathscr{L}$ a *left transversal*, of Γ_2 in Γ_1. If $|\mathscr{R}|$ is finite, so is $|\mathscr{L}|$ and

$$|\mathscr{R}| = |\mathscr{L}| = [\Gamma_1 : \Gamma_2],$$

the index of Γ_2 in Γ_1. Note that, if $T \in \mathscr{R}$ and $-I \in \Gamma_2$, then $-T \notin \mathscr{R}$.

Similar notations can be used with inhomogeneous groups. We note also that, if $-I \in \Gamma_2 \subseteq \Gamma_1 \subseteq \Theta$ and $\Gamma_1 = \Gamma_2 \cdot \mathscr{R}$, then $\hat{\Gamma}_1 = \hat{\Gamma}_2 \cdot \hat{\mathscr{R}}$, where there is a one-to-one correspondence between matrices in $\mathscr{R}$ and transformations in $\hat{\mathscr{R}}$; for this reason we shall usually write not only $\Gamma_1 = \Gamma_2 \cdot \mathscr{R}$ but also $\hat{\Gamma}_1 = \hat{\Gamma}_2 \cdot \mathscr{R}$.

Theorem 1.1.1. *Let Γ_2 and Γ_1 be subgroups of Θ with $\Gamma_2 \subseteq \Gamma_1$. Then, if $\Gamma_1 = \Gamma_2 \cdot \mathscr{R}$ and S is any member of Γ_1, $\Gamma_1 = \Gamma_2 \cdot (\mathscr{R}S)$. A similar result holds in the inhomogeneous case.*

Proof. For $\Gamma_1 = \Gamma_1 S = (\Gamma_2 \cdot \mathscr{R})S = \Gamma_2 \cdot (\mathscr{R}S)$.

Theorem 1.1.2. *Let Γ_2 be a subgroup of finite index μ in a group Γ_1, and let S be a fixed member of Γ_1. Then there exist a finite number of*

elements $L_1, L_2, \ldots, L_m$, say, in Γ_1 and m disjoint sets

$$\mathscr{S}_i = \bigcup\{L_i S^k : 0 \le k < \sigma_i\} \quad (1 \le i \le m),$$

where

$$\sigma_i = \min\{k : S^k \in L_i^{-1}\Gamma_2 L_i, k \in \mathbb{Z}^+\}, \tag{1.1.14}$$

such that

$$\mu = \sigma_1 + \sigma_2 + \cdots + \sigma_m \tag{1.1.15}$$

and

$$\Gamma_1 = \Gamma_2 \cdot \bigcup_{i=1}^{m} \mathscr{S}_i.$$

Moreover, if S has finite order σ, then σ_i divides σ for $1 \le i \le m$. Also, if Γ_2 is normal in Γ_1, then $\sigma_i = \sigma_0$, say, for $1 \le i \le m$, and so

$$\mu = m\sigma_0. \tag{1.1.16}$$

Proof. Take any $L_1 \in \Gamma_1$ and define σ_1 by (1.1.14); since $L_1^{-1}\Gamma_2 L_1$ has finite index μ in Γ_1, σ_1 is a finite positive number and the members of $\mathscr{S}_1$ belong to σ_1 different right cosets of Γ_2 in Γ_1. If $\mu = \sigma_1$, this completes the proof and $m = 1$ in this case. If $\mu > \sigma_1$ we take any L_2 not belonging to $\Gamma_2\mathscr{S}_1$ and define σ_2 by (1.1.14). As before, the σ_2 elements $L_2 S^k$ $(0 \le k < \sigma_2)$ belong to different right cosets of Γ_2. Moreover $L_2 S^k \notin \Gamma_2\mathscr{S}_1$; for if $L_2 S^k \in \Gamma_2\mathscr{S}_1$ then $L_2 \in \Gamma_2\mathscr{S}_1 S^{-k} = \Gamma_2\mathscr{S}_1$, which is false. If $\mu = \sigma_1 + \sigma_2$ the theorem follows; if $\mu > \sigma_1 + \sigma_2$, we take an $L_3 \notin \Gamma_2(\mathscr{S}_1 \cup \mathscr{S}_2)$ and proceed similarly. Since μ is finite and $\sigma_i > 0$ for each i, there exists a positive integer m such that (1.1.15) holds and the process then terminates, giving the required result. The two final sentences in the enunciation are immediate consequences.

Note that, when $L_1, L_2, \ldots, L_r$ $(r < m)$ have been chosen, L_{r+1} may be any member of Γ_1 not in $\Gamma_2 \bigcup_{i=1}^{r} \mathscr{S}_i$. Also, although the elements $L_1, \ldots, L_m$ are not uniquely determined, the integer m is the same and so are the numbers $\sigma_1, \sigma_2, \ldots, \sigma_m$ (in some order) for all choices of $L_1, \ldots, L_m$.

Theorem 1.1.3. *Let Γ_2 be a normal subgroup of finite index μ in a group Γ_1, and let S be a fixed member of Γ_1. Let σ be the least positive integer such that $S^\sigma \in \Gamma_2$ and write*

$$\mathscr{S} = \bigcup\{S^k : 0 \le k < \sigma\}.$$

Then there exist $m = \mu/\sigma$ distinct elements $L_1, L_2, \ldots, L_m$ of Γ_1 such that

$$\Gamma_1 = \mathscr{S} \cdot \Gamma_2 \cdot \mathscr{L}, \tag{1.1.17}$$

where $\mathscr{L} = \bigcup\{L_i : 1 \le i \le m\}$. Also, if Γ_1^* and Γ_2^* are the subgroups of Γ_1 generated by S and S^σ, respectively, and $\Gamma_2 = \Gamma_2^* \cdot \mathscr{R}$, then

$$\Gamma_1 = \Gamma_1^* \cdot \mathscr{R} \cdot \mathscr{L}. \tag{1.1.18}$$

Proof. The proof of (1.1.17) is similar to that of theorem 1.1.2. The normality of Γ_2 in Γ_1 comes in when we infer from

$$S^k T_2 L_i = S^l T_2' L_j \quad (0 \le l \le k < \sigma;\ T_2, T_2' \in \Gamma_2),$$

that $L_j \in \mathscr{S}\Gamma_2 L_i$, and so $L_j = L_i$, $l = k$ and $T_2' = T_2$. We then have, since $\Gamma_1^* = \mathscr{S} \cdot \Gamma_2^*$,

$$\Gamma_1 = \mathscr{S} \cdot \Gamma_2 \cdot \mathscr{L} = \mathscr{S} \cdot \Gamma_2^* \cdot \mathscr{R} \cdot \mathscr{L} = \Gamma_1^* \cdot \mathscr{R} \cdot \mathscr{L},$$

which is (1.1.18).

We note that, by (1.1.11),

$$\operatorname{tr} ST = \operatorname{tr} TS \tag{1.1.19}$$

whenever S and T belong to Θ. In particular, we deduce that

$$\operatorname{tr} L^{-1}TL = \operatorname{tr} T \tag{1.1.20}$$

whenever L and T belong to Θ; i.e. conjugate elements have the same trace.

If A and B are elements of a group, we write $\langle A, B\rangle$ for the subgroup generated by them, and use similar notations for any number of generators.

We conclude this subsection by introducing another convenient notation. We shall write

$$x :=^{!} y \quad \text{or} \quad y =: x$$

to denote that x is a new symbol, which is defined to be equal to y; y will often be a rather complicated expression.

1.2. The modular group. We write

$$\Gamma(1) := \{T \in \Omega : a, b, c, d \in \mathbb{Z}\}. \tag{1.2.1}$$

It follows from (1.1.5, 11) that this set of matrices is a group, and it is known as the *(homogeneous) modular group.*† The corresponding group of mappings is denoted by $\hat{\Gamma}(1)$ and is called the *(inhomogeneous) modular group.* By (1.1.12), $\hat{\Gamma}(1) \cong \Gamma(1)/\Lambda$. Alternative notations are

$$\Gamma(1) = \mathrm{SL}(2, \mathbb{Z}), \quad \hat{\Gamma}(1) = \mathrm{LF}(2, \mathbb{Z}).$$

The following special matrices belonging to $\Gamma(1)$ occur with great frequency:

$$U := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad V := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad W := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{1.2.2}$$

The corresponding mappings are given by

$$Uz = z + 1, \quad Vz = -1/z, \quad Wz = \frac{z}{z+1}.$$

We note that, for any $k \in \mathbb{Z}$,

$$U^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}. \tag{1.2.3}$$

Further,

$$V^2 = -I, \quad P^3 = -I, \tag{1.2.4}$$

where

$$P := VU = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}. \tag{1.2.5}$$

Also

$$P^2 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \quad W = UVU. \tag{1.2.6}$$

The mappings V and P therefore have periods 2 and 3 respectively.

We denote by Γ_U the subgroup of $\Gamma(1)$ generated by $\pm U$; it consists of all matrices $\pm U^n$ $(n \in \mathbb{Z})$. The corresponding mappings are the translations

$$w = z + n \quad (n \in \mathbb{Z}).$$

† German: *Modulgruppe.*

We write similarly, Γ_{U^k} $(k \in \mathbb{Z}^+)$ for the group generated by $\pm U^k$; the corresponding mappings are the translations

$$w = x + k \quad (n \in \mathbb{Z}).$$

We now consider the group $\Gamma(1)$ more closely. Let c, d be any two coprime integers. Then we can find integers a, b such that $ad - bc = 1$; i.e. we can find a $T \in \Gamma(1)$ with second row $[c, d]$. Further if a', b' is any other pair of integers with $a'd - b'c = 1$, then, by elementary number theory,

$$a' = a + nc, \quad b' = b + nd$$

for some $n \in \mathbb{Z}$, and conversely. Thus the only matrices in $\Gamma(1)$ with second row $[c, d]$ are the matrices

$$T' = U^n T \quad (n \in \mathbb{Z}).$$

Hence, if S is any member of the right coset $\Gamma_U T$ of Γ_U in $\Gamma(1)$, then $[\gamma, \delta] = \pm[c, d]$, and conversely. We deduce

Theorem 1.2.1. *Let $\mathcal{R}$ be a set of matrices $T \in \Gamma(1)$ with the property that, for each $S \in \Gamma(1)$ there is exactly one $T \in \mathcal{R}$ such that $[c, d] = \pm[\gamma, \delta]$. Then*

$$\Gamma(1) = \Gamma_U \cdot \mathcal{R}. \tag{1.2.7}$$

Conversely, if (1.2.7) holds, then $\mathcal{R}$ has the property stated. A similar result holds for the inhomogeneous groups $\hat{\Gamma}(1)$, $\hat{\Gamma}_U$ and a set of mappings $\hat{\mathcal{R}}$ with corresponding properties.

The next theorem shows that if T is any element of $\Gamma(1)$ we can find a conjugate element $L^{-1}TL$ of a certain simple form.

Theorem 1.2.2. *If $T \in \Gamma(1)$ and $\mathrm{tr}\ T = t$, then there exists an $L \in \Gamma(1)$ such that, if $S = L^{-1}TL$, then*

$$|\alpha - \tfrac{1}{2}t| \leq \tfrac{1}{2}|\gamma|, \quad |\delta - \tfrac{1}{2}t| \leq \tfrac{1}{2}|\gamma|, \quad |\gamma| \leq |\beta|, \quad 3\gamma^2 \leq |t^2 - 4|,$$

where S is given by (1.1.1). Further, L belongs to the subgroup generated by U and V.

Proof. For any $n \in \mathbb{Z}$

$$T_1 := U^{-n}TU^n = \begin{bmatrix} 1 & -n \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a-nc & b+n(a-d)-n^2c \\ c & d+nc \end{bmatrix} =: \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}.$$

Note that $\operatorname{tr} T_1 = t$. We choose n to make

$$|a_1 - \tfrac{1}{2}t| = |\tfrac{1}{2}(a-d) - nc| = |d_1 - \tfrac{1}{2}t| \leq \tfrac{1}{2}|c| = \tfrac{1}{2}|c_1|;$$

this is possible even when $c = 0$, since then $a = d$. If $|c_1| \leq |b_1|$ we stop the process here; if $|c_1| > |b_1|$ we form

$$T_2 := V^{-1}T_1V = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} d_1 & -c_1 \\ -b_1 & a_1 \end{bmatrix}$$

$$=: \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$$

and note that we then have $|c_2| < |c_1|$ and $\operatorname{tr} T_2 = t$. We now form $T_3 := U^{-m}T_2U^m$, choosing m as above to make

$$|a_3 - \tfrac{1}{2}t| = |d_3 - \tfrac{1}{2}t| \leq \tfrac{1}{2}|c_3| = \tfrac{1}{2}|c_2| = \tfrac{1}{2}|b_1|,$$

and stop the process here if $|c_3| \leq |b_3|$; if not, then $|c_3| > |b_3|$ and we proceed as above, obtaining a matrix T_4 with $|c_4| < |c_3| < |c_1|$. The process must stop after a finite number k of steps when we reach a matrix $S = T_{2k+1}$ with

$$|\alpha - \tfrac{1}{2}t| = |\delta - \tfrac{1}{2}t| \leq \tfrac{1}{2}|\gamma| \quad \text{and} \quad |\gamma| \leq |\beta|.$$

Hence

$$|t^2 - 4| = |(\alpha+\delta)^2 - 4 = |4\beta\gamma + (\alpha-\delta)^2|$$

$$\geq 4|\beta||\gamma| - |\alpha-\delta|^2 \geq 4\gamma^2 - \gamma^2 = 3\gamma^2.$$

Theorem 1.2.3. *Let $T \in \Gamma(1)$. If T is of finite order then T is conjugate to one of the matrices*

$$\pm I, \pm V, \pm P, \pm P^2$$

and $|\operatorname{tr} T| \leq 2$. Conversely, if $|\operatorname{tr} T| \leq 2$, then T is conjugate to one of these matrices or else T is a conjugate of $\pm U^k$ for some $k \in \mathbb{Z}$.

Proof. From the inequalities in theorem 1.2.2 we get the possibilities shown in table 1 for S when $|t| \leq 2$. If $t > 2$, we can put

Table 1

$t = \alpha + \delta$	α	β	γ	δ	S
0	0	∓ 1	± 1	0	$\pm V$
1	0	∓ 1	± 1	1	$P, -V^{-1}P^2V$
	1			0	$-P^2, V^{-1}PV$
−1	0	∓ 1	± 1	−1	$-P, V^{-1}P^2V$
	−1			0	$P^2, -V^{-1}PV$
2	1	k	0	1	U^k
−2	−1	$-k$	0	−1	$-U^k$

$t = 2\cosh\theta$, for some $\theta > 0$, so that, by (1.1.9),

$$\operatorname{tr}(S^q) = 2\cosh q\theta > 2$$

for all $q \in \mathbb{Z}^+$. It follows that, if $|t| > 2$, then $S^q \neq \pm I$ for all $q \in \mathbb{Z}^+$.

Theorem 1.2.4. *$\Gamma(1)$ is generated by U and V; every element $T \in \Gamma(1)$ can be written in the form*

$$T = U^{q_0}VU^{q_1}V \cdots VU^{q_n} \tag{1.2.8}$$

where $q_i \in \mathbb{Z}$ $(0 \leq i \leq n)$; this representation is not unique.

Proof. Let $T \in \Gamma(1)$ and take $S = L^{-1}TL$ as in theorem 1.2.2. By theorem 1.2.3, we may suppose that $|t| > 2$, where $t = \operatorname{tr} T$. Then $S \notin \Gamma_U$, and so we can choose $q \in \mathbb{Z}$ so that, if $t_1 = \operatorname{tr} U^qS$,

$$|t_1| = |t + q\gamma| \leq \tfrac{1}{2}|\gamma|;$$

this is possible since $\gamma \neq 0$. Then

$$|t_1| \leq \tfrac{1}{2}|\gamma| \leq \tfrac{1}{2}\{\tfrac{1}{3}|t^2-4|\}^{\frac{1}{2}} < |t|,$$

and some transform of U^qS will satisfy the inequalities in theorem 1.2.2. In this way we can derive an element S' of trace t', where $|t'| \leq 2$, and where S' is derived from T by multiplication on left and right by powers of U and V. Since, by table 1, S' is also a product of matrices U and V, the required result follows. The representation is not unique, since

$$P^6 = (VU)^6 = I.$$

Theorem 1.2.5. *$\Gamma(1)$ is generated by V and P^2; every element $T \in \Gamma(1)$ can be written uniquely in the form*

$$T = (-1)^r P^{2p_0} V P^{2p_1} V \cdots V P^{2p_n} \tag{1.2.9}$$

where

$$0 \le r \le 1, \quad 0 \le p_i \le 2 \quad (0 \le i \le n), \quad p_i > 0 \quad (0 < i < n). \tag{1.2.10}$$

(*When $n = 0$, the expression on the right-hand side of (1.2.9) is* $(-1)^r P^{2p_0}$.)

Proof. It is clear from (1.2.4, 5, 8) that T can be written in the form (1.2.9) subject to the conditions (1.2.10). The representation is unique if we can have

$$I = (-1)^r P^{2p_0} V P^{2p_1} \cdots V P^{2p_n} \tag{1.2.11}$$

only when $n = r = p_0 = 0$. We call the right-hand side of (1.2.11) a *word of length n* and suppose that (1.2.11) holds for some $n > 0$ and that n is the least positive integer for which this holds. Then

$$I = (-1)^r V P^{2p_1} V P^{2p_2} \cdots V P^{2p_n + 2p_0} \tag{1.2.12}$$

and $p_n + p_0 \not\equiv 0 \pmod 3$. For if $p_n + p_0 \equiv 0 \pmod 3$, then $n \ge 2$ and

$$I = (-1)^{r-1} P^{2p_1} V \cdots V P^{2p_{n-1}},$$

which is a word of shorter length $n-2$; hence $n = 2$ and so $p_1 = 0$, which is false. Since $P^6 = I$, we may assume that $p_n + p_0 = 1$ or 2, so that we deduce from (1.2.12) that $\pm I$ can be represented as a product of n factors each of which is either

$$VP^4 = U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \text{or} \quad -VP^2 = W = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{1.2.13}$$

Since the entries in U and W are all non-negative the same is true of the product, which must be $+I$. But, if either

$$I = US \quad \text{or} \quad I = WS,$$

then S has one negative entry, and this gives the desired contradiction.

It follows from theorem 1.2.5 that

$$\Gamma(1) = \langle P^2, V \rangle * \langle P, V \rangle \tag{1.2.14}$$

with the notation introduced at the end of §1.1. Further, since $P^4 = -P$, it is clear that theorem 1.2.5 remains true, possibly with a different value of r, if each exponent $2p_i$ $(0 \le i \le n)$ in (1.2.9) is replaced by p_i $(0 \le i \le n)$.

Since the matrices T and $-T$ give rise to the same bilinear mapping, it is clear that we have

Theorem 1.2.6. *$\hat{\Gamma}(1)$ is generated by the mappings V and P, which have orders 2 and 3, respectively; i.e.*

$$\hat{\Gamma}(1) = \langle P, V \rangle. \tag{1.2.15}$$

Further, every mapping T of $\hat{\Gamma}(1)$ can be written uniquely in the form

$$T = P^{p_0} V P^{p_1} \cdots V P^{p_n} \tag{1.2.16}$$

as a composition of mappings, where

$$0 \le p_i \le 2 \quad (0 \le i \le n), \quad p_i > 0 \quad (0 < i < n). \tag{1.2.17}$$

Theorem 1.2.6 can be expressed by stating that $\hat{\Gamma}(1)$ is the free product of the cyclic groups $\langle V \rangle$ and $\langle P \rangle$. We use an asterisk (*) to denote a free product, so that we have

$$\hat{\Gamma}(1) = \langle V \rangle * \langle P \rangle. \tag{1.2.18}$$

The integer n occurring in (1.2.9) is called the *length* of the group element or *word* T and we write $l(T) := n$. Thus $l(T) = 0$ if and only if $T = P^q$ for some integer q.

We conclude this section by investigating the automorphism groups of $\Gamma(1)$ and $\hat{\Gamma}(1)$; see Hua and Reiner (1951). We write Aut G and Inn G for the groups of all automorphisms and all inner automorphisms, respectively, of a group G. We also recall that Klein's four-group is the direct product of two cyclic groups of order 2.

Theorem 1.2.7. *Let ψ be an automorphism of $\Gamma(1)$. Then ψ is determined uniquely by its action on the generators V and P, and we must have, for some $L \in \Gamma(1)$;*

$$\psi(V) = L^{-1} V^{\mu} L, \quad \psi(P) = L^{-1} P^{\nu} L, \tag{1.2.19}$$

where $\mu = \pm 1$, $\nu = \pm 1$. Accordingly, Aut $\Gamma(1)$/Inn $\Gamma(1)$ *is isomorphic to the four-group.*

Proof. Clearly, if $\psi \in$ Aut $\Gamma(1)$, then, for some $L_1, L_2 \in \Gamma(1)$,

$$\psi(V) = L_1^{-1} V^{\mu} L_1, \quad \psi(P) = L_2^{-1} P^{\nu} L_2.$$

by theorem 1.2.3, where $\mu = \pm 1$, $\nu = \pm 1$. By applying an inner automorphism to ψ we obtain an automorphism $\varphi \in \mathrm{Aut}\, \Gamma(1)$ such that

$$\varphi(V) = L^{-1} V^{\mu} L, \quad \varphi(P) = P^{\nu} \tag{1.2.20}$$

for some $L \in \Gamma(1)$. Further, we may assume either that $L = I$, or else that

$$m := l(L) \geq 1,$$

in which case L has the canonical form

$$L = P^{2q_0} V P^{2q_1} V \cdots V P^{2q_m}, \tag{1.2.21}$$

where $0 < q_0 \leq 2$ and $q_m = 0$; for V^{μ} and P^{ν} are unaltered when conjugated by V and P, respectively. Note that $l(L^{-1}VL) = 2m+1$.

Now take any $T \in \Gamma(1)$ as in (1.2.9), so that $n = l(T)$. From (1.2.9, 20, 21) we can work out $\varphi(T)$ and express it in canonical form. We find that

$$l(\varphi(T)) = n(2m+1).$$

From this it is clear that φ is a surjection if and only if $m = 0$, i.e. $L = I$. The theorem follows.

Note that, when $\mu = -1$, $\nu = 1$, we deduce from (1.2.19) that, for all $T \in \Gamma(1)$,

$$\psi(T) = (-1)^{l(T)} L^{-1} T L \quad \text{for some } L \in \Gamma(1). \tag{1.2.22}$$

Write

$$J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{1.2.23}$$

so that J is its own inverse and does not belong to Θ. Note that

$$J^{-1} T J = \begin{bmatrix} a & -b \\ -c & d \end{bmatrix}. \tag{1.2.24}$$

and that

$$J^{-1} V J = -V, \quad J^{-1} P J = V^{-1} P^{-1} V. \tag{1.2.25}$$

We now deduce immediately the following theorem from theorem 1.2.7.

Theorem 1.2.8. $\mathrm{Aut}\, \hat{\Gamma}(1) = \langle J \rangle \mathrm{Inn}\, \hat{\Gamma}(1)$. *Thus* $\mathrm{Aut}\, \hat{\Gamma}(1)/\mathrm{Inn}\, \hat{\Gamma}(1)$ *is a cyclic group of order* 2.

In fact, if $\psi \in \mathrm{Aut}\, \hat{\Gamma}(1)$, then, for each $T \in \hat{\Gamma}(1)$,

$$\psi(T) = L^{-1} T L \quad \text{or} \quad \psi(T) = J^{-1} L^{-1} T L J$$

for some fixed $L \in \hat{\Gamma}(1)$.

In conclusion, we note that the automorphism

$$T \mapsto T' = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$$

is an outer automorphism, since

$$T' = (VJ)^{-1} T (VJ).$$

Further, the outer automorphism $T \mapsto J^{-1} T J$ is closely associated with the non-analytic map

$$z \mapsto J^*(z) := -\bar{z}, \tag{1.2.26}$$

where the bar denotes the complex conjugate. For

$$(J^*)^{-1} T J^*(z) = J^{-1} T J(z) = \frac{az - b}{-cz + d}.$$

1.3. The subgroups Γ^2, Γ^3, Γ^4 and $\Gamma''(1)$. Let $T \in \Gamma(1)$, and suppose that T is expressed as in theorem 1.2.5. We define

$$h(T) = n + 2r, \quad p(T) = p_0 + p_1 + \cdots + p_n. \tag{1.3.1}$$

It then follows from (1.2.4) that, for any T_1, $T_2 \in \Gamma(1)$,

$$h(T_1 T_2) \equiv h(T_1) + h(T_2) \pmod 4 \tag{1.3.2}$$

and

$$p(T_1 T_2) \equiv p(T_1) + p(T_2) \pmod 3. \tag{1.3.3}$$

Thus h and p are homomorphisms of $\Gamma(1)$ onto the additive groups of residue classes modulis 4 and 3, respectively. The kernels of these homomorphisms, namely

$$\Gamma^4 := \{T: T \in \Gamma(1), h(T) \equiv 0 \pmod 4\} \tag{1.3.4}$$

and

$$\Gamma^3 := \{T: T \in \Gamma(1), p(T) \equiv 0 \pmod 3\}, \tag{1.3.5}$$

are therefore normal subgroups of $\Gamma(1)$ of indices 4 and 3, respectively. We also write

$$\Gamma^2 := \{T: T \in \Gamma(1), h(T) \equiv 0 \pmod 2\} \tag{1.3.6}$$

so that $\Gamma^4 \subseteq \Gamma^2 \subseteq \Gamma(1)$, and Γ^2 is a normal subgroup of $\Gamma(1)$ of index 2.

From (1.2.9) and the definitions of the three subgroups it is easily seen that Γ^2 is generated by P and

$$P_1 := V^{-1}PV = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}; \tag{1.3.7}$$

thus

$$\Gamma^2 = \langle P, P_1 \rangle.$$

Similarly,

$$\Gamma^4 = \langle P^2, P_1^2 \rangle, \tag{1.3.8}$$

and

$$\Gamma^3 = \langle V, V_1, V_2 \rangle, \tag{1.3.9}$$

where

$$\begin{aligned} V_1 &:= P^{-1}VP = P^2VP^4 = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}, \\ V_2 &:= P^{-2}VP^2 = P^4VP^2 = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}. \end{aligned} \tag{1.3.10}$$

Note that $-I$ belongs to Γ^2 and Γ^3 but not to Γ^4. Thus

$$\Gamma^4 \cong \Gamma^2/\Lambda.$$

It is clear from theorem 1.2.5 that

$$\Gamma^4 = \langle P^2 \rangle * \langle P_1^2 \rangle. \tag{1.3.11}$$

We now consider the commutator group $\Gamma'(1)$ of $\Gamma(1)$. This group is generated by the commutators

$$[S, T] := STS^{-1}T^{-1}, \tag{1.3.12}$$

for $S, T \in \Gamma(1)$. We prove

Theorem 1.3.1. *The commutator group $\Gamma'(1)$ is a free group of index 12 in $\Gamma(1)$. $\Gamma'(1)$ has rank 2 and is generated by UW and WU.*

Further $\Gamma'(1) = \Gamma^3 \cap \Gamma^4$, so that $T \in \Gamma'(1)$ if and only if

$$h(T) \equiv 0 \pmod 4 \quad \textit{and} \quad p(T) \equiv 0 \pmod 3. \tag{1.3.13}$$

The factor group $\Gamma(1)/\Gamma'(1)$ is a cyclic group of order 12 and is generated by the coset $\Gamma'(1)U$. Also, writing Γ' for $\Gamma'(1)$, we have

$$-I \in \Gamma' U^6, \quad V \in \Gamma' U^9, \quad P \in \Gamma' U^{10}, \quad W \in \Gamma' U^{-1}. \tag{1.3.14}$$

In fact,

$$U^6 = -[UW, (WU)^{-1}].$$

Finally,

$$\Gamma(1) = \Gamma^3\Gamma^4, \quad \textit{and} \quad \bar{\Gamma}'(1) = \Gamma^2 \cap \Gamma^3.$$

Proof. For any $S, T \in \Gamma(1)$ it follows from (1.3.2, 3, 12) that

$$h([S, T]) \equiv 0 \pmod 4, \quad p([S, T]) \equiv 0 \pmod 3.$$

It follows from this that every $T \in \Gamma'(1)$ satisfies (1.3.13).

Conversely, if $T \in \Gamma(1)$ and T satisfies (1.3.13), we can express T in the form (1.2.9), where (1.2.10) holds. Write

$$q_m = p_0 + p_1 + \cdots + p_m \quad (0 \leq m \leq n).$$

Then n is even and it is easily verified that

$$T = [P^{2q_0}, V][V, P^{2q_1}][P^{2q_2}, V][V, P^{2q_3}] \ldots [P^{2q_{n-2}}, V][V, P^{2q_{n-1}}],$$

so that $T \in \Gamma'(1)$. Now

$$[V, P] = [V, P^4] = [P^4, V]^{-1} = UW$$

and

$$[V, P^2] = [P^2, V]^{-1} = WU,$$

so that $\Gamma'(1)$ is generated by UW and WU.

From these facts it is clear that $\Gamma'(1)$ is a free group of rank 2 having UW and WU as generators; for otherwise we could express I as a product of the form (1.2.9) with $n > 0$, and this is impossible. The index of $\Gamma'(1)$ in $\Gamma(1)$ is 12 since each coset corresponds to a different pair (h, p) of residue classes and consists of those $T \in \Gamma(1)$ for which

$$h(T) \equiv h \pmod 4 \quad \text{and} \quad p(T) \equiv p \pmod 3.$$

In fact

$$\Gamma(1) = \Gamma''(1) \cdot \mathcal{R},$$

where $\mathcal{R}$ consists of the 12 matrices $P^{2p}V^h$ $(0 \le p < 3, 0 \le h < 4)$.

If we use $\sim$ to denote the equivalence relation of belonging to the same coset of $\Gamma''(1)$ in $\Gamma(1)$, we deduce from

$$V = -U^3[U^{-2}, V][V, U^{-1}]$$

that $-I \sim V^2 \sim U^6$, $V \sim U^9$, $P \sim VU \sim U^{10}$ and $W \sim WUU^{-1} \sim U^{-1}$. In fact we have

$$P^{2p}V^h \sim U^{8p+9h} \quad (0 \le p < 3, 0 \le h < 4).$$

Finally, it is obvious from the earlier part of the proof that $\Gamma''(1) = \Gamma^3 \cap \Gamma^4$, and so $\bar{\Gamma}''(1) = \Gamma^2 \cap \Gamma^3$. Also $\Gamma(1) = \Gamma^3\Gamma^4$ since $\Gamma^3\Gamma^4$ contains the generators V and P^2.

We now state the corresponding results for the associated inhomogeneous groups. For this purpose we use the results already obtained, together with (1.1.12, 13).

Theorems 1.3.2. *The associated normal inhomogeneous groups have the following properties*:

$$\hat{\Gamma}^2 = \hat{\Gamma}^4 = \langle P^2 \rangle * \langle P_1^2 \rangle, \quad \hat{\Gamma}^4 \cong \Gamma^4, \quad [\hat{\Gamma}(1) : \hat{\Gamma}^2] = 2, \tag{1.3.15}$$

$$\hat{\Gamma}^3 = \langle V \rangle * \langle V_1 \rangle * \langle V_2 \rangle, \quad \hat{\Gamma}^3 \cong \Gamma^3/\Lambda, \quad [\hat{\Gamma}(1) : \hat{\Gamma}^3] = 3, \tag{1.3.16}$$

$$\hat{\Gamma}''(1) = \hat{\Gamma}^2 \cap \hat{\Gamma}^3 \cong \Gamma''(1), \quad [\hat{\Gamma}(1) : \hat{\Gamma}''(1)] = 6, \tag{1.3.17}$$

$$\hat{\Gamma}(1) = \hat{\Gamma}^2\hat{\Gamma}^3. \tag{1.3.18}$$

The factor group $\hat{\Gamma}(1)/\hat{\Gamma}''(1)$ is a cyclic group of order 6 generated by the coset containing U.

We conclude by giving alternative definitions of Γ^2, Γ^3 and $\bar{\Gamma}''(1)$. For this purpose we define

$$Q := [W, U] = VU^3 = \begin{bmatrix} 0 & -1 \\ 1 & 3 \end{bmatrix}. \tag{1.3.19}$$

Theorem 1.3.3. *For $\nu = 2$, 3 and 6 define*

$$G^\nu := \{T \in \Gamma(1) : Q^{-1}TQ \equiv \pm T \pmod{\nu}\}. \tag{1.3.20}$$

Then $G^2 = \Gamma^2$, $G^3 = \Gamma^3$ and $G^6 = \bar{\Gamma}''(1)$.

Proof. G^2 and G^3 are certainly subgroups of $\Gamma(1)$ and are proper subgroups since they do not contain U. Further, G^2 contains the generators P and P_1 of Γ^2, while G contains the generators V, V_1 and V_2 of Γ^3. We deduce that $G^2 = \Gamma^2$ and $G^3 = \Gamma^3$. Accordingly

$$G^6 = G^2 \cap G^3 = \Gamma^2 \cap \Gamma^3 = \bar{\Gamma}''(1).$$

1.4. The level of a subgroup; congruence subgroups. Let Γ be any subgroup of $\Gamma(1)$, not necessarily of finite index. For each $L \in \Gamma(1)$, define n_L to be the least positive integer such that

$$U^{n_L} \in L\Gamma L^{-1}. \tag{1.4.1}$$

When Γ has infinite index in $\Gamma(1)$, no such finite positive n_L need exist and we put $n_L = \infty$ in this case; on the other hand, when Γ has finite index in $\Gamma(1)$, n_L always exists and is finite. Consider the set $\{n_L : L \in \Gamma(1)\}$. If the numbers in this set have a finite least common multiple n, we call n the *level*† of Γ and write

$$\operatorname{lev} \Gamma = n. \tag{1.4.2}$$

In all other cases we put $\operatorname{lev} \Gamma = \infty$.

In particular, $\operatorname{lev} \Gamma$ is always finite when Γ has finite index in $\Gamma(1)$, since then Γ has only a finite number of conjugate subgroups in $\Gamma(1)$. When (1.4.2) holds for finite n we have

$$U^n \in L\Gamma L^{-1} \quad \text{for all } L \in \Gamma(1). \tag{1.4.3}$$

We now write $\Delta(n)$ for the normal closure of the cyclic group $\langle U^n \rangle$; i.e.

$$\Delta(n) = \langle L^{-1}U^nL : L \in \Gamma(1) \rangle \tag{1.4.4}$$

and is the smallest normal subgroup containing $\langle U^n \rangle$. Clearly n is the smallest positive integer such that $\Delta(n) \subseteq \Gamma$.

We make exactly similar definitions for inhomogeneous groups. Thus, if $\hat{\Gamma}$ is any subgroup of $\hat{\Gamma}(1)$, $\operatorname{lev} \hat{\Gamma}$ is defined and the group $\hat{\Delta}(n)$ is the normal closure of the group $\langle U^n \rangle$ of mappings. By way of example we note that

$$\operatorname{lev} \Gamma(1) = \operatorname{lev} \hat{\Gamma}(1) = 1, \quad \operatorname{lev} \Delta(n) = \operatorname{lev} \hat{\Delta}(n) = n,$$

and $\operatorname{lev} \langle U^n \rangle = \infty$ for both homogeneous and inhomogeneous groups $\langle U^n \rangle$. Also, by theorems 1.3.1, 2,

$$\operatorname{lev} \Gamma''(1) = 12, \quad \operatorname{lev} \hat{\Gamma}''(1) = 6.$$

† German: *Stufe*.

An important class of subgroups of the modular group consists of what are called congruence subgroups. If S and $T \in \Gamma(1)$, we write

$$S \equiv T \pmod{n},$$

where $n \in \mathbb{Z}^+$, if and only if

$$\alpha \equiv a, \quad \beta \equiv b, \quad \gamma \equiv c \quad \text{and} \quad \delta \equiv d \pmod{n}.$$

It follows at once from (1.1.5) and (1.1.11) that if

$$S_1 \equiv S_2 \pmod{n} \quad \text{and} \quad T_1 \equiv T_2 \pmod{n},$$

then

$$T_1^{-1} \equiv T_2^{-1} \pmod{n} \quad \text{and} \quad S_1 T_1 \equiv S_2 T_2 \pmod{n}. \tag{1.4.5}$$

We write

$$\Gamma(n) := \{S \in \Gamma(1) : S \equiv I \pmod{n}\}; \tag{1.4.6}$$

this agrees with the previous definition when $n = 1$. Then $\Gamma(n)$ is a group; for if $S \equiv T \equiv I \pmod{n}$, then

$$ST^{-1} \equiv II^{-1} \equiv I \pmod{n}.$$

We also write

$$\bar{\Gamma}(n) := \Lambda\Gamma(n) = \{S \in \Gamma(1) : S \equiv \pm I \pmod{n}\}, \tag{1.4.7}$$

and can prove similarly that this is a group. The two homogeneous groups $\Gamma(n)$ and $\bar{\Gamma}(n)$ give rise to the same inhomogeneous group, which we denote by $\hat{\Gamma}(n)$. Each of the groups $\bar{\Gamma}(n)$ and $\hat{\Gamma}(n)$ is called a *principal congruence group*,† and the same title is sometimes conferred on $\Gamma(n)$.

Since $-I \in \Gamma(n)$ if and only if $n = 1$ or 2, we have

$$\hat{\Gamma}(n) \cong \Gamma(n)/\Lambda \cong \bar{\Gamma}(n)/\Lambda \quad (n = 1, 2), \tag{1.4.8}$$

$$\hat{\Gamma}(n) \cong \Gamma(n) \cong \bar{\Gamma}(n)/\Lambda \quad (n \geq 3). \tag{1.4.9}$$

Both $\Gamma(n)$ and $\bar{\Gamma}(n)$ are normal subgroups of $\Gamma(1)$, and $\hat{\Gamma}(n)$ is a normal subgroup of $\hat{\Gamma}(1)$. For, if $S \in \Gamma(n)$ and $T \in \Gamma(1)$, then

$$T^{-1}ST \equiv T^{-1}IT \equiv I \pmod{n},$$

and the proof is similar in the other cases.

† German: *Hauptkongruenzgruppe.*

Clearly

$$\operatorname{lev} \Gamma(n) = \operatorname{lev} \bar{\Gamma}(n) = \operatorname{lev} \hat{\Gamma}(n) = n,$$

and also

$$\Delta(n) \subseteq \Gamma(n), \quad \hat{\Delta}(n) \subseteq \hat{\Gamma}(n).$$

Further, since $J^{-1}U^nJ = U^{-n}$ and $l(U^n) = n$, it follows from (1.2.22) that $\Delta(n)$ and $\Gamma(n)$ are invariant under the automorphism $T \mapsto J^{-1}TJ$, but not under the automorphism (1.2.22) unless n is even. On the other hand, $\bar{\Gamma}(n)$ is invariant under all automorphisms of $\Gamma(1)$, and $\hat{\Delta}(n)$ and $\hat{\Gamma}(n)$ are invariant under all automorphisms of $\hat{\Gamma}(1)$.

The factor groups

$$G(n) := \Gamma(1)/\Gamma(n), \quad \bar{G}(n) := \Gamma(1)/\bar{\Gamma}(n) \tag{1.4.10}$$

and

$$\hat{G}(n) := \hat{\Gamma}(1)/\hat{\Gamma}(n) \tag{1.4.11}$$

are called *modulary*† groups of level n. By (1.4.8–11) we have

$$\hat{G}(n) \cong \bar{G}(n) \cong G(n) \quad (n = 1, 2) \tag{1.4.12}$$

and

$$\hat{G}(n) \cong \bar{G}(n) \cong G(n)/\Lambda \quad (n \geq 3). \tag{1.4.13}$$

Since the number of incongruent matrices T modulo n is clearly less than or equal to n^4, the three modulary groups are clearly finite and their orders are denoted by $\mu(n)$, $\bar{\mu}(n)$ and $\hat{\mu}(n)$, respectively; $\mu(n)$ should not be confused with the Möbius function.

Theorem 1.4.1.

$$\mu(n) = n^3 \prod_{p|n} \left(1 - \frac{1}{p^2}\right),$$

where the product is taken over all primes p dividing n, and

$$\hat{\mu}(n) = \bar{\mu}(n) = \mu(n) \quad (n = 1, 2), \quad \hat{\mu}(n) = \bar{\mu}(n) = \tfrac{1}{2}\mu(n) \quad (n \geq 3).$$

Proof. By (1.4.12, 13), it is enough to find $\mu(n)$, i.e. the number of incongruent matrices S modulo n. Our proof is similar to that given by Gunning (1962). We say that a pair of integers c, d is a *primitive*

† German: *Modulargruppe.*

pair modulo n if and only if $(c, d, n) = 1$, and denote by $\lambda(n)$ the number of incongruent primitive pairs modulo n.

Lemma 1. *If c, d is a primitive pair modulo n there exists an $S \in \Gamma(1)$ such that $\gamma \equiv c$, $\delta \equiv d \pmod{n}$. Conversely, if $S \in \Gamma(1)$, then γ, δ is a primitive pair modulo n.*

Proof. The last part is obvious since $(\gamma, \delta) = 1$, so that $(\gamma, \delta, n) = 1$.

As usual we write $q|r$ to mean that q divides r. Suppose that c, d is a primitive pair modulo n, so that $(c, d, n) = 1$. Take $\gamma = c$ and write

$$c = c_1 c_2,$$

where c_1 is the largest divisor of c that is prime to n. Take $m \in \mathbb{Z}$ so that

$$\delta := d + mn \equiv 1 \pmod{c_1},$$

which is possible since $(n, c_1) = 1$. Then $(\gamma, \delta) = 1$. For if p is a prime divisor of γ and δ, then p divides both c and $d + mn$. If $p|c_1$ then, since $\delta \equiv 1 \pmod{c_1}$, $p \nmid \delta$, which is false. If $p|c_2$, then $p|n$ and so $p|d$; thus $p|(c, d, n)$, which is false. Hence we have found γ, δ with $\gamma \equiv c$, $\delta \equiv d \pmod{n}$ and $(\gamma, \delta) = 1$. We can now find $\alpha, \beta \in \mathbb{Z}$ such that $\alpha\delta - \beta\gamma = 1$ and so $S \in \Gamma(1)$.

Lemma 2. *For each primitive pair, c, d of integers modulo n there are exactly n matrices $S \in \Gamma(1)$ modulo n for which*

$$\gamma \equiv c, \quad \delta \equiv d \pmod{n}.$$

Proof. Let S_1 and S_2 be two members of $\Gamma(1)$ with second rows congruent to $[c, d]$ modulo n. Then

$$S_1 S_2^{-1} \equiv \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \pmod{n},$$

for some $k \in \mathbb{Z}$. As there are only n incongruent values of k modulo n, it follows that there are exactly n incongruent matrices $S \in \Gamma(1)$ with $[\gamma, \delta] \equiv [c, d] \pmod{n}$.

Lemma 3. *$\lambda(n)$ is a multiplicative function; i.e., if $(n_1, n_2) = 1$, then $\lambda(n_1 n_2) = \lambda(n_1)\lambda(n_2)$.*

Proof. Let γ_j, δ_j be a primitive pair modulo n_j $(j = 1, 2)$. Then $\gamma_1 n_2 + \gamma_2 n_1$, $\delta_1 n_2 + \delta_2 n_1$ is a primitive pair modulo $n_1 n_2$, since $(n_1, n_2) = 1$. Also, incongruent pairs for n_1 and n_2 lead to incongruent pairs for $n_1 n_2$; i.e., if

$$\gamma_1' n_2 + \gamma_2' n_1 \equiv \gamma_1 n_2 + \gamma_2 n_1 \pmod{n_1 n_2}$$

and

$$\delta_1' n_2 + \delta_2' n_1 \equiv \delta_1 n_2 + \delta_2 n_1 \pmod{n_1 n_2},$$

then $\gamma_1' \equiv \gamma_1 \pmod{n_1}$, $\delta_1' \equiv \delta_1 \pmod{n_1}$, $\gamma_2' \equiv \gamma_2 \pmod{n_2}$ and $\delta_2' \equiv \delta_2 \pmod{n_2}$. Thus $\lambda(n_1)\lambda(n_2) \leq \lambda(n_1 n_2)$. Conversely, let γ, δ be a primitive pair modulo $n_1 n_2$; then γ, δ is a primitive pair modulo n_1 and modulo n_2. Also, since $(n_1, n_2) = 1$, incongruent pairs modulo $n_1 n_2$ cannot give rise to congruent ones modulo n_1 and modulo n_2. Thus $\lambda(n_1 n_2) \leq \lambda(n_1)\lambda(n_2)$.

Lemma 4. *If p is prime and $k \geq 1$, $\lambda(p^k) = p^{2k}(1 - p^{-2})$.*

Proof. There are $p^k(1 - 1/p)$ incongruent integers c modulo p^k such that $(c, p) = 1$. For any one of these, each of the p^k incongruent values of d will give a primitive pair. Since these pairs are all incongruent modulo p^k we have $p^{2k}(1 - 1/p)$ such pairs. There are p^{k-1} values of c such that $(c, p) = p$. To each of these correspond $p^k(1 - 1/p)$ values of d incongruent modulo p^k and such that $(d, p) = 1$. This gives $p^{2k-1}(1 - 1/p)$ primitive pairs. Addition gives

$$\lambda(p^k) = p^{2k}\left(1 - \frac{1}{p^2}\right).$$

From lemmas 3 and 4 we obtain

$$\lambda(n) = n^2 \prod_{p|n} \left(1 - \frac{1}{p^2}\right)$$

and, since $\mu(n) = n\lambda(n)$, by lemma 2, the theorem follows.

Theorem 1.4.2. *If m and n are positive integers,*

$$\Gamma(m) \cap \Gamma(n) = \Gamma(\{m, n\}) \tag{1.4.14}$$

and

$$\Gamma(m)\Gamma(n) = \Gamma((m, n)), \tag{1.4.15}$$

where $\{m, n\}$ is the least common multiple of m and n.

Proof. Write $h=(m,n)$, $l=\{m,n\}$. We have $T\in\Gamma(m)\cap\Gamma(n)$ if and only if $T\equiv I \pmod m$ and $T\equiv I \pmod n$; this holds if and only if $T\equiv I \pmod l$. This proves (1.4.14).

If $T\in\Gamma(m)\Gamma(n)$ then $T=S_1S_2$ where $S_1\equiv I \pmod m$ and $S_2\equiv I \pmod n$. Hence $S_1\equiv I \pmod h$ and $S_2\equiv I \pmod h$, so that $T\equiv I \pmod h$. It follows that

$$\Gamma(m)\Gamma(n)\subseteq\Gamma(h). \tag{1.4.16}$$

Now, by (1.4.14) and one of the isomorphism theorems,

$$\{\Gamma(m)\Gamma(n)\}/\Gamma(n)\cong\Gamma(m)/\Gamma(l), \tag{1.4.17}$$

so that

$$\begin{aligned}\mu(n)&=[\Gamma(1):\Gamma(n)]=[\Gamma(1):\Gamma(m)\Gamma(n)][\Gamma(m)\Gamma(n):\Gamma(n)]\\&=[\Gamma(1):\Gamma(m)\Gamma(n)][\Gamma(m):\Gamma(l)]\\&=[\Gamma(1):\Gamma(m)\Gamma(n)]\mu(l)/\mu(m).\end{aligned}$$

Since $\mu(l)\mu(h)=\mu(m)\mu(n)$ by the formula in theorem 1.4.1 it follows that

$$[\Gamma(1):\Gamma(m)\Gamma(n)]=\mu(h)$$

and this combined with (1.4.16) gives (1.4.15).

Theorem 1.4.3. *If* $(m,n)=1$, *then*

$$G(mn)\cong G(m)\times G(n).$$

Proof. By (1.4.15, 17),

$$\Gamma(1)/\Gamma(n)\cong\Gamma(m)/\Gamma(mn) \tag{1.4.18}$$

so that we can identify $G(n)$ with $\Gamma(m)/\Gamma(mn)$ and $G(m)$ with $\Gamma(n)/\Gamma(mn)$. These are both subgroups of $\Gamma(1)/\Gamma(mn)=G(mn)$. We have, by theorem 1.4.2,

$$G(m)\cap G(n)=\{I\},\quad G(m)nG(n)=G(mn),$$

where I is the identity in $G(mn)$. From this the theorem follows; see theorem 2.5.1 of Hall (1959).

By repeated applications of theorem 1.4.3 we see that, if

$$n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k},$$

where the p_i are different primes and $\alpha_i>0$ $(1\le i\le k)$, then $G(n)$ is the direct product of the k groups $G(p_i^{\alpha_i})$. Thus the structure of $G(n)$ is known when we know the structure of $G(p^k)$ for each prime p and $k\in\mathbb{Z}^+$. It follows from (1.4.8, 9) that some knowledge of the structure of $\hat{G}(n)$ may be obtained when we know the structure of the groups $\hat{G}(p^k)$.

So far we have considered principal congruence groups only. If Γ is a subgroup of $\Gamma(1)$ and if, for some $n\in\mathbb{Z}^+$,

$$\Gamma(n)\subseteq\Gamma\subseteq\Gamma(1), \tag{1.4.19}$$

we say that Γ is a *congruence group*; Γ is then necessarily of finite index in $\Gamma(1)$. Further, since $U^n\in\Gamma$, it follows that the level of Γ is a divisor of n.

Suppose that Γ is a congruence group satisfying (1.4.19) and that

$$\Gamma(1)=\Gamma(n)\cdot\mathscr{R}.$$

It follows that $\Gamma=\Gamma(n)\cdot\mathscr{R}_0$, where $\mathscr{R}_0$ is a subset of $\mathscr{R}$; i.e. Γ is the set of all matrices T that are congruent to a member of $\mathscr{R}_0$ modulo n. Congruence groups are often defined in this way in terms of a finite set $\mathscr{R}_0$ of matrices.

As an example of a non-principal congruence group, let

$$\Gamma_0(n)=\{T\in\Gamma(1): c\equiv 0 \pmod n\}. \tag{1.4.20}$$

That $\Gamma_0(n)$ is a group follows from (1.1.5, 11). Clearly

$$\hat{\Gamma}_0(n)\cong\Gamma_0(n)/\Lambda.$$

Similarly, we define

$$\Gamma^0(n)=\{T\in\Gamma(1): b\equiv 0 \pmod m\} \tag{1.4.21}$$

and

$$\begin{aligned}\Gamma_0^0(n)&=\{T\in\Gamma(1): b\equiv c\equiv 0 \pmod n\}\\&=\Gamma_0(n)\cap\Gamma^0(n).\end{aligned} \tag{1.4.22}$$

Since

$$V^{-1}TV=\begin{bmatrix} d & -c\\ -b & a\end{bmatrix},$$

$\Gamma_0(n)$ and $\Gamma^0(n)$ are conjugate subgroups of $\Gamma(1)$ and

$$\Gamma^0(n)=V^{-1}\Gamma_0(n)V,\quad \hat{\Gamma}^0(n)=V^{-1}\hat{\Gamma}_0(n)V.$$

Since $\Gamma(n) \in \Gamma^0(n)$ and $U^r \in \Gamma^0(n)$ only when n divides r, it follows that

$$n = \operatorname{lev} \Gamma^0(n) = \operatorname{lev} \Gamma_0(n) = \operatorname{lev} \Gamma_0^0(n)$$

and similar results hold for the corresponding inhomogeneous groups.

We now calculate the index of $\Gamma_0(n)$ in $\Gamma(1)$.

If $c \equiv 0 \pmod{n}$ and $ad - bc = 1$, it follows that $(d, n) = 1$, and this holds for

$$\phi(n) := n \prod_{p \mid n} \left(1 - \frac{1}{p}\right)$$

incongruent values of d modulo n. Each of the $\phi(n)$ pairs c, d is a primitive pair modulo n, and lemmas 1 and 2 show that there are exactly $n\phi(n)$ incongruent matrices modulo n $\Gamma_0(n)$. It follows that

$$\psi(n) := [\Gamma(1) : \Gamma_0(n)] = n \prod_{p \mid n} \left(1 + \frac{1}{p}\right) \tag{1.4.23}$$

and that

$$\hat{\psi}(n) := [\hat{\Gamma}(1) : \hat{\Gamma}_0(n)] = \psi(n). \tag{1.4.24}$$

Values of $\hat{\mu}(n)$ and $\hat{\psi}(n)$ are given for $n \leq 16$ in table 2.

Table 2

n	$\hat{\mu}(n)$	$\hat{\psi}(n)$	n	$\hat{\mu}(n)$	$\hat{\psi}(n)$
1	1	1	9	324	12
2	6	3	10	360	18
3	12	4	11	660	12
4	24	6	12	576	24
5	60	6	13	1092	14
6	72	12	14	1008	24
7	168	8	15	1440	24
8	192	12	16	1536	24

Theorems 1.4.4. *For any positive integers n and N,*

$$\Gamma(N)\Gamma^0(n) = \Gamma^0((N, n)), \quad \Gamma(N)\Gamma_0(n) = \Gamma_0((N, n)). \tag{1.4.25}$$

Proof. It suffices to prove the first identity. Write $N_1 = (N, n)$. Since $\Gamma(N) \subseteq \Gamma^0(N_1)$ and $\Gamma^0(n) \subseteq \Gamma^0(N_1)$ we need only prove that

$$\Gamma^0(N_1) \subseteq \Gamma(N)\Gamma^0(n).$$

Take any $S \in \Gamma^0(N_1)$, so that $\beta \equiv 0 \pmod{N_1}$ and therefore

$$\alpha\delta \equiv 1 \pmod{N_1}.$$

It follows that $(\delta, N_1) = 1$. For any integer r write

$$S' = W^{rN}S = \begin{bmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{bmatrix},$$

say. Then

$$\delta' = \delta + \beta rN = (\delta, N)(A + rB),$$

say, where $(A, B) = 1$. We can therefore choose r so that $A + rB$ is a prime greater than n. It follows that

$$(\delta', n) = ((\delta, N), n) = (\delta, N, n) = (\delta, N_1) = 1,$$

and therefore

$$(N\delta', n) = (N, n) = N_1.$$

Accordingly we can choose an integer s such that

$$\beta + sN\delta' \equiv 0 \pmod{n}.$$

It follows that, for this choice of r and s,

$$U^{sN}W^{rN}S \in \Gamma^0(n)$$

and so $S \in \Gamma(N)\Gamma^0(n)$.

Observe that we have in fact proved slightly more, namely that

$$\Delta(N)\Gamma^0(n) = \Gamma^0((N, n)). \tag{1.4.26}$$

We also require to know the index in $\Gamma(1)$ of the group

$$\Gamma_0(m, n) := \Gamma_0(m) \cap \Gamma^0(n). \tag{1.4.27}$$

In particular, we note that $\Gamma_0^0(n) = \Gamma_0(n, n)$ and that $[\Gamma_0^0(n) : \Gamma(n)]$ is $\varphi(n)$, since it is the number of solutions modulo n of the congruence $ad \equiv 1 \pmod{n}$.

Now the m matrices U^{nr} $(0 \leq r < m)$ are easily seen to form a left transversal of $\Gamma_0(m, mn)$ in $\Gamma_0(m, n)$. Hence

$$[\Gamma_0(m, n) : \Gamma_0(m, mn)] = m,$$

and, similarly,

$$[\Gamma_0(m, mn):\Gamma_0(mn, mn)] = n.$$

Accordingly,

$$[\Gamma_0(m, n):\Gamma(mn)] = mn\varphi(mn),$$

from which we deduce that

$$[\hat{\Gamma}(1):\hat{\Gamma}_0(m, n)] = [\Gamma(1):\Gamma_0(m, n)] = \psi(mn). \quad (1.4.28)$$

In particular, the index of $\Gamma_0^0(n)$ in $\Gamma(1)$ is

$$\psi(n^2) = n\psi(n).$$

Theorem 1.4.5. *Let m, n and q be positive integers. Then*

$$\Gamma_0(m, n) = \langle \Gamma_0(m, nq), U^n \rangle.$$

Proof. It is enough to prove that, if $S \in \Gamma_0(m, n)$, integers r and s can be found such that

$$S' := U^m S U^{sn} \in \Gamma_0(m, nq).$$

We have, in an obvious notation,

$$\beta' = (\alpha + rn\gamma)sn + \beta + rn\delta = n\{(\alpha + rn\gamma)s + \beta_1 + r\delta\},$$

say. Now $(\alpha, n\gamma) = (\alpha, n) = 1$, since $\alpha\delta \equiv 1 \pmod{n}$, so that we can choose r to make $\alpha + rn\gamma$ a prime greater than q. Then $(\alpha + rn\gamma, q) = 1$, and therefore s can be chosen to make $\beta' \equiv 0 \pmod{nq}$, as required.

1.5. Groups of level 2. We shall need to study such groups when we introduce theta functions. We note that the following six matrices, defined in (1.1.4) and (1.2.2, 5, 6) form a set of coset representatives of $\Gamma(1)$ modulo $\Gamma(2)$.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad V = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

$$W = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \quad P^2 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

The corresponding mappings form a transversal of $\hat{\Gamma}(2)$ in $\hat{\Gamma}(1)$.

We have

$$P^3 \equiv U^2 \equiv V^2 \equiv W^2 \equiv I \pmod{2},$$

$$P \equiv VU, \quad P^2 \equiv UV, \quad W = UVU \equiv VUV \pmod{2}.$$

This illustrates the fact that $G(2)$ is isomorphic to the dihedral group of order 6, i.e. the symmetric group on three symbols. In fact we can set up a correspondence between

$$I, U, V, W, P, P^2$$

and the identical permutation e, (12), (13), (23), (123), (132), respectively; we use here the usual cycle notation.

Between $\Gamma(1)$ and $\Gamma(2)$ we have four groups

$$\Gamma_P(2), \Gamma_U(2), \Gamma_V(2), \Gamma_W(2)$$

as in fig. 1, where the index is marked and where normal subgroups

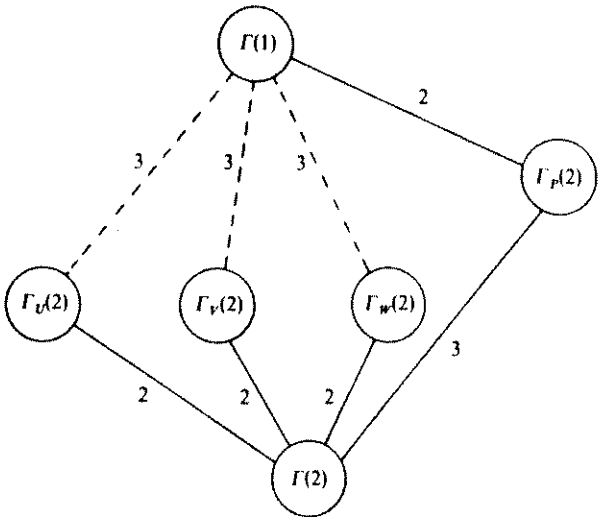


Fig. 1. Groups between $\Gamma(1)$ and $\Gamma(2)$.

are indicated with a continuous line. Here

$$\Gamma_U(2) := \{S \in \Gamma(1): S \equiv I \text{ or } U \pmod{2}\}, \quad (1.5.1)$$

and $\Gamma_V(2)$, $\Gamma_W(2)$ are defined similarly, while

$$\Gamma_P(2) := \{S \in \Gamma(1): S = I, P \text{ or } P^2 \pmod{2}\}. \quad (1.5.2)$$

Note that $\Gamma_U(2)=\Gamma_0(2)$, $\Gamma_W(2)=\Gamma^0(2)$. The three subgroups $\Gamma_U(2)$, $\Gamma_V(2)$ and $\Gamma_W(2)$ are conjugate; in fact

$$\Gamma_V(2)=P^{-1}\Gamma_U(2)P, \quad \Gamma_W(2)=P^{-2}\Gamma_U(2)P^2. \tag{1.5.3}$$

Each of the four groups just defined clearly has level 2.

The matrix $-I$ belongs to all six homogeneous groups mentioned above, so that an exactly similar figure can be drawn for the associated inhomogeneous groups.

The group $\Gamma_P(2)$ is, in fact, the group Γ^2 defined in (1.3.6). To prove this it is enough to show that $\Gamma_P(2)\subseteq\Gamma^2$, since both groups have index 2 in $\Gamma(1)$. This follows since Γ^2 is generated by P and P_1, where P_1 is defined by (1.3.7), and $P_1\equiv P^2 \pmod 2$. We have also

$$\hat{\Gamma}_P(2)=\hat{\Gamma}^2.$$

For certain purposes it is convenient to define a homogeneous group $\Gamma^*(2)$, which does not contain $-I$ and whose associated inhomogeneous group $\hat{\Gamma}^*(2)$ is identical with $\hat{\Gamma}(2)$. We define

$$\Gamma^*(2):=\{S\in\Gamma(1):\alpha\equiv\delta\equiv 1 \pmod 4,\ \beta\equiv\gamma\equiv 0 \pmod 2\}. \tag{1.5.4}$$

It is clear that this group has the properties stated and that

$$\Gamma(4)\subseteq\Gamma^*(2)\subseteq\Gamma(2).$$

$\Gamma^*(2)$ is of index 12 in $\Gamma(1)$ and contains $\Gamma(4)$ as a subgroup of index 4; see fig. 6 (p. 82). It is easily verified that $\Gamma^*(2)$ has normalizer $\Gamma_V(2)$. Also $\Gamma^*(2)$ has two conjugate subgroups in $\Gamma(1)$, namely $P^{-1}\Gamma^*(2)P$ and $P^{-2}\Gamma^*(2)P^2$, whose normalizers are $\Gamma_U(2)$ and $\Gamma_W(2)$, respectively.

We note in conclusion that although both $\hat{\Gamma}^*(1)$ and $\hat{\Gamma}(2)$ are normal subgroups of index 6 in $\hat{\Gamma}(1)$, they are distinct, since their quotient groups differ.

1.6. Groups of level 3. It is easily verified that the transformations associated with the following twelve matrices constitute a transversal of $\hat{\Gamma}(3)$ in $\hat{\Gamma}(1)$.

$$\left.\begin{array}{cccc} I, & V, & V_1 & V_2, \\ U, & U^2, & W, & W^2, \\ P, & P^2, & P_1, & P_1^2. \end{array}\right\} \tag{1.6.1}$$

Here P_1, V_1, V_2 are defined by (1.3.7, 10) and the remaining matrices are defined in the previous section. $\hat{G}(3)$ is isomorphic to the alternating group on four symbols and the transformations listed in (1.6.1) can be put into one-to-one correspondence with the permutations

$$\begin{array}{cccc} e, & (12)(34), & (13)(24), & (14)(23), \\ (234), & (243), & (134), & (143), \\ (124), & (142), & (132), & (123), \end{array}$$

respectively. We write

$$\Gamma_U(3)=\{T\in\Gamma(1):\pm T=I, U \text{ or } U^2 \pmod 3\} \tag{1.6.2}$$

and conjugate subgroups $\Gamma_W(3)$, $\Gamma_P(3)$ and $\Gamma_{P_1}(3)$ are defined similarly, U being replaced by W, P, P_1 at each occurrence in (1.6.2). Then

$$\Gamma_W(3)=V^{-1}\Gamma_U(3)V, \quad \Gamma_P(3)=V_2^{-1}\Gamma_U(3)V_2,$$
$$\Gamma_{P_1}=V_1^{-1}\Gamma_U(3)V_1. \tag{1.6.3}$$

Clearly $\Gamma_U(3)=\Gamma_0(3)$ and $\Gamma_W(3)=\Gamma^0(3)$.

Further,

$$\Gamma^3=\{T\in\Gamma(1):\pm T\equiv I, V, V_1 \text{ or } V_2 \pmod 3\}; \tag{1.6.4}$$

for the set Γ on the right-hand side of (1.6.4) is clearly a subgroup of $\Gamma(1)$ containing $\bar{\Gamma}(3)$ as a subgroup of index 4, and so has index 3 in $\Gamma(1)$. Further, by (1.3.9), Γ^3 is contained in this subgroup Γ and, since $[\Gamma(1):\Gamma^3]=3$, $\Gamma^3=\Gamma$. The factor group $\Gamma^3/\bar{\Gamma}(3)$ is isomorphic to the four-group. There are three conjugate groups between $\bar{\Gamma}(3)$ and Γ^3, namely the groups

$$\Gamma_V(3), \quad \Gamma_{V_1}(3) \quad \text{and} \quad \Gamma_{V_2}(3)$$

where

$$\Gamma_V(3):=\{T\in\Gamma(1):\pm T\equiv I \text{ or } V \pmod 3\}, \tag{1.6.5}$$

and the other groups are defined similarly. We have

$$\Gamma_{V_1}(3)=P^{-1}\Gamma_V(3)P, \quad \Gamma_{V_2}(3)=P^{-2}\Gamma_V(3)P^2. \tag{1.6.6}$$

It is easy to see that

$$\Gamma_V(3)=\{T\in\Gamma(1):Q^{-1}TQ\equiv T \pmod 3\}, \tag{1.6.7}$$

where Q is given by (1.3.19).

Corresponding results for the associated inhomogeneous groups and their relative structure is illustrated in fig. 2. With the exception of the modular group itself, all the groups considered in this section have level 3.

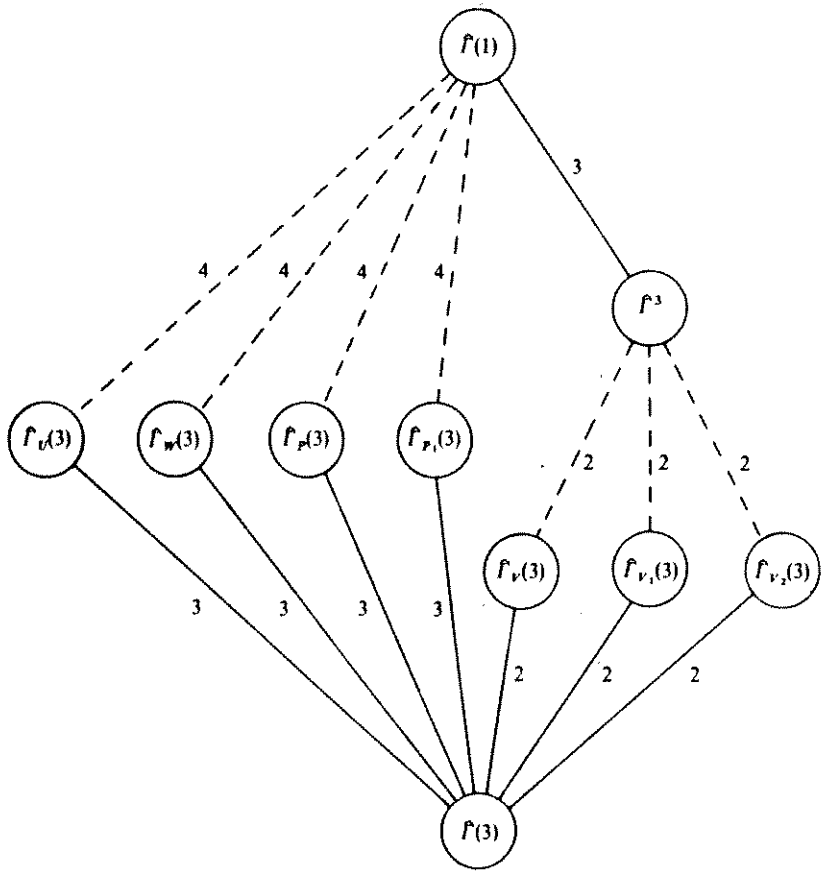


Fig. 2. Groups between $\hat{\Gamma}(1)$ and $\hat{\Gamma}(3)$.

1.7. Further results. We give here a brief sketch of further results of a similar kind, and some references for further reading.

From the monumental treatise by Klein and Fricke (1890, 1892) a great deal of information about particular subgroups of the modular group can be excavated by the persevering reader, and some of this information is given in a more easily assimilated form by Vivanti (1906, 1910). For example, the following alternative definitions can be given.

$$\hat{\Gamma}^2 = \{T: T \in \hat{\Gamma}(1),\ ab + bc + cd \equiv 0 \pmod 2\}, \tag{1.7.1}$$

$$\hat{\Gamma}^3 = \{T: T \in \hat{\Gamma}(1),\ ab + cd \equiv 0 \pmod 3\}, \tag{1.7.2}$$

$$\hat{\Gamma}''(1) = \hat{\Gamma}^2 \cap \hat{\Gamma}^3 = \{T: T \in \hat{\Gamma}(1),\ ab + 3bc + cd \equiv 0 \pmod 6\}. \tag{1.7.3}$$

See Klein and Fricke, vol. 1, p. 627; Weber (1909), §§54, 71; Petersson (1953); Wohlfahrt (1964). These definitions can be deduced from theorem 1.3.3 by writing the congruence defining G^ν in the form

$$T^{-1}QT \equiv \pm Q \pmod \nu.$$

We have seen in (1.4.20, 21) that each of the congruence subgroups $\Gamma_0(n)$ and $\Gamma^0(n)$ can be defined by a single linear congruence satisfied by entries of the matrices belonging to the group in question. The same holds for each group conjugate to $\Gamma_0(n)$. Thus, when $n = 2$ there are three conjugate subgroups, namely $\Gamma_0(2) = \Gamma_U(2)$, $\Gamma^0(2) = \Gamma_W(2)$ and

$$\Gamma_V(2) = P^{-1}\Gamma_0(2)P = \{T \in \Gamma(1): a + b - c - d \equiv 0 \pmod 2\}. \tag{1.7.4}$$

Similarly, when $n = 4$ we have again the three groups

$$\Gamma_0(4) =: \Gamma_U(4), \quad \Gamma^0(4) =: \Gamma_W(4)$$

and

$$\Gamma_V(4) := P^{-1}\Gamma_0(1)P = \{T \in \Gamma(1): a + b - c - d \equiv 0 \pmod 4\}. \tag{1.7.5}$$

Also, when $n = 3$, there are four conjugate subgroups listed in (1.6.2, 3), namely $\Gamma_U(3) = \Gamma_0(3)$, $\Gamma_W(3) = \Gamma^0(3)$ and also

$$\Gamma_P(3) = \{T \in \Gamma(1): a + b - c - d \equiv 0 \pmod 3\} \tag{1.7.6}$$

and

$$\Gamma_{P_1}(3) = \{T \in \Gamma(1): a - b + c - d \equiv 0 \pmod 3\}. \tag{1.7.7}$$

See Rankin (1973*b*).

Newman (1962) has studied a family of groups, which we denote by $\hat{\Gamma}^m$ ($m \in \mathbb{Z}^+$); here $\hat{\Gamma}^m$ denotes the subgroup of $\hat{\Gamma}(1)$ generated by the mth powers of elements of $\hat{\Gamma}(1)$. For $m = 2$, 3 it can be shown that the groups $\hat{\Gamma}^2$, $\hat{\Gamma}^3$ are the ones we have been studying.

For each $m \in \mathbb{Z}^+$, $\hat{\Gamma}^m$ is normal in $\hat{\Gamma}(1)$ and Newman has shown that

$$\hat{\Gamma}^m \hat{\Gamma}^n = \hat{\Gamma}^{(m,n)}, \quad \hat{\Gamma}^m = \hat{\Gamma}(1) \quad \text{if } (m, 6) = 1, \tag{1.7.8}$$

$$\hat{\Gamma}^{2m} = \hat{\Gamma}^2 \quad \text{if } (m, 3) = 1, \quad \hat{\Gamma}^{3m} = \hat{\Gamma}^3 \quad \text{if } (m, 2) = 1. \tag{1.7.9}$$

When $m \equiv 0 \pmod 6$, the structure of $\hat{\Gamma}^m$ is not known in all cases. Thus $\hat{\Gamma}^6$ is known to be a subgroup of index 216 in $\hat{\Gamma}(1)$ lying between $\hat{\Gamma}'(1)$ and $\hat{\Gamma}''(1)$, while, for $n > 1$, $[\hat{\Gamma}^6 : \hat{\Gamma}^{6n}] \geq n^{37}$ and may be infinite. See also Rankin (1969).

In the further algebraic study of the subgroups of the modular group the following three group-theoretic theorems are useful.

***Theorem 1.7.1* (Kurosh).** *Let the group G be a free product of subgroups A_α. We write this as $G = \prod_\alpha^* A_\alpha$. Then, if H is a subgroup of G, we have*

$$H = F * \prod_\beta^* B_\beta,$$

where F is a free group and, for each β, B_β is conjugate to one of the subgroups A_α.

We note that the free product $\prod_\beta^* B_\beta$ can be empty and that F can be the trivial group. However, if $\prod_\beta^* B_\beta$ is not empty, then F must have infinite index in G; for otherwise some power of an element of $\prod_\beta^* B_\beta$ (other than the identity) would belong to F.

***Theorem 1.7.2* (Schreier).** *Let H be a subgroup of finite index μ in a free group G of finite rank R. Then the rank r of H is also finite and*

$$r = 1 + \mu(R - 1).$$

We note that the rank of a free group is the number of free generators of the group. For proofs of both these theorems see Kurosh (1960).

***Theorem 1.7.3* (Nielsen).** *Let the group G have identity element e and be the free product of cyclic groups G_i of orders m_i $(1 \leq i \leq n)$. Then the commutator group G' is a free group of index $m = m_1 m_2 \ldots m_n$ in G and the rank of G' is*

$$1 + m\left\{-1 + \sum_{i=1}^{n}\left(1 - \frac{1}{m_i}\right)\right\}.$$

G' is generated by the set of all commutators $[xax^{-1}, xbx^{-1}]$, where $a \in G_i$, $b \in G_j$, $a \neq e \neq b$, $1 \leq i \leq j \leq n$ and

$$x = a_1 \ldots a_{i-1} a_{i+1} \ldots a_{j-1},$$

for any a_k in G_k $(1 \leq k \leq n)$. The factor group G/G' is isomorphic to the direct product of the cyclic groups $G_1, G_2, \ldots, G_n$.

See Nielsen (1948) and also Lyndon (1973); for the commutator notation see (1.3.12). Theorem 1.7.3 provides an alternative method of proving theorems 1.3.1 and 1.3.2.

With the help of theorem 1.7.1 we easily derive the following result of Newman (1964).

Theorem 1.7.4. *A subgroup of $\hat{\Gamma}(1)$ is free if and only if it contains no elements of finite order other than the identity. If $\hat{\Gamma}$ is a normal subgroup of $\hat{\Gamma}(1)$ different from $\hat{\Gamma}(1)$, $\hat{\Gamma}^2$ and $\hat{\Gamma}^3$, then $\hat{\Gamma}$ is a free group.*

The following theorem is a simple deduction from Schreier's theorem.

Theorem 1.7.5. *Let $\hat{\Gamma}$ be a free subgroup of $\hat{\Gamma}(1)$ of finite index μ. Then $\mu \equiv 0 \pmod 6$ and $\hat{\Gamma}$ has rank $1 + \frac{1}{6}\mu$. In particular, the index of any normal subgroup of $\hat{\Gamma}(1)$ other than $\hat{\Gamma}(1)$, $\hat{\Gamma}^2$ and $\hat{\Gamma}^3$ is divisible by 6.*

Proof. We shall prove the last sentence by analytic methods in chapter 2. The fact that $1 + \frac{1}{6}\mu$ is the rank of $\hat{\Gamma}$ is due to Mason (1969).

Write $\hat{\Delta} = \hat{\Gamma} \cap \hat{\Gamma}'(1)$ and put

$$\lambda = [\hat{\Gamma} : \hat{\Delta}], \quad \nu = [\hat{\Gamma}'(1) : \hat{\Delta}],$$

so that

$$\lambda\mu = [\hat{\Gamma}(1) : \hat{\Delta}] = 6\nu.$$

By theorem 1.7.2 applied to $\hat{\Delta}$ as a subgroup of $\hat{\Gamma}$ and $\hat{\Gamma}(1)$ we have

$$1 + (r - 1)\lambda = 1 + \nu,$$

where r is the rank of $\hat{\Gamma}$. Thus, $\mu = 6(r - 1)$, from which the theorem follows.

Fig. 3 displays all the normal subgroups between $\hat{\Gamma}(1)$ and $\hat{\Gamma}(6)$. The two groups $\hat{\Gamma}^{2'}$ and $\hat{\Gamma}^{3'}$ are the commutator groups of $\hat{\Gamma}^2$ and $\hat{\Gamma}^3$. By applying theorem 1.7.3 their indices in the groups immediately above them can be verified.

A complete study of all subgroups between $\hat{\Gamma}(1)$ and $\hat{\Gamma}(4)$ has been made by Petersson (1963). His list contains thirty groups, including $\hat{\Gamma}(4)$ and $\hat{\Gamma}(1)$; among these thirty groups there are contained, of course, the six groups in fig. 1. Petersson (1953) has also studied subgroups of finite index μ in $\hat{\Gamma}(1)$ for which

$$\hat{\Gamma}(1) = \hat{\Gamma} \cdot \bigcup_{k=0}^{\mu-1} U^k;$$

such a subgroup $\hat{\Gamma}$ is called a *cycloidal* subgroup of $\hat{\Gamma}(1)$.

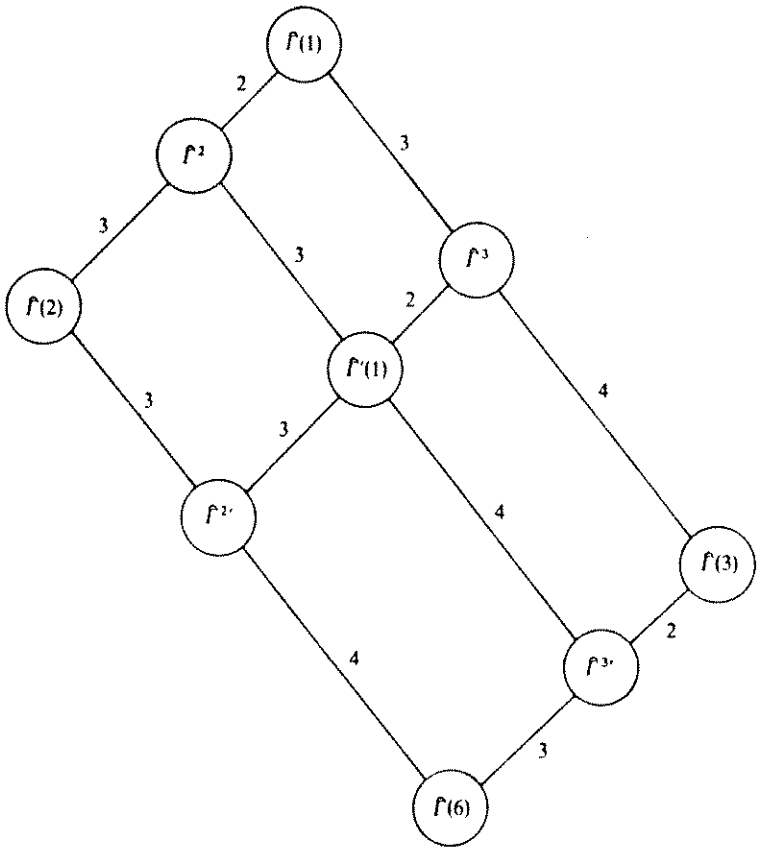


Fig. 3. Normal subgroups between $\hat{\Gamma}(1)$ and $\hat{\Gamma}(6)$.

We have already remarked that $\hat{G}(2)$ is isomorphic to the dihedral group of order 6 and that $\hat{G}(3)$ is isomorphic to the alternating group on four symbols, which is the tetrahedral group. It can be shown that $\hat{G}(4)$ and $\hat{G}(5)$ are isomorphic to the octahedral and icosahedral groups, respectively. In fact, for $n = 2, 3, 4, 5$, $\hat{G}(n)$ is generated by

$$u = U\hat{\Gamma}(n), \quad v = V\hat{\Gamma}(n),$$

where

$$u^n = v^2 = (uv)^3 = e,$$

e being the group identity.

A great deal of information about the modulary groups $\hat{G}(n)$ and $\hat{G}(p^k)$ can be found in Klein and Fricke (1890, 1892) and in Vivanti (1906, 1910); see also Frasch (1933). A complete classification of all the normal subgroups of $\hat{G}(n)$ has been given by McQuillan (1965).

The definition of level given in §1.4 is due to Wohlfahrt (1964). Previously the concept of level was only defined for congruence groups. Thus a congruence subgroup Γ of $\Gamma(1)$ was defined to be of level n if $\Gamma(n) \subseteq \Gamma$ and if n is the least positive integer for which this inclusion is valid. That the two definitions agree for congruence groups is a consequence of the relations

$$\Delta(n)\Gamma(qn) = \Gamma(n), \quad \hat{\Delta}(n)\hat{\Gamma}(qn) = \hat{\Gamma}(n), \qquad (1.7.10)$$

which were essentially proved by Wohlfahrt; see also Rankin (1969) and Leutbecher (1970). Here q is any positive integer.

If we now assume that $\operatorname{lev}\Gamma = n$, so that $\Delta(n) \subseteq \Gamma \subseteq \Gamma(1)$, and that Γ is a congruence group, then, for some $q \in \mathbb{Z}^+$, $\Gamma(q) \subseteq \Gamma$. From (1.7.10) we deduce that $\Gamma(n) \subseteq \Gamma$ and it is clear that n is the least positive integer for which this holds. Thus the two definitions agree.

It has been proved by various authors that $\Delta(n) = \Gamma(n)$ for $1 \leq n \leq 5$ and that $[\Gamma(n) : \Delta(n)] = \infty$ for $n \geq 6$. In fact $\Delta(6) = \Gamma''(1)$ and $[\Gamma(n) : \Delta(n)\Gamma'(n)] = \infty$ for $n \geq 6$. Similar results hold in the inhomogeneous case. It follows that every subgroup of level 2, 3, 4 or 5 must be a congruence subgroup. Also the groups described in §§1.5, 6 are the only subgroups of levels 2 and 3.

There exist subgroups of finite index in $\hat{\Gamma}(1)$ that are not congruence subgroups. That this is the case was first stated by

Klein in 1879; see Klein and Fricke (1890) (pp. 308, 418, 659–63), where a number of examples of groups and their associated functions are given. Proofs were given by Fricke (1887) and Pick (1887), who produced an example of such a subgroup of index 54. All the subgroups of $\hat{\Gamma}(1)$ of index $\mu \leq 7$ are listed in Rankin (1969). For $\mu \leq 6$ these are all congruence subgroups, but 28 of the 42 subgroups of index 7 are not congruence subgroups and fall into four classes each containing seven conjugate subgroups. In recent years various authors have discovered large families of 'non-congruence' subgroups. See, for example, Reiner (1958), Newman (1965, 1968), Rankin (1967*b*) and Mason (1969).

It is possible to give a rather complicated formula for the number of subgroups of $\hat{\Gamma}(1)$ of given index μ and to give methods, in certain cases, for specifying the number of normal subgroups; see Newman (1967).

2: Mapping properties

2.1. Conformal mappings. We begin by recalling some properties of meromorphic functions on the extended complex plane.

A *domain* $\mathbb{D}$ is a subset of $\bar{\mathbb{C}}$ that is open and connected in the topology on $\bar{\mathbb{C}}$; in particular, $\mathbb{H}$ is a domain.

Let f be a function defined on a domain $\mathbb{D}$. In order to discuss the behaviour of f at ∞, when $\infty \in \mathbb{D}$, we make use of the homeomorphism

$$z^* = 1/z$$

between neighbourhoods of $z = \infty$ and $z^* = 0$, and write

$$f^*(z^*) = f(z),$$

so that f^* is defined on a neighbourhood of 0 when $\infty \in \mathbb{D}$.

The statement that f is *holomorphic* on $\mathbb{D}$ means that (i) f maps $\mathbb{D}$ into $\mathbb{C}$, (ii) f is differentiable (with a finite derivative) on $\mathbb{D} - \{\infty\}$, (iii) if $\infty \in \mathbb{D}$, f^* is differentiable on some neighbourhood of 0. Note that, when $\infty \in \mathbb{D}$, the derivatives of f and f' are related by the equation

$$z^* f^{*\prime}(z^*) = -z f'(z)$$

on punctured neighbourhoods of $z^* = 0$ and $z = \infty$.

We say that f is holomorphic at a point $p \in \bar{\mathbb{C}}$ if f is holomorphic on some domain $\mathbb{D}$ containing p. In particular, f is holomorphic at ∞ if f^* is holomorphic on some domain containing 0; such a domain can always be taken in the form $|z^*| < r$ for some $r > 0$.

We say that f is meromorphic at a point $p \in \mathbb{C}$ if there exists an integer $q \geq 0$ such that h is holomorphic at p, where

$$h(z) \stackrel{!}{=} (z-p)^q f(z);$$

this implies in particular that h is continuous at p and that

$$h(p) = \lim_{z \to p} (z-p)^q f(z).$$

We say that f is meromorphic at ∞ if f^* is meromorphic at 0. A point p at which f is meromorphic but not holomorphic (so that

$q>0$) is called a *pole* of f. The smallest value of q that satisfies the conditions is called the *order* of the pole. If we can take $q=1$, the pole is simple. At a pole p of f we have $f(p)=\infty$.

More generally, if f is meromorphic at $p\in\mathbb{C}$, we define $\operatorname{ord}(f,p)$ to be the greatest integer q such that $(z-p)^{-q}f(z)$ is holomorphic at p. If $q>0$, q is the order of the zero of f at p; if $q<0$, $-q$ is the order of the pole of f at p.

Suppose that f maps a domain $\mathbb{D}$ onto a subset $\mathbb{D}_1$ of $\bar{\mathbb{C}}$. we say that f maps $\mathbb{D}$ *conformally* onto $\mathbb{D}_1$, or that f is conformal on $\mathbb{D}$, if and only if f is meromorphic on $\mathbb{D}$ and f is a bijective map of $\mathbb{D}$ onto $\mathbb{D}_1$. If f is conformal on the domain $\mathbb{D}$, it follows that its derivative f' exists and is non-zero at all points of $\mathbb{D}$ that are not poles of f. Also, if p is a pole of f in $\mathbb{D}$, then p is a simple pole. Further, the inverse function f^{-1} of a conformal mapping f is again conformal. The proofs of these facts involve the use of Rouché's theorem.

Since a holomorphic function is continuous, it follows that every conformal mapping f is a homeomorphism of its domain $\mathbb{D}$ onto $f(\mathbb{D})$, which is also a domain.

Conformal mappings have the property that angles between intersecting curves in the domain $\mathbb{D}$ are preserved in $f(\mathbb{D})$ both as regards magnitude and direction.

We now consider the particular case when

$$f(z):=T(z)=\frac{az+b}{cz+d},$$

where $T\in\hat{\Theta}$. As we saw in §1.1, T is a bijection of $\bar{\mathbb{C}}$ onto itself.

If $c=0$, then

$$f(z)=\frac{a}{d}z+\frac{b}{d}.$$

The mapping T is holomorphic on $\mathbb{C}$ and there is a simple pole at ∞. If $c\neq 0$, the mapping T is holomorphic on $\bar{\mathbb{C}}-\{-d/c\}$ and there is a simple pole at $-d/c$. That is, in every case T is holomorphic on $\bar{\mathbb{C}}-\{-d/c\}$ with a simple pole at $-d/c$. Hence T maps $\bar{\mathbb{C}}$ conformally onto $\bar{\mathbb{C}}$. Note also that, for $z\neq -d/c$ and $z\neq\infty$,

$$T'(z)=\frac{1}{(cz+d)^2}. \tag{2.1.1}$$

The converse result that, if f maps $\bar{\mathbb{C}}$ conformally onto $\bar{\mathbb{C}}$, then $f\in\hat{\Theta}$, also holds.

It can be shown similarly that f maps $\mathbb{C}$ conformally onto $\mathbb{C}$ if and only if $f\in\hat{\Theta}$ and $f(\infty)=\infty$. Further, f maps $\mathbb{H}$ conformally onto $\mathbb{H}$ if and only if $f\in\hat{\Omega}$. In these results the 'if' parts are straightforward, while the 'only if' parts require the use of Schwarz's lemma.

Since we shall not require the 'only if' parts, we content ourselves by showing that, if $T\in\hat{\Omega}$, then $T\mathbb{H}=\mathbb{H}$. For this purpose we write

$$w=:u+iv=T(z)=\frac{az+b}{cz+d},\quad z=:x+iy,$$

where u, v, x and y are real. Now

$$v=\operatorname{Im}T(z)=\frac{y}{|cz+d|^2}, \tag{2.1.2}$$

and conversely

$$y=\frac{v}{|cw-a|^2}. \tag{2.1.3}$$

It follows that $w\in\mathbb{H}$ if and only if $z\in\mathbb{H}$.

We note also that, if $T\in\hat{\Omega}$, then

$$T\bar{\mathbb{R}}=\bar{\mathbb{R}},\quad T\bar{\mathbb{H}}=\bar{\mathbb{H}}$$

and, if $T\in\hat{\Gamma}(1)$,

$$T\mathbb{P}=\mathbb{P},\quad T\mathbb{H}'=\mathbb{H}'.$$

Further, circles and straight lines are mapped by T onto circles or straight lines.

It is sometimes convenient to write the transformation $T\in\hat{\Theta}$ in a different form. Let z_1 and z_2 be two different finite complex numbers having finite images $w_j=Tz_j$ $(j=1,2)$. Then the transformation

$$w=\frac{az+b}{cz+d}$$

can be written in the form

$$\frac{w-w_1}{w-w_2}=K\frac{z-z_1}{z-z_2}, \tag{2.1.4}$$

where K is a constant. Further, if z_3 is another finite point having a known finite image w_3, then K is uniquely determined, namely

$$K=\frac{w_3-w_1}{w_3-w_2}\Big/\frac{z_3-z_1}{z_3-z_2}.$$

Alternatively, since $\infty = T(-d/c)$ we have

$$K = \frac{cz_2 + d}{cz_1 + d}, \tag{2.1.5}$$

and this holds also if $c = 0$.

2.2. Fixed points. For any $T \in \Theta$ and $z \in \mathbb{C}$ we define

$$T\!:\! z = cz + d. \tag{2.2.1}$$

Observe that, although Tz and $(-T)z$ are the same,

$$(-T)\!:\! z = -(T\!:\! z).$$

In particular, $I\!:\! z = 1$, $(-I)\!:\! z = -1$, and, more generally,

$$\pm U^n\!:\! z = \pm 1 \quad (n \in \mathbb{Z}).$$

By (1.1.4), for any S, $T \in \hat{\Theta}$ and $z \neq \infty$, $T^{-1}\infty$,

$$\begin{aligned} ST\!:\! z &= (\gamma a + \delta c)z + \gamma b + \delta d = (\gamma Tz + \delta)(cz + d) \\ &= (S\!:\! Tz)(T\!:\! z). \end{aligned}$$

We thus have the important identity

$$ST\!:\! z = (S\!:\! Tz)(T\!:\! z). \tag{2.2.2}$$

A particular case of this is

$$1 = (T^{-1}\!:\! Tz)(T\!:\! z). \tag{2.2.3}$$

The equation (2.2.2) holds, in particular, for any S, $T \in \Omega$ and all $z \in \mathbb{H}$. If $T \in \Omega$ and $z \in \mathbb{H}$, $T\!:\! z$ is finite and non-zero.

A point $z \in \bar{\mathbb{C}}$ is called a *fixed point* of a mapping $T \in \hat{\Theta}$ if and only if $Tz = z$. We suppose in the first place that $c \neq 0$, so that $z \neq \infty$. It then follows by induction from (2.2.2, 3) that

$$T^n\!:\! z = (T\!:\! z)^n \tag{2.2.4}$$

for any $n \in \mathbb{Z}$, where T is the associated matrix.

The equation $Tz = z$ is equivalent to

$$cz^2 + (d-a)z - b = 0,$$

which has two, not necessarily distinct, roots, namely

$$z_1, z_2 := \frac{(a-d) \pm [(a+d)^2 - 4]^{\frac{1}{2}}}{2c}. \tag{2.2.5}$$

We note also that

$$z_1 z_2 = -b/c, \quad z_1 + z_2 = (a-d)/c,$$

so that

$$(T\!:\! z_1)(T\!:\! z_2) = (cz_1 + d)(cz_2 + d) = 1. \tag{2.2.6}$$

It is clear that, when $T \in \hat{\Omega}$, the nature of the roots z_1, z_2 depends upon the sign of the real number $(a+d)^2 - 4$. If (i) $|\mathrm{tr}\ T| < 2$, then $(a+d)^2 - 4 < 0$ and z_1 and z_2 are conjugate complex numbers, one of which, say z_1, lies in $\mathbb{H}$. T is then called an *elliptic transformation* and z_1 and z_2 are called elliptic fixed points. If (ii) $|\mathrm{tr}\ T| = 2$, then $z_1 = z_2$ and we have one real fixed point. T is called a *parabolic transformation* and z_1 is called a parabolic fixed point. Finally (iii), if $|\mathrm{tr}\ T| > 2$, then z_1 and z_2 are distinct real numbers and T is called a *hyperbolic transformation* and z_1 and z_2 are called hyperbolic fixed points.

We now examine these possibilities in greater detail, for the case when $T \in \Gamma(1)$, still making the assumption that $c \neq 0$. Our object is to express the mapping T in the form (2.1.4), i.e.

$$\frac{w - w_1}{w - w_2} = K\frac{z - z_1}{z - z_2} = \frac{T\!:\! z_2}{T\!:\! z_1} \cdot \frac{z - z_1}{z - z_2}, \tag{2.2.7}$$

where this is possible; see (2.1.5). Here $w = Tz$ and $w_j = Tz_j$ $(j = 1, 2)$. Note that, although $T\!:\! z_2$ and $T\!:\! z_1$ change sign when we replace the matrix T by $-T$, their ratio remains unchanged.

(i) *Elliptic transformations.* Here $|\mathrm{tr}\ T| < 2$, and, by theorem 1.2.3, there are two possibilities:

(a) $\mathrm{tr}\ T = 0$, $T = \pm L^{-1} V L$ for some $L \in \Gamma(1)$,

or

(b) $\mathrm{tr}\ T = \pm 1$, $T = \pm L^{-1} P^s L$ for $s = 1, 2$ and some $L \in \Gamma(1)$.

In case (a) $Tz = z$ is equivalent to $V(Lz) = Lz$; i.e. Lz is a fixed point for V and so $Lz_1 = i$, $Lz_2 = -i$. Hence

$$z_1 = L^{-1} i, \quad z_2 = L^{-1}(-i) = \bar{z}_1. \tag{2.2.8}$$

Here the bar denotes the complex conjugate. Since $T^2 = V^2 = -I$, we have, by (2.2.4), $(T\!:\! z_j)^2 = -1$ $(j = 1, 2)$ and therefore, by

(2.2.6), $K=(T:z_2)/(T:z_1)=-1$. Hence

$$\frac{w-z_1}{w-z_2}=-\frac{z-z_1}{z-z_2}. \tag{2.2.9}$$

The transformation T is of order 2.

In case (b), $Tz=z$ is equivalent to $P^s(Lz)=Lz$, where $s=1$ or 2. An elementary calculation shows that both P and P^2 have fixed points

$$\rho=e^{2\pi i/3} \tag{2.2.10}$$

and $\bar{\rho}$, so that $Lz_1=\rho$, $Lz_2=\bar{\rho}$. Hence

$$z_1=L^{-1}\rho, \quad z_2=L^{-1}\bar{\rho}. \tag{2.2.11}$$

Since $T^3=\pm P^{3s}=\pm I$, (2.2.4) gives $T:z_1=\pm\rho$ or $\pm\rho^2$ and so $T:z_2=\pm\rho^2$ or $\pm\rho$, by (2.2.6). Hence $K=\rho$ or ρ^2, and the transformation takes the form

$$\frac{w-z_1}{w-z_2}=K\frac{z-z_1}{z-z_2} \quad (K=\rho \text{ or } \rho^2). \tag{2.2.12}$$

It is a transformation of order 3.

(ii) *Parabolic transformations.* Here $\operatorname{tr} T=\pm 2$ and, by theorem 1.2.3, $T=\pm L^{-1}U^qL$ for some $q\in\mathbb{Z}$ $(q\neq 0)$ and $L\in\Gamma(1)$. Thus $Tz=z$ is equivalent to $U^q(Lz)=Lz$, i.e. $(Lz)+q=Lz$. Hence $Lz=\infty$ and the single fixed point is

$$z_1=L^{-1}\infty. \tag{2.2.13}$$

Since $c\neq 0$, z_1 is a finite rational number.

Put $a+d=2\varepsilon$, where $\varepsilon=\pm 1$. Then, by (2.2.5), $z_1=(a-d)/(2c)$ so that $T:z_1=\varepsilon=\frac{1}{2}\operatorname{tr} T$. If $w=Tz$,

$$\frac{1}{w-z_1}=\frac{cz+d}{(a-cz_1)(z-z_1)}=\frac{c(z-z_1)+\varepsilon}{\varepsilon(z-z_1)}.$$

Hence the transformation T can be expressed in the form

$$\frac{1}{w-z_1}=\frac{1}{z-z_1}+c\varepsilon. \tag{2.2.14}$$

So far we have assumed that $c\neq 0$. When $c=0$, $T=\pm U^q$ for some $q\in\mathbb{Z}$ and $\operatorname{tr} T=\pm 2$. If $q=0$ we obtain the identical transformation under which every point is fixed. When $q\neq 0$, $T=\pm L^{-1}U^qL$ (with $L=I$) and we include T among the set of parabolic transformations. There is just one fixed point, namely $z_1=\infty$. In place of (2.2.14), the transformation T takes the canonical form

$$w=z+q. \tag{2.2.15}$$

(iii) *Hyperbolic transformations.* Here $|\operatorname{tr} T|>2$ and, since z_2 is real, (2.1.5) and (2.2.6) give $K=(T:z_2)^2>0$. The transformation has the canonical form (2.2.7) with $K>0$; clearly $K\neq 1$, since $z_1\neq z_2$.

From the above analysis we see that the transformations $T\in\hat{\Gamma}(1)$ can be divided into five classes:

1. *The identity transformation*, whose matrix $T=\pm I$.

2. *Elliptic transformations of order 2.* The matrix T of such a mapping is conjugate to $\pm V$ and satisfies $T^2=-I$. We denote by

$$\mathbb{E}_2=\{z\in\mathbb{C}: z=L^{-1}i,\ L\in\hat{\Gamma}(1)\} \tag{2.2.16}$$

the set of all elliptic fixed points of order 2 in $\mathbb{H}$.

3. *Elliptic transformations of order 3.* The matrix T of such a mapping is conjugate to $\pm P$ or $\pm P^2$ and satisfies $T^3=\pm I$. We denote by

$$\mathbb{E}_3=\{z\in\mathbb{C}: z=L^{-1}\rho,\ L\in\hat{\Gamma}(1)\} \tag{2.2.17}$$

the set of all elliptic fixed points of order 3 in $\mathbb{H}$. Here ρ is defined by (2.2.10).

4. *Parabolic transformations.* The matrix T of such a mapping is conjugate to $\pm U^q$ $(q\in\mathbb{Z}, q\neq 0)$. The set of all parabolic fixed points is $\mathbb{P}$, since $\mathbb{P}=\{z\in\bar{\mathbb{C}}: z=L^{-1}\infty,\ L\in\hat{\Gamma}(1)\}$. Parabolic fixed points are also called *cusps* for a reason that will be clear later on.

5. *Hyperbolic transformations.* The fixed points of such transformations are less important in the theory. It is easy to see that they are all irrational numbers. Hyperbolic transformations are of infinite order.

Note that $T\mathbb{E}_2=\mathbb{E}_2$, $T\mathbb{E}_3=\mathbb{E}_3$ for all $T\in\hat{\Gamma}(1)$.

We also write

$$\mathbb{E}=\mathbb{E}_2\cup\mathbb{E}_3. \tag{2.2.18}$$

We now suppose that Γ is a subgroup of $\Gamma(1)$. The mappings $T\in\hat{\Gamma}$ can be divided into five classes in a similar way, but some of

these classes may be empty. Thus $\hat{\Gamma}$ will contain elliptic transformations only if it contains a mapping conjugate to V or P. For $m = 2, 3$ we denote by $\mathbb{E}_m(\Gamma)$ the set of all fixed points in $\mathbb{H}$ of elliptic transformations of order m belonging to $\hat{\Gamma}$. It follows that

$$\mathbb{E}_2(\Gamma) = \{z \in \mathbb{C}: z = L^{-1}i, L \in \hat{\Gamma}(1), L^{-1}VL \in \hat{\Gamma}\}, \quad (2.2.19)$$

$$\mathbb{E}_3(\Gamma) = \{z \in \mathbb{C}: z = L^{-1}\rho, L \in \hat{\Gamma}(1), L^{-1}PL \in \hat{\Gamma}\}. \quad (2.2.20)$$

Note that we can omit $L^{-1}P^2L$ since $L^{-1}P^2L \in \hat{\Gamma}$ if and only if $L^{-1}PL \in \hat{\Gamma}$ and has the same fixed points. We write

$$\mathbb{E}(\Gamma) = \mathbb{E}_2(\Gamma) \cup \mathbb{E}_3(\Gamma). \quad (2.2.21)$$

If $\hat{\Gamma}$ is normal in $\hat{\Gamma}(1)$ it is clear that $\mathbb{E}_m(\Gamma)$ is either $\mathbb{E}_m$ or the null set $\varnothing$.

We now suppose that z is any point of $\mathbb{H}'$. The *stabilizer* of $z \pmod{\Gamma}$ is defined to be the subset Γ_z of Γ consisting of all $T \in \Gamma$ for which $Tz = z$. Clearly Γ_z is a subgroup of Γ. The stabilizer of $z \pmod{\Gamma(1)}$ is denoted by $\Gamma_z(1)$; the corresponding inhomogeneous groups are denoted by $\hat{\Gamma}_z$ and $\hat{\Gamma}_z(1)$. Evidently $\hat{\Gamma}_z$ is a subgroup of $\hat{\Gamma}_z(1)$. Further, if $L \in \hat{\Gamma}(1)$ then

$$L^{-1}\hat{\Gamma}_{Lz}L = (L^{-1}\hat{\Gamma}L)_z. \quad (2.2.22)$$

The preceding discussion of fixed points shows that

$$\hat{\Gamma}_\infty(1) = \hat{\Gamma}_U$$

in the notation of §1.2. Also $\hat{\Gamma}_i(1)$ and $\hat{\Gamma}_\rho(1)$ are the cyclic groups of orders 2 and 3 generated by the mappings V and P, respectively. It follows immediately from this and (2.2.22) that

$$\hat{\Gamma}_z(1) = \begin{cases} L^{-1}\hat{\Gamma}_U L, & \text{when } z = L^{-1}\infty, \\ L^{-1}\hat{\Gamma}_i(1)L, & \text{when } z = L^{-1}i, \\ L^{-1}\hat{\Gamma}_\rho(1)L, & \text{when } z = L^{-1}\rho, \\ \hat{\Lambda} = \{I\} & \text{otherwise.} \end{cases} \quad (2.2.23)$$

Here L is any member of $\hat{\Gamma}(1)$, and $z \in \mathbb{H}'$. Note that, in every case, $\hat{\Gamma}_z(1)$ is a cyclic group.

We now define the *order* of $z \pmod{\Gamma}$ to be

$$n(z, \Gamma) := [\hat{\Gamma}_z(1) : \hat{\Gamma}_z]. \quad (2.2.24)$$

The index on the right is possibly infinite. However, if $\hat{\Gamma}$ has finite index in $\hat{\Gamma}(1)$ then $n(z, \Gamma)$ is finite. For if the cyclic group $\hat{\Gamma}_z(1)$ is generated by S it follows that $S^n \in \hat{\Gamma}$ for some finite positive integer n, and $n(z, \Gamma)$ is in fact the least such integer n. We note also that, if we take $\Gamma_1 = \hat{\Gamma}(1)$, $\Gamma_2 = \hat{\Gamma}$ in theorem 1.1.2, with S as above then, for $1 \le i \le m$,

$$\sigma_i = n(z, L_i^{-1}\Gamma L_i) = n(L_i z, \Gamma). \quad (2.2.25)$$

The transformation S is parabolic, elliptic or the identity according as $z \in \mathbb{P}$, $\mathbb{E}$ or $\mathbb{H}' - \mathbb{P} \cup \mathbb{E}$, respectively; in fact, we can take $S = L^{-1}UL, L^{-1}VL, L^{-1}PL$ and I in the four cases listed in (2.2.23). In the latter case $n(z, r) = 1$. When $z \in \mathbb{P}$, we also call $n(z, \Gamma)$ the *width of the cusp* $z \pmod{\Gamma}$; if $z = L^{-1}\infty$, $\hat{\Gamma}_z$ is generated by $L^{-1}U^nL$, where $n = n(z, r)$. When $z \in \mathbb{E}_m$ $(m = 2, 3)$, $n(z, \Gamma) = 1$ or m according as z does or does not belong to $\mathbb{E}_m(\Gamma)$.

It can be shown similarly that, when $\hat{\Gamma}$ has finite index in $\hat{\Gamma}(1)$, some positive power of every hyperbolic transformation belongs to $\hat{\Gamma}$, and is, of course, hyperbolic. Hence $\hat{\Gamma}$ always contains hyperbolic transformations.

We note, in conclusion, that if $\Gamma = \Gamma'(1)$ or $\Gamma(N)$ for $N > 1$, then $\mathbb{E}(\Gamma) = \varnothing$ and so $n(z, \Gamma) = m$ if $z \in \mathbb{E}_m$ $(m = 2, 3)$. Further, if $z \in \mathbb{P}$,

$$n(z, \Gamma'(1)) = 6 \quad \text{and} \quad n(z, \Gamma(N)) = N.$$

2.3. Fundamental regions. Let Γ be a subgroup of $\Gamma(1)$, so that the mappings T in $\hat{\Gamma}$ map $\mathbb{H}'$ onto itself. In what follows we shall only be interested in subsets of $\mathbb{H}'$, so that we are not, for example, interested in hyperbolic fixed points.

Two points z_1, z_2 in $\mathbb{H}'$ are said to be *congruent*, or *equivalent*, $(\text{mod } \Gamma)$, if there exists a $T \in \hat{\Gamma}$ such that

$$z_2 = Tz_1.$$

It is easily verified that this is an equivalence relation, and we write

$$z_2 \equiv z_1 \pmod{\Gamma}.$$

The equivalence class containing a point $z \in \mathbb{H}'$ is called the *orbit* of $z \pmod{\Gamma}$ and is denoted by $\hat{\Gamma}z$; instead of $\hat{\Gamma}(1)z$ we may write $[z]$. Thus $\mathbb{E}_2 = [i]$, $\mathbb{E}_3 = [\rho]$ and, if $\hat{\Gamma}$ is of finite index in $\hat{\Gamma}(1)$, $\mathbb{P} = \hat{\Gamma}\infty = [\infty]$. Clearly $n(z_1, \Gamma) = n(z_2, \Gamma)$ when $z_1 \equiv z_2 \pmod{\Gamma}$.

A subset $\mathbb{F}$ of $\mathbb{H}'$ is called a *proper fundamental region* for $\hat{\Gamma}$ if $\mathbb{F}$ contains exactly one point from each orbit $\hat{\Gamma}z$. By giving $\mathbb{F}$ the quotient topology induced by the topology on $\mathbb{H}$ (compactified

suitably at points of $\mathbb{P}$) and the equivalence relation, $\mathbb{F}$ can be made into a connected Hausdorff space which is, in fact, the Riemann surface associated with the group $\hat{\Gamma}$. We shall not, however, use any Riemann surface theory, although we may mention this theory at various points. In practice it is usually convenient to impose further conditions on $\mathbb{F}$, such as that it is a simply connected subset of $\mathbb{H}'$ and is bounded by curves of a prescribed form.

Theorem 2.3.1. *Let $\mathbb{F}$ be a proper fundamental region for a subgroup $\hat{\Gamma}$ of $\hat{\Gamma}(1)$ and suppose that $\mathbb{F}=\bigcup_{n=1}^{\infty}\mathbb{F}_n$, where the sets $\mathbb{F}_n$ are disjoint. Then, if $T_n\in\hat{\Gamma}$ for each $n\in\mathbb{Z}^+$, the set $\bigcup_{n=1}^{\infty}T_n\mathbb{F}_n$ is also a proper fundamental region for $\hat{\Gamma}$.*

This is obvious, as all we have done is to choose Tz as a representative of $\hat{\Gamma}z$ rather than z for some $T\in\hat{\Gamma}$. The theorem is useful since it enables us to piece together fundamental regions in alternative ways that may be convenient for special purposes. Usually the number of non-null regions $\mathbb{F}_n$ is finite.

Theorem 2.3.2. *Let $\mathbb{F}$ be a proper fundamental region for a subgroup $\hat{\Gamma}$ of $\hat{\Gamma}(1)$ and suppose that $T\in\hat{\Gamma}(1)$. Then* (i) *$T\mathbb{F}$ is a proper fundamental region for the conjugate group $T\hat{\Gamma}T^{-1}$.* (ii) *In particular, if $T\in\hat{\Gamma}$ and $T\neq\pm I$, then $T\mathbb{F}$ is a proper fundamental region for $\hat{\Gamma}$ and $\mathbb{F}\cap T\mathbb{F}$ is either empty or consists of a single point ζ, which is a fixed point for T.* (iii) *A fixed point ζ for a mapping $T\in\hat{\Gamma}$ cannot be an interior point of $\mathbb{F}$.* (iv) *The regions $T\mathbb{F}$ for $T\in\hat{\Gamma}$ cover $\mathbb{H}'$ without overlapping; if T_1 and T_2 are different transformations in $\hat{\Gamma}$, then $T_1\mathbb{F}$ and $T_2\mathbb{F}$ have at most one point in common.*

Proof. (i) If $z\in\mathbb{H}'$ and $T\in\hat{\Gamma}(1)$, then $T^{-1}z\in\mathbb{H}'$ and so there exists an $S\in\hat{\Gamma}$ such that $ST^{-1}z\in\mathbb{F}$. Then $TST^{-1}z\in T\mathbb{F}$. Further, if $TS_1T^{-1}z$ and $TS_2T^{-1}z$ are two points of $T\mathbb{F}$, where S_1 and S_2 are in $\hat{\Gamma}$, then $S_1T^{-1}z$ and $S_2T^{-1}z$ are points in the same orbit $\hat{\Gamma}T^{-1}z$ lying in $\mathbb{F}$ and so are identical. The original points $TS_1T^{-1}z$ and $TS_2T^{-1}z$ are therefore also identical.

(ii) Let $T\in\hat{\Gamma}$, $T\neq\pm I$ and suppose that $\zeta\in\mathbb{F}\cap T\mathbb{F}$. Then $\zeta\in\mathbb{F}$ and $T^{-1}\zeta\in\mathbb{F}$ and, since these points are congruent (mod Γ), $\zeta=T^{-1}\zeta$; i.e. $T\zeta=\zeta$. Hence ζ is a fixed point for T and there is only one such point in $\mathbb{H}'$.

(iii) If ζ is an interior point of $\mathbb{F}$, then there exists a neighbourhood $\mathbb{N}$ of ζ with $\mathbb{N}\subseteq\mathbb{F}$. But $T\mathbb{N}$ is a neighbourhood of $T\zeta=\zeta$ and so

therefore is

$$\mathbb{N}'=\mathbb{N}\cap T\mathbb{N}.$$

But $\mathbb{N}'\subseteq\mathbb{N}\subseteq\mathbb{F}$ and $\mathbb{N}'\subseteq T\mathbb{N}\subseteq T\mathbb{F}$, so that $\mathbb{N}'\subseteq\mathbb{F}\cap T\mathbb{F}$, which is false since $\mathbb{F}\cap T\mathbb{F}=\{\xi\}$.

(iv) That $\mathbb{H}'\subseteq\bigcup_{T\in\hat{\Gamma}}T\mathbb{F}$ is obvious. Further $\zeta\in T_1\mathbb{F}\cap T_2\mathbb{F}$ if and only if $T_1^{-1}\zeta\in\mathbb{F}\cap T_1^{-1}T_2\mathbb{F}$ and the last part follows from this and (ii).

When $\hat{\Gamma}=\hat{\Gamma}_{U^k}$ it is easy to find a fundamental region.

Theorem 2.3.3. *Let $n\in\mathbb{Z}^+$, $\delta\geq 0$ and put*

$$\mathbb{S}_k(\delta):=\{z\in\mathbb{H}':\ -\tfrac{1}{2}k\leq\operatorname{Re}z<\tfrac{1}{2}k,\ \operatorname{Im}z\geq\delta\}. \tag{2.3.1}$$

Also put

$$\mathbb{S}_k:=\mathbb{S}_k(0). \tag{2.3.2}$$

Then $\mathbb{S}_k$ is a proper fundamental region for $\hat{\Gamma}_{U^k}$. Further, for each $\zeta\in\mathbb{H}'$, $[\zeta]\cap\mathbb{S}_k(\delta)$ is a finite set when $\delta>0$.

Proof. Since $\hat{\Gamma}_{U^k}$ consists of the transformations

$$w=z+kn\quad(n\in\mathbb{Z}),$$

it is obvious that $\mathbb{S}_k$ is a fundamental region for $\hat{\Gamma}_{U^k}$.

Take $\zeta\in\mathbb{H}'$ and $\delta>0$, and put $\zeta=\xi+i\eta$, where ξ, η are real. If $\eta=0$, then $[\zeta]\cap\mathbb{S}_k(\delta)=\varnothing$, so that we may assume that $\eta>0$. If $T\zeta\in\mathbb{S}_k(\delta)$ for any $T\in\hat{\Gamma}(1)$, then, by (2.1.2),

$$c^2(\xi^2+\eta^2)+2cd\xi+d^2\leq\eta/\delta.$$

For given ξ, η, δ the number of pairs of integers c, d that satisfy this inequality is finite; for the inequality states that the point (c, d) lies in a certain ellipse. For each coprime pair of integers c, d satisfying the inequality we can find a matrix $T'\in\Gamma(1)$ with second row $[c, d]$; any other such matrix T is given by $T=U^nT'$ for some $n\in\mathbb{E}$. If $T\zeta\in\mathbb{S}_k(\delta)$, then $T'\zeta+n\in\mathbb{S}_k(\delta)$ and this can hold for at most k different values of n. This completes the proof.

Corollary 2.3.3. *The set $\mathbb{E}$ is a countable subset of isolated points of $\mathbb{H}$.*

It follows immediately from the theorem that each point of $\mathbb{E}$ is isolated. Further,

$$\mathbb{E} = \bigcup_{n=1}^{\infty} \{([i] \cup [\rho]) \cap \mathbb{S}_n(1/n)\}$$

from which countability follows, as each subset $([i] \cup [\rho]) \cap \mathbb{S}_n(1/n)$ is finite.

Let $\hat{\Gamma}$ be a subgroup of $\hat{\Gamma}(1)$. A subset $\mathbb{F}$ of $\mathbb{H}'$ is called a *fundamental region* for $\hat{\Gamma}$ if $\mathbb{F}$ contains at least one point of every orbit $\hat{\Gamma}z$ ($z \in \mathbb{H}'$) and exactly one point whenever $z \notin \mathbb{E} \cup \mathbb{F}$. A proper fundamental region is therefore a fundamental region, and a fundamental region only differs from a proper fundamental region in the possible inclusion of a countable number of fixed points of $\hat{\Gamma}(1)$.

Theorem 2.3.4. *Theorem 2.3.1 holds with the word 'proper' omitted. So does theorem 2.3.2, except that in (ii) and (iv) the two different fundamental regions may intersect in more than one point of* $\mathbb{E} \cup \mathbb{P}$.

This follows immediately. We note that in the proof of part (iii) of theorem 2.3.2, $\mathbb{F} \cap T\mathbb{F}$ is a set of isolated points and so cannot contain $\mathbb{N}'$.

Theorem 2.3.5. *Let $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ be subgroups of $\hat{\Gamma}(1)$ and suppose that $\hat{\Gamma} \subseteq \hat{\Gamma}_1 = \hat{\Gamma}_2 \cdot \mathscr{R}$. Then, if $\mathbb{F}_1$ is a fundamental region for $\hat{\Gamma}_1$,*

$$\mathbb{F}_2 = \bigcup_{T \in \mathscr{R}} T\mathbb{F}_1$$

is a fundamental region for $\hat{\Gamma}_2$.

Proof. Let $z \in \mathbb{H}'$; then there exists an $S_1 \in \hat{\Gamma}_1$ such that $S_1 z \in \mathbb{F}_1$. Write $S_1^{-1} = S_2^{-1} T$, where $S_2 \in \hat{\Gamma}_2$ and $T \in \mathscr{R}$. Then $S_2 z = TS_1 z \in T\mathbb{F}_1 \subseteq \mathbb{F}_2$.

Conversely, suppose that z and $z' \in \mathbb{F}_2$, where $z' = S_2 z$ for $S_2 \in \hat{\Gamma}_2$, and that neither z nor z' is a fixed point of $\hat{\Gamma}(1)$. Then $z \in T\mathbb{F}_1$, $z' \in T'\mathbb{F}_1$, where T, $T' \in \mathscr{R}$. Hence $T^{-1}z$ and $T'^{-1}z'$ are congruent (mod Γ_1), lie in $\mathbb{F}_1$ and are not fixed points of $\hat{\Gamma}(1)$. They are therefore identical; i.e. $z = S_2^{-1}z' = S_2^{-1}T'T^{-1}z$. Since z is not a fixed point, we must have $S_2^{-1}T'T^{-1} = \pm I$; i.e. $T' \in \hat{\Gamma}_2 T$. This implies that $T' = T$ and $S_2 = \pm I$; i.e. $z' = z$.

Corollary 2.3.5. *Under similar assumptions, if $\hat{\Gamma}_1 = \mathscr{L} \cdot \hat{\Gamma}_2$, $\bigcup_{T \in \mathscr{L}} T^{-1}\mathbb{F}_1$ is a fundamental region for $\hat{\Gamma}_2$.*

This is proved similarly.

2.4. Construction of fundamental regions for $\hat{\Gamma}(1)$ and its subgroups. We denote by $\mathbb{F}_I$ the set

$$\mathbb{F}_I = \mathbb{F}^{(1)} \cup \mathbb{F}^{(2)}, \tag{2.4.1}$$

where

$$\mathbb{F}^{(1)} := \{z \in \bar{\mathbb{C}} : -\tfrac{1}{2} \leq \operatorname{Re} z \leq 0, |z| \geq 1\} \tag{2.4.2}$$

and

$$\mathbb{F}^{(2)} := \{z \in \mathbb{C} : 0 < \operatorname{Re} z < \tfrac{1}{2}, |z| > 1\}. \tag{2.4.3}$$

We include ∞ in $\mathbb{F}^{(1)}$, but not in $\mathbb{F}^{(2)}$. The closure of any transform $T\mathbb{F}^{(1)}$ or $T\mathbb{F}^{(2)}$ ($T \in \hat{\Gamma}(1)$) we call a *triangle*.

Theorem 2.4.1. *$\mathbb{F}_I$ is a proper fundamental region for $\hat{\Gamma}(1)$.*

Proof. Let ζ be any point of $\mathbb{H}'$. We show first that some member of the orbit $[\zeta]$ lies in $\mathbb{F}_I$. For this purpose we may assume that ζ is not congruent to any point on the frontier $\partial\mathbb{F}_I$ of $\mathbb{F}_I$, since every such point z either lies in $\mathbb{F}_I$, or else one of the congruent points $z - 1$, $-1/z$ does. We call $\eta = \operatorname{Im} \zeta$ the *height* of ζ.

We observe first that, for some $n \in \mathbb{Z}$, $\zeta_1 = U^n\zeta \in \mathbb{S}_1$ (see (2.3.2)) and has the same height as ζ. If $\zeta_1 \notin \mathbb{F}_I$, then $|\zeta| < 1$ and so $V\zeta_1 \in [\zeta]$ and has height greater than ζ; for

$$\operatorname{Im}(-1/\zeta_1) = (\operatorname{Im} \zeta_1)/|\zeta_1|^2.$$

Further, for some $m \in \mathbb{Z}$, $\zeta_2 = U^m V\zeta_1 = U^m VU^n\zeta \in \mathbb{S}_1$ and has height greater than η. Either $\zeta_2 \in \mathbb{F}_I$ or else we can continue the process and find a congruent point $\zeta_3 \in \mathbb{S}_1$ of greater height than ζ_2. Since $\mathbb{S}_1(\eta) \cap [\zeta]$ is finite, the process ultimately terminates after a finite number k of stages in the finding of a point $\zeta_k \in [\zeta] \cap \mathbb{F}_I$.

It remains to prove that each orbit contains only one point of $\mathbb{F}_I$. For suppose that z_1 and z_2 are different congruent points of $\mathbb{F}_I$. Then both are finite and we may assume that

$$y_2 = \operatorname{Im} z_2 \geq y_1 = \operatorname{Im} z_1.$$

Let $z_2 = Tz_1$, where $T \in \Gamma(1)$, so that

$$y_2 = \frac{y_1}{|cz_1 + d|^2}.$$

Hence $|cz_1 + d| \leq 1$.

We cannot have $|c| \geq 2$, since no circle of radius $r \leq \frac{1}{2}$ and orthogonal to $\mathbb{R}$ meets $\mathbb{F}_I$; also, if $c = 0$, then $T = U^k$, which is clearly impossible. We may therefore assume that $c = 1$. The only circles of unit radius centred at points of $\mathbb{Z}$ that meet $\mathbb{F}_I$ are the circles

$$|z| = 1 \quad \text{and} \quad |z+1| = 1.$$

There are thus two cases:

(i) $c = 1, d = 0, |z_1| = 1$, (ii) $c = d = 1, z_1 = \rho$.

In case (i) we must have $T = U^k V$ and either $k = 0$, $z_1 = z_2 = i$, or $k = -1$ and $z_1 = z_2 = \rho$. In case (ii), $T = U^k P$ and we must have $k = 0$, $z_1 = z_2 = \rho$. Hence, in both cases $z_1 = z_2$ and this completes the proof of the theorem.

It follows from theorem 2.3.1 that

$$\hat{\mathbb{F}}_I = \mathbb{F}^{(1)} \cup \{U^{-1}\mathbb{F}^{(2)}\} \tag{2.4.4}$$

is also a proper fundamental region for $\hat{\Gamma}(1)$. For each $T \in \hat{\Gamma}(1)$ we write

$$\mathbb{F}_T = T\mathbb{F}_I, \quad \hat{\mathbb{F}}_T = T\hat{\mathbb{F}}_I. \tag{2.4.5}$$

It follows from theorem 2.3.2 that

$$\mathbb{H}' = \bigcup_{T \in \hat{\Gamma}(1)} \mathbb{F}_T = \bigcup_{T \in \hat{\Gamma}(1)} \hat{\mathbb{F}}_T.$$

We note that the boundary (frontier) of $\mathbb{F}_I$ consists of four 'sides' l_U, Ul_U, l_V and Vl_V, where

$$l_U = \{z = x + yi : x = -\tfrac{1}{2}, \tfrac{1}{2}\sqrt{3} \leq y\} \tag{2.4.6}$$

and

$$l_V = \{z = x + yi : -\tfrac{1}{2} \leq x \leq 0, |z| = 1, y > 0\}. \tag{2.4.7}$$

l_U and l_V are contained in $\mathbb{F}_I$ but $Ul_U - \{\infty\}$ and $Vl_V - \{i\}$ are not.

Fig. 4 shows how the regions $\mathbb{F}_T$ fit together. The angles between sides contained in the regions in question are marked. That the

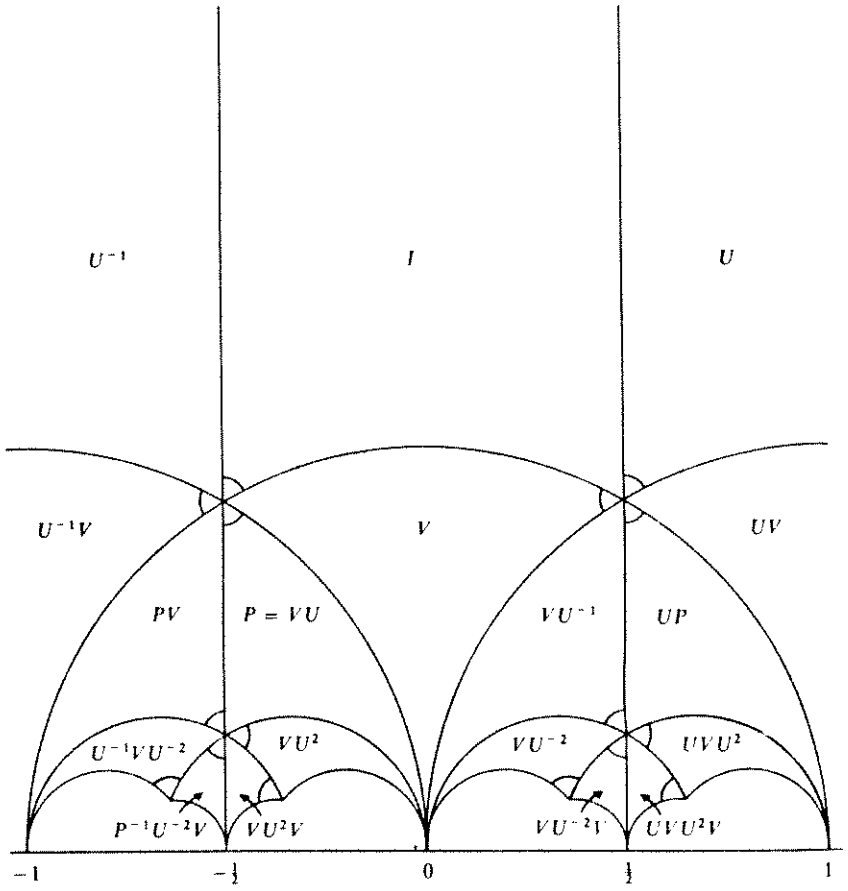


Fig. 4

regions cluster closer and closer to the real axis is shown by the following theorem.

Theorem 2.4.2. *If $\mathbb{F}_T$ contains a point ζ with $\eta = \operatorname{Im}\zeta \geq \delta$, then $|cd| \leq 1/\delta$.*

Proof. Suppose that $\zeta = Tz$ is such a point, so that $z \in \mathbb{F}_I$. Write $r = |z|$, $z = re^{i\theta}$ where $\frac{1}{3}\pi \leq \theta \leq \frac{2}{3}\pi$. If $\eta \geq \delta$ we have, by (2.1.2),

$$|cz + d|^2 = (cx + d)^2 + c^2y^2 \leq y/\delta;$$

i.e.

$$(cr + d)^2 \cos^2(\tfrac{1}{2}\theta) + (cr - d)^2 \sin^2(\tfrac{1}{2}\theta) \leq (r/\delta)\sin\theta.$$

Since $\cos^2(\frac{1}{2}\theta) \geq \frac{1}{4}$, $\sin^2(\frac{1}{2}\theta) \geq \frac{1}{4}$, we have

$$\frac{r}{\delta} \geq \frac{r \sin\theta}{\delta} \geq \frac{c^2r^2+d^2}{2} \geq |cd|r,$$

from which the theorem follows.

We note that since $(c, d) = 1$, the condition $|cd| \leq 1/\delta$ is satisfied by only a finite number of pairs c, d.

By theorem 1.2.1, $\hat{\Gamma}(1) = \hat{\Gamma}_U \cdot \mathscr{R}$, where $\mathscr{R}$ is a set of mappings T such that, for each $S \in \Gamma(1)$, there is exactly one $T \in \mathscr{R}$ with $[c, d] = \pm[\gamma, \delta]$. We choose such a set in the following way. We first take $T = I$, which corresponds to $\gamma = 0$, $\delta = 1$. Then, for each pair c, d with $c > 0$, $(c, d) = 1$ we choose a and b to satisfy

$$ad - bc = 1, \quad \left|\frac{a}{c}\right| \leq \frac{1}{2}.$$

The equation $a' = a + nc$ on p. 9 shows that this is possible. This determines T uniquely except when $c = 2$, in which case $d \neq 0$ and we take $a = \operatorname{sgn} d$, so that

$$\operatorname{Re} Ti = \frac{ac+bd}{c^2+d^2} = \frac{1}{2}\left\{1 - \frac{|d|}{4+d^2}\right\} \operatorname{sgn} d;$$

it follows that

$$-\tfrac{1}{2} < \operatorname{Re} Ti < \tfrac{1}{2} \tag{2.4.8}$$

in this case.

The set of all such T we call $\mathscr{R}_0$, and it is clearly a right transversal of $\hat{\Gamma}_U$ in $\hat{\Gamma}(1)$.

We now apply theorem 2.3.5 with this choice of $\mathscr{R}_0$ and $\Gamma_1 = \hat{\Gamma}(1)$, $\Gamma_2 = \hat{\Gamma}_U$, taking $\mathbb{F}_1 = \mathbb{F}_I$. We obtain a region

$$\mathbb{S}_U = \bigcup_{T \in \mathscr{R}_0} T\mathbb{F}_I = \bigcup_{T \in \mathscr{R}_0} \mathbb{F}_T, \tag{2.4.9}$$

which is a fundamental region for Γ_U. We show that $\mathbb{S}_U$ differs from the proper fundamental region $\mathbb{S}_1$ of theorem 2.3.3 only in the respect that the straight line segment

$$\lambda := \{z = x + iy : x = -\tfrac{1}{2}, 0 \leq y < \tfrac{1}{2}\}$$

in $\mathbb{S}_1$ is replaced by the congruent segment $U\lambda$, and the fixed point $Wi = \frac{1}{2}(1+i)$ is added.

To prove this we need only observe from fig. 4 that the line $l = \{z : x = -\frac{1}{2}\}$ is made up of sides of the regions $\mathbb{F}_I$, $\mathbb{F}_P$ and $\mathbb{F}_{VU^2V}$ and that no points of λ belong to any of these regions, while $U\lambda$ and $\frac{1}{2}(1+i)$ are contained in $\mathbb{F}_{VU^{-1}} \cup \mathbb{F}_{VU^{-2}V}$. Each of the five regions mentioned is a subset of $\mathbb{S}_U$. In particular, the lines l and Ul contain no interior points of any region $\mathbb{F}_T$. Hence every region $\mathbb{F}_T$ is either wholly contained in $\mathbb{S}_U$ or has no points in common with it except possibly the nine fixed points that lie on l and Ul; these are ∞, ρ, $-\rho^2$, $\frac{1}{2}(1-i)$, $\frac{1}{2}(1+i)$, $\frac{1}{6}(\rho-2)$, $\frac{1}{6}(\rho+4)$, $-\frac{1}{2}$ and $\frac{1}{2}$. Hence, if $|a/c| = |T\infty| < \frac{1}{2}$, then $\mathbb{F}_T \subseteq \mathbb{S}_U$, while, if $|T\infty| = \frac{1}{2}$, this is also true by our choice of T in (2.4.8) to make $|\operatorname{Re} Ti| < \frac{1}{2}$.

We now construct a fundamental region for a subgroup $\hat{\Gamma}$ of $\hat{\Gamma}(1)$ of finite index μ by applying theorem 2.3.5 with

$$\hat{\Gamma}_1 = \hat{\Gamma}(1), \quad \hat{\Gamma}_2 = \hat{\Gamma}.$$

We take $\mathbb{F}_1$ to be $\mathbb{F}_I$. The regions that we shall progressively construct will have boundaries consisting of a finite number of arcs of circles and segments of straight lines, these lines and circles being orthogonal to the real axis. Each arc or segment is called a *side* and has two endpoints called *vertices* which are points of $\mathbb{E} \cup \mathbb{P}$. Two such regions are said to be adjacent at a side l if their interiors are disjoint and if l is a side of each region.

The *vertex angle* at a vertex of a region is the interior angle between the sides meeting at the vertex in question. If ζ is a vertex of a fundamental region $\mathbb{F}$ for Γ, the *vertex set* $\mathbb{V}(\zeta, \mathbb{F})$ is defined to be the set of all vertices of $\mathbb{F}$ that are congruent to $\zeta \pmod{\Gamma}$. If $z_1 \equiv z_2 \pmod{\Gamma}$, where $z_1 \in \mathbb{E} \cup \mathbb{P}$, then z_1 and z_2 have the same order $\pmod{\Gamma}$. Hence points in the same vertex set have the same order $\pmod{\Gamma}$.

For every choice of $\mathscr{R}$ in theorem 2.3.5 we get a fundamental region $\mathbb{F}_2$ for $\hat{\Gamma}$, but it may happen that, if $\mathscr{R}$ is not chosen suitably, $\mathbb{F}_2$ will consist of several disjoint components. It is possible, however, to choose $\mathscr{R}$ in such a way that $\mathbb{F}_2$ is a connected subset of $\bar{\mathbb{C}}$. To do this we proceed as follows.

Let ζ_1 be any point of $\mathbb{P}$ and take any $L_1 \in \hat{\Gamma}(1)$ such that $\zeta_1 = L_1\infty$. Let $n_1 = n(\zeta_1, \Gamma)$, the order of $\zeta_1 \pmod{\Gamma}$ (see (2.2.24)), so that the mappings $L_1U^{k_1}$ $(0 \leq k_1 \leq n_1)$ belongs to different right cosets of $\hat{\Gamma}$ in $\hat{\Gamma}(1)$. Write

$$\mathbb{D}_1 := \bigcup_{k_1=0}^{n_1-1} L_1U^{k_1}\mathbb{F}_I.$$

$\mathbb{D}_1$ is formed from n_1 adjacent fundamental regions for $\hat{\Gamma}(1)$ and contains ζ_1 as a vertex. If $n_1 < \mu$, we choose, if possible, an $L_2 \in \Gamma(1)$ such that $L_2 \mathbb{F}_I$ is adjacent to $\mathbb{D}_1$ and such that L_2 belongs to none of the n_1 cosets mentioned above. Let n_2 be the order of $\zeta_2 = L_2\infty \pmod{\Gamma}$. Then the mappings $L_2 U^{k_2}$ $(0 \le k_2 < n_2)$, together with $L_1 U^{k_1}$ $(0 \le k_1 < n_1)$ all belong to different right cosets, since $L_2 U^s L_1^{-1} \notin \hat{\Gamma}$ for all $s \in \mathbb{Z}$. Write

$$\mathbb{D}_2 := \bigcup_{k_2=0}^{n_2-1} L_2 U^{k_2} \mathbb{F}_I.$$

Then $\mathbb{D}_2$ is formed from n_2 adjacent fundamental regions for $\hat{\Gamma}(1)$ and is adjacent to $\mathbb{D}_1$. We proceed in this way. We ultimately obtain a connected set

$$\mathbb{F} := \bigcup_{i=1}^{\lambda} \mathbb{D}_i,$$

where

$$\mathbb{D}_i := \bigcup_{k_i=0}^{n_i-1} L_i U^{k_i} \mathbb{F}_I,$$

$n_i = n(\zeta_i, \Gamma)$ and $\zeta_i = L_i\infty$ $(1 \le i \le \lambda)$. The mappings $L_i U^{k_i}$ $(0 \le k_i < n_i,\ 1 \le i \le \lambda)$ all belong to different right cosets of $\hat{\Gamma}$ in $\hat{\Gamma}(1)$ and

$$\mu' := \sum_{i=1}^{\lambda} n_i \le \mu.$$

Further, if $\mathbb{F}_T$ is adjacent to $\mathbb{F}$, for any $T \in \hat{\Gamma}(1)$, then $T \in \Sigma$, where Σ is the union of the μ' right cosets mentioned above.

Now $\mathbb{F}$ is a subset of a fundamental region for $\hat{\Gamma}$ and so, by theorem 2.3.4, the different regions $S\mathbb{F}$ for $S \in \hat{\Gamma}$ can only overlap at points of $\mathbb{E} \cup \mathbb{P}$. Further, each region $S\mathbb{F}$ has a finite number of sides and is adjacent at each side to some other region $S'\mathbb{F}$ $(S' \in \hat{\Gamma})$. It follows that, if $\zeta \in \mathbb{E}$, the finite number (2 or 6) of regions $\mathbb{F}_T$ that meet at ζ either are all contained in

$$\mathbb{H}_1 := \bigcup_{T \in \Sigma} \mathbb{F}_T,$$

or all contained in

$$\mathbb{H}_2 := \bigcup_{T \in \hat{\Gamma}(1) - \Sigma} \mathbb{F}_T.$$

Thus every point $z \in \mathbb{H}$, whether it is an interior or a boundary point of a region $\mathbb{F}_T$, is an interior point of $\mathbb{H}_1$ or $\mathbb{H}_2$. Since $\mathbb{H}$ is a connected open subset of $\bar{\mathbb{C}}$ and $\mathbb{H} \cap \mathbb{H}_1 \ne \emptyset$, it follows that $\mathbb{H} \cap \mathbb{H}_2 = \emptyset$. Hence $\Sigma = \hat{\Gamma}(1)$ and it follows that

$$\sum_{i=1}^{\lambda} n_i = \mu \tag{2.4.10}$$

and that $\mathbb{F}$ is a fundamental region for $\hat{\Gamma}$.

We note also that the cusps ζ_1 and ζ_2 are incongruent $(\bmod \Gamma)$. For otherwise, for some $S \in \hat{\Gamma}$, $L_1\infty = SL_2\infty$, which implies that $L_1 U^k L_2^{-1} \in \hat{\Gamma}$ for some $k \in \mathbb{Z}$; this is false. It follows that the λ cusps $\zeta_i = L_i\infty (1 \le i \le \lambda)$ of $\mathbb{F}$ are incongruent $(\bmod \Gamma)$. Hence the orbit $[\infty] (\bmod \Gamma(1))$ splits up into λ different orbits $\hat{\Gamma} L_i\infty \pmod{\Gamma}$ $(1 \le i \le \lambda)$. The number $\lambda = \lambda(\Gamma)$ depends only on $\hat{\Gamma}$ and is called the number of cusps of the group $\hat{\Gamma}$. Note that each of the λ orbits $\hat{\Gamma} L_i\infty$ is represented by a single vertex of $\mathbb{F}$ and so, for $1 \le i \le \lambda$, $\mathbb{V}(\zeta_i, \mathbb{F})$ consists of the single point ζ_i. In the general case, for arbitrary $\mathcal{R}$, λ is the number of different parabolic vertex sets.

It is clear that the same process could have been carried out by using $\hat{\mathbb{F}}_I$ in place of $\mathbb{F}_I$. Further, instead of grouping fundamental regions $\mathbb{F}_T$ at parabolic fixed points, we could have grouped them together at points of $\mathbb{E}_2$ (using $\mathbb{F}_I$) or $\mathbb{E}_3$ (using $\hat{\mathbb{F}}_I$); for the above analysis goes through similarly with V or P in place of U.

For $m = 2, 3$ let $\mathbb{E}_m$ split into $\varepsilon_m = \varepsilon_m(\Gamma)$ orbits $\hat{\Gamma} \zeta_i^{(m)} \pmod{\Gamma}$. Then we obtain, analogously to (2.4.10),

$$\mu = \sum_{i=1}^{e_2} n_i^{(2)} = \sum_{i=1}^{e_3} n_i^{(3)}, \tag{2.4.11}$$

where $n_i^{(m)} = n(\zeta_i^{(m)}, \Gamma)$; i.e. $n_i^{(m)}$ is 1 or m according as the point $\zeta_i^{(m)}$ is or is not a fixed point for $\hat{\Gamma}$. The number of incongruent sets of elliptic fixed points of order m for $\hat{\Gamma}$ is denoted by $e_m = e_m(\Gamma)$ $(m = 2, 3)$. Clearly

$$e_m\left(1 - \frac{1}{m}\right) = \sum_{i=1}^{\varepsilon_m} \left(1 - \frac{n_i^{(m)}}{m}\right) = \varepsilon_m - \frac{\mu}{m}. \tag{2.4.12}$$

Note that the equations (2.4.10, 11) are particular cases of (1.1.15) in the cases $S = U$, V and P.

We have therefore proved the following theorem.

Theorem 2.4.3. *Let $\hat{\Gamma}$ be a subgroup of $\hat{\Gamma}(1)$ of finite index μ and suppose that $\hat{\Gamma}(1) = \hat{\Gamma} \cdot \mathcal{R}$. Then* (i) *the regions*

$$\mathbb{F} := \bigcup_{T \in \mathcal{R}} \mathbb{F}_T \quad \text{and} \quad \hat{\mathbb{F}} := \bigcup_{T \in \mathcal{R}} \hat{\mathbb{F}}_T$$

are fundamental regions for $\hat{\Gamma}$. Further the order of any vertex of $\mathbb{F}$ (mod Γ) is equal to half the number of triangles in $\bar{\mathbb{F}}$ containing points of its vertex set.

(ii) *$\mathcal{R}$ can be chosen so that $\mathbb{F}$ is a connected subset of $\bar{\mathbb{C}}$ and so that one of the following conditions hold:*

(a) Each one of the $\lambda(\Gamma)$ cusps of $\hat{\Gamma}$ is represented by a single vertex ζ_i of $\mathbb{F}$, and $n(\zeta_i, \Gamma)$ fundamental regions for $\hat{\Gamma}(1)$ meet in $\mathbb{F}$ at ζ_i.

(b) Each of the $e_2(\Gamma)$ orbits (mod Γ) *of elliptic fixed points in $\mathbb{E}_2(\Gamma)$ is represented by a single vertex of $\mathbb{F}$ of vertex angle π. No points of $\mathbb{E}_2 - \mathbb{E}_2(\Gamma)$ lie on the boundary of $\mathbb{F}$.*

(iii) *$\mathcal{R}$ can be chosen so that $\hat{\mathbb{F}}$ is a connected subset of $\bar{\mathbb{C}}$ and so that one of the two following conditions hold:*

(a) Each one of the $\lambda(\Gamma)$ cusps of $\hat{\Gamma}$ is represented by a single vertex ζ_i of $\hat{\mathbb{F}}$ and $n(\zeta_1, \Gamma)$ fundamental regions for $\hat{\Gamma}(1)$ meet in $\hat{\mathbb{F}}$ at ζ_i.

(b) Each of the $e_3(\Gamma)$ orbits (mod Γ) *of elliptic fixed points in $\mathbb{E}_3(\Gamma)$ is represented by a single vertex of $\hat{\mathbb{F}}$ of vertex angle $\frac{2}{3}\pi$. No points of $\mathbb{E}_2 - \mathbb{E}_3(\Gamma)$ lie on the boundary of $\hat{\mathbb{F}}$.*

Theorem 2.4.4. *Let $\hat{\Gamma}$ be a subgroup of $\hat{\Gamma}(1)$ of finite index μ and suppose that $\hat{\Gamma}(1) = \hat{\Gamma} \cdot \mathcal{R}$ and that*

$$\mathbb{F} = \bigcup_{T \in \mathcal{R}} \mathbb{F}_T.$$

Then (i) *the sides of $\mathbb{F}$ can be grouped into pairs λ_j, λ_j' $(j = 1, 2, \ldots, s)$ in such a way that $\lambda_j \subseteq \mathbb{F}$, $\lambda_j' \cap \mathbb{F} \subseteq \mathbb{E} \cup \mathbb{P}$ and $\lambda_j' = L_j \lambda_j$, where $L_j \in \hat{\Gamma}$ $(j = 1, 2, \ldots, s)$.* (ii) *No side in any one pair is congruent* (mod Γ) *to any side in another pair.* (iii) *The regions $L_j^{-1}\mathbb{F}$ and $\mathbb{F}$ are adjacent at λ_j while $\mathbb{F}$ and $L_j\mathbb{F}$ are adjacent at λ_j'.* (iv) *If a point z describes λ_j in such a way that the interior of $\mathbb{F}$ is on the left, then $L_j z$ describes λ_j' with the interior of $\mathbb{F}$ on the right.* (v) *$\hat{\Gamma}$ is generated by the s transformations $L_1, L_2, \ldots, L_s$.*

A similar result holds for the fundamental region $\hat{\mathbb{F}}$ of theorem 2.4.3 (i).

Proof. Let $\lambda = Tl$ be a side of $\mathbb{F}$, where $l = l_R$ and R is U or V; see (2.4.6, 7). We suppose that $\lambda \subseteq \mathbb{F}$. Then $\lambda \subseteq \mathbb{F}_T \subseteq \mathbb{F}$ and $T \in \mathcal{R}$. Now $\mathbb{F}_{TR^{-1}}$ and $\mathbb{F}_T$ are adjacent at λ and so, since λ is a side of $\mathbb{F}$, $\mathbb{F}_{TR^{-1}} \not\subseteq \mathbb{F}$. Now, for some $L \in \hat{\Gamma}$ and $S \in \mathcal{R}$

$$TR^{-1} = L^{-1}S$$

and $\mathbb{F}_S \subseteq \mathbb{F}$. Then $\lambda' := L\lambda = SRl$, which is a side of $\mathbb{F}_S$ not contained in $\mathbb{F}_S$. λ' is also a side of $\mathbb{F}$, since the adjacent region $\mathbb{F}_{SR}$ is not contained in $\mathbb{F}$; for SR belongs to the same right coset as T and, if $\mathbb{F}_{SR} \subseteq \mathbb{F}$, then $SR = T$ and so $\mathbb{F}_{TR^{-1}} = \mathbb{F}_S \subseteq \mathbb{F}$, which is false. Since λ' is a side of $\mathbb{F}$, $\lambda = L^{-1}\lambda'$ is a side also of $L^{-1}\mathbb{F}$. Similarly λ' is also a side of $L\mathbb{F}$. Thus with each side λ of $\mathbb{F}$ contained in $\mathbb{F}$ there is associated $L \in \hat{\Gamma}$ such that $\lambda' = L\lambda$ is a side of $\mathbb{F}$ not contained in $\mathbb{F}$.

Each different side λ_j $(j \leq s)$ of $\mathbb{F}$ contained in $\mathbb{F}$ determines in this way a congruent side λ_j' $(j \leq s)$ of $\mathbb{F}$ not contained in $\mathbb{F}$; since $\mathbb{F}$ is a fundamental region for $\hat{\Gamma}$, λ_i and λ_j' are not congruent to λ_j' if $i \neq j$. Further, the method of construction of $\mathbb{F}$ shows that the number of sides λ_j' not contained in $\mathbb{F}$ cannot exceed s. This proves parts (i), (ii) and (iii), and (iv) is obvious by the conformal property of bilinear mappings.

Finally, if $\mathbb{F}, S_1\mathbb{F}, S_2\mathbb{F}, \ldots, S_n\mathbb{F} (S_i \in \hat{\Gamma})$ is any sequence of images of $\mathbb{F}$, each adjacent to its successor, it follows from (iii) that S_n belongs to the group generated by $L_1, L_2, \ldots, L_s$. Also the set of all points z in $\mathbb{H}$ belonging to regions $S_n\mathbb{F}$ that can be reached by such sequences is open, and so also is its complement in $\mathbb{H}$, which must therefore be empty. This completes the proof of the theorem.

The last part of the theorem shows that, if we can construct a fundamental region for a group $\hat{\Gamma}$, we can find mappings in $\hat{\Gamma}$ that generate $\hat{\Gamma}$. For this purpose it is particularly convenient to choose fundamental regions with as few sides as possible, such as are constructed in parts (ii) and (iii) of theorem 2.4.3, since then the number of generators is small. Also we may, if we wish, include λ_j' in $\mathbb{F}$ instead of λ_j, for any value of j. Further, if λ_i and λ_j $(i \neq j)$ are consecutive sides on the same arc or straight line segment, and so are λ_i' and λ_j', we may count each of $\lambda_i \cup \lambda_j$, $\lambda_i' \cup \lambda_j'$ as a single side. We note also that, if $\mathbb{F}$ is a fundamental region for $\hat{\Gamma}$, we may obtain a proper fundamental region for $\hat{\Gamma}$ by omitting a finite number of points of $\mathbb{E} \cup \mathbb{P}$ from $\mathbb{F}$.

If, in theorem 2.4.4, $\mathscr{R}$ is chosen so that no two of the fundamental regions $\mathbb{F}_T$ are adjacent, then the number of pairs of congruent sides will be 2μ, since each $\mathbb{F}_T$ has four sides. We note, however, that for every choice of $\mathscr{R}$ the arguments used in the proof yield the following:

Corollary 2.4.4. *Under the assumptions of theorem 2.4.4, the 4μ sides of the regions $\mathbb{F}_T$ ($T \in \mathscr{R}$) can be grouped in two families $\mathbb{L}_U$ and $\mathbb{L}_V$ of pairs of sides (λ, λ'), where $\lambda' = S\lambda$ for some $S \in \hat{\Gamma}$. Each family contains μ pairs and, for $R = U$ or V,*

$$\mathbb{L}_R = \{(\lambda, \lambda') : \lambda = Tl_R, \lambda' = T'Rl_R;\ T, T' \in \mathscr{R},\ T'RT^{-1} \in \Gamma\}. \tag{2.4.13}$$

Theorems 2.4.3, 4 enable us to construct fundamental regions for groups and to find their generators. We give several examples of this in table 3, but first obtain some general results on normal subgroups of $\hat{\Gamma}(1)$. If $\hat{\Gamma}$ is a normal subgroup of $\hat{\Gamma}(1)$ with finite index μ, then the order $n(z, \Gamma)$ of each point $z \in \mathbb{P}$ (mod Γ) is the same; in fact, for each $z \in \mathbb{P}$, $n(z, \Gamma) = n_\infty$, where n_∞ is the smallest positive integer r for which $U^r \in \hat{\Gamma}$; i.e. $n_\infty = \operatorname{lev} \hat{\Gamma}$. In exactly the same way

$$n(z, \Gamma) = n_m \quad \text{for each } z \in \mathbb{E}_m \quad (m = 2, 3),$$

where $n_m = 1$ or m; n_2 and n_3 are respectively the smallest positive integers r for which V^r and $P^r \in \hat{\Gamma}$. We can therefore classify normal subgroups $\hat{\Gamma}$ of $\hat{\Gamma}(1)$ according to their type or *branch schema* $\{n_2, n_3, n_\infty\}$. We have:

Theorem 2.4.5. *If $\hat{\Gamma}$ is a normal subgroup of finite index μ in $\hat{\Gamma}(1)$ and is of branch schema $\{n_2, n_3, n_\infty\}$, then*

$$\mu = n_2\varepsilon_2 = n_3\varepsilon_3 = n_\infty\lambda, \tag{2.4.14}$$

where ε_2, ε_3 and λ are the number of orbits into which $\mathbb{E}_2$, $\mathbb{E}_3$ and $\mathbb{P}$ split (mod Γ), *respectively.*

Further, only the following four branch schemata occur:

(i) $\{1, 1, 1\}$, (ii) $\{2, 1, 2\}$, (iii) $\{1, 3, 3\}$, (iv) $\{2, 3, n\}$.

Only one group exists of each of the first three schemata, namely $\hat{\Gamma}(1)$, $\hat{\Gamma}^2$, and $\hat{\Gamma}^3$. In case (iv) $\mu \equiv 0$ (mod 6). (*Cf. theorem* 1.7.5.)

Proof. The equation (2.4.14) follows from (2.4.10, 11) and shows that, if $\hat{\Gamma}$ has branch schema $\{2, 3, n\}$, then $\mu \equiv 0$ (mod 6).

We prove that if a group $\hat{\Gamma}$ has schema $\{1, 3, n\}$, then $n = 3$. In the first place, by (2.4.14),

$$\mu = n\lambda = \varepsilon_2 = 3\varepsilon_3,$$

so that $\mu \equiv 0$ (mod 3). Let the fundamental region $\mathbb{F}$ of $\hat{\Gamma}$, if it exists, be formed by adjoining images of $\hat{\mathbb{F}}_I$. Since I, P, P^2 belong to different right cosets of $\hat{\Gamma}$ in $\hat{\Gamma}(1)$, we may assume that the region

$$\mathbb{F}_3 = \hat{\mathbb{F}}_I \cup \hat{\mathbb{F}}_P \cup \hat{\mathbb{F}}_{P^2} \tag{2.4.15}$$

is a subset of $\mathbb{F}$. Each of its six sides has one endpoint at a point of $\mathbb{E}_2$, and such vertices cannot be interior points of $\mathbb{F}$, since they are fixed points for $\hat{\Gamma}$; further only two triangles in $\mathbb{F}$ can contain such a vertex. It follows that $\mathbb{F} = \mathbb{F}_3$ and that $\mu = 3$.

The transformations L mapping congruent sides of $\mathbb{F}_3$ into each other cannot map any one of these three elliptic fixed points i, $\frac{1}{2}(i-1)$, $i-1$ into another, so that the six sides fall into three pairs $\lambda_\nu, \lambda'_\nu$ ($\nu = 1, 2, 3$) with $\lambda'_\nu = L_\nu\lambda_\nu$, $L_\nu \in \hat{\Gamma}$ ($\nu = 1, 2, 3$), where L_1, L_2 and L_3 have i, $\frac{1}{2}(i-1)$ and $i-1$ as fixed points, respectively. We take

$$L_1 := V, \quad L_2 := PVP^{-1} = V_2, \quad L_2 := P^2VP^{-2} = V_1 \tag{2.4.16}$$

and it follows from (1.3.9) that these three mappings do in fact generate the normal subgroup $\hat{\Gamma}^3$ of index 3 in $\hat{\Gamma}(1)$; further I, P, P^2 is a right transversal of $\hat{\Gamma}^3$ in $\hat{\Gamma}(1)$.

It can be shown similarly that, when $n_3 = 1$, the only groups are $\hat{\Gamma}(1)$ and $\hat{\Gamma}^2$. We already know that the latter is a normal subgroup of $\hat{\Gamma}(1)$ of index 2, and arguments of a similar nature to those given for $\hat{\Gamma}^3$ show that it has $\mathbb{F}_I \cup \mathbb{F}_V$ as a fundamental region and is generated by P and $P_1 := V^{-1}PV$. This completes the proof of the theorem.

In fig. 5 the fundamental regions of some of the groups that we have discussed are shown. We note that, since $\{I, P, P^2\}$ is a right transversal for the groups $\hat{\Gamma}_U(2)$, $\hat{\Gamma}_V(2)$, $\hat{\Gamma}_W(2)$ and $\hat{\Gamma}_3$, the region $\mathbb{F}_3$, defined by (2.4.15), is a fundamental region for each of these four groups. The transformations L_i (see theorem 2.4.4) that map a side λ_i into a corresponding side λ'_i are, however, different in each

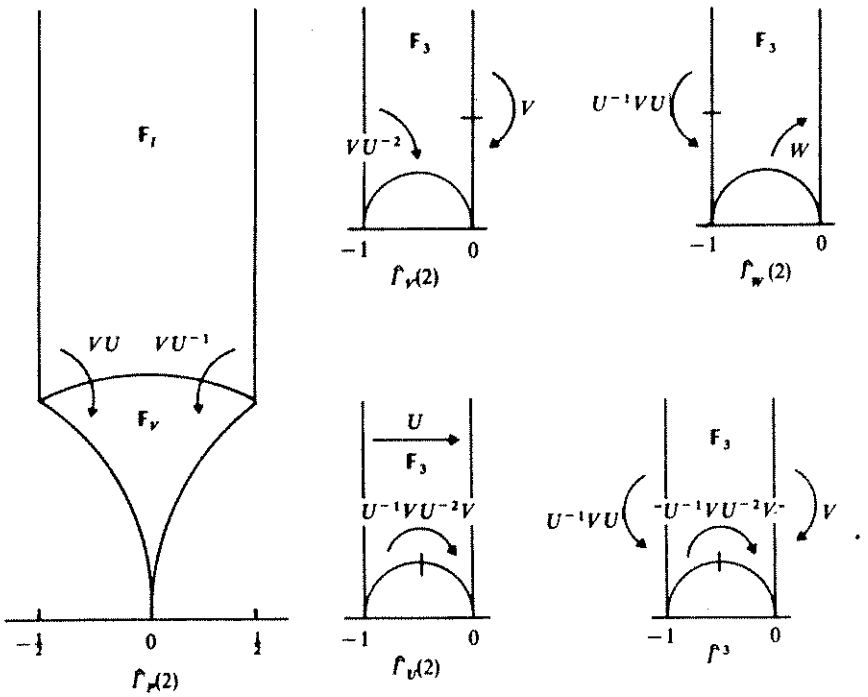


Fig. 5

of the four cases. The generators of the four groups, as given in the penultimate column of table 3, are derived from the mappings L_j illustrated in fig. 5. Theorem 2.4.4 can be applied in a similar way to find the generators of other subgroups of $\hat{\Gamma}(1)$ such as $\hat{\Gamma}(2)$ and $\hat{\Gamma}^*(1)$. Note that $\hat{\Gamma}(2)$ has schema $\{2, 3, 2\}$ and that both $\hat{\Gamma}(6)$ and $\hat{\Gamma}^*(1)$ have schema $\{2, 3, 6\}$.

Table 4 contains a variety of information about the nine normal subgroups of $\hat{\Gamma}(1)$ lying between $\hat{\Gamma}(6)$ and $\hat{\Gamma}(1)$. The rank of each subgroup is the minimum number of generators. The relations between the groups are illustrated in fig. 3.

We conclude by observing that, for $n > 1$, the principle congruence group $\hat{\Gamma}(n)$ has schema $\{2, 3, n\}$, and that the number $\hat{\lambda}(n)$ of incongruent cusps is

$$\hat{\lambda}(n) = \frac{\hat{\mu}(n)}{n} = \begin{cases} 3 & (n = 2), \\ \frac{1}{2}n^2 \prod_{p|n} \left(1 - \frac{1}{p^2}\right) & (n > 2). \end{cases} \tag{2.4.17}$$

Table 3. Groups of small index

$\hat{\Gamma}$	μ	F	Vertex sets in E_2	Order	Vertex sets in E_3	Order	Vertex sets in P	Order	Generators	Alternative generators
$\hat{\Gamma}(1)$	1	F_1	i	1	$\rho, -\rho^2$	1	∞	1	U, V	V, P
$\hat{\Gamma}^2_\rho(2)$	2	$F_1 \cup F_V$	—	—	ρ $-\rho^2$	1 1	$\infty, 0$	2	VU, VU^{-1}	$P, V^{-1}PV$
$\hat{\Gamma}_U(2)$	3	F_3	$\frac{1}{2}(i-1)$ $i, i-1$	1 2	—	—	∞ $0, -1$	1 2	$U, U^{-1}VU^{-2}V$	U, PVP^{-1}
$\hat{\Gamma}_V(2)$	3	F_3	i $i-1, \frac{1}{2}(i-1)$	1 2	—	—	-1 $0, \infty$	1 2	V, VU^2	V, U^2
$\hat{\Gamma}_W(2)$	3	F_3	$i-1$ $i, \frac{1}{2}(i-1)$	1 2	—	—	0 $-1, \infty$	1 2	$W, U^{-1}VU$	W, U^2
$\hat{\Gamma}^3$	3	F_3	i $i-1$ $\frac{1}{2}(i-1)$	1 1 1	—	—	$0, -1, \infty$	3	$V, U^{-1}VU,$ $U^{-1}VU^{-2}V$	$V, PVP^{-1}, P^{-1}VP$
$\hat{\Gamma}(2)$	6	$F_3 \cup UF_3$	—	—	—	—	0 ∞	2 2	U^2, VU^2V	
$\hat{\Gamma}(1)$	6	$F_3 \cup UF_3$	—	—	—	—	$1, -1$ $0, \infty, 1, -1$	2 6	UW, WU	

Table 4. Normal subgroups between $\hat{\Gamma}(1)$ and $\hat{\Gamma}(6)$

$\hat{\Gamma}$	μ	Rank	n_∞	λ	Generators	Congruence relations for $T \in \Gamma$: $\pm T \equiv$
$\hat{\Gamma}(1)$	1	2	1	1	$V=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$, $P=\begin{bmatrix}0 & -1\\ 1 & 1\end{bmatrix}$	
$\hat{\Gamma}^2$	2	2	2	1	$P=\begin{bmatrix}0 & -1\\ 1 & 1\end{bmatrix}$, $V^{-1}PV=\begin{bmatrix}1 & -1\\ 1 & 0\end{bmatrix}=P_1$	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0 & -1\\ 1 & 1\end{bmatrix}, \begin{bmatrix}-1 & -1\\ 1 & 0\end{bmatrix}=P^2 \pmod 2$. See also (1.7.1)
$\hat{\Gamma}^3$	3	3	3	1	$V=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$, $P^{-1}VP=\begin{bmatrix}-1 & -2\\ 1 & 1\end{bmatrix}$, $P^{-2}VP=\begin{bmatrix}-1 & -1\\ 2 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}, \begin{bmatrix}-1 & -1\\ 2 & 1\end{bmatrix}, \begin{bmatrix}1 & 2\\ -1 & -1\end{bmatrix} \pmod 3$. See also (1.7.2)
$\hat{\Gamma}'(1)$	6	2	6	1	$UW=[V,P]=\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}$, $WU=[V,P^2]=\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}$	See (1.7.3)
$\hat{\Gamma}(2)$	6	2	2	3	$U^2=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}$, $V^{-1}U^{-2}V=\begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix} \pmod 2$
$\hat{\Gamma}(3)$	12	3	3	4	$U^3=\begin{bmatrix}1 & 3\\ 0 & 1\end{bmatrix}$, $P^{-1}U^3P=\begin{bmatrix}4 & 3\\ -3 & -2\end{bmatrix}$, $P^{-2}U^3P^2=\begin{bmatrix}1 & 0\\ -3 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \pmod 3$

$\hat{\Gamma}^{2'}$	18	4	6	3	$[P, VPV]=\begin{bmatrix}1 & -2\\ -2 & 5\end{bmatrix}$, $[P, VP^2V]=\begin{bmatrix}3 & -2\\ -4 & 3\end{bmatrix}$, $[P^2, VPV]=\begin{bmatrix}3 & -4\\ -2 & 3\end{bmatrix}$, $[P^2, VP^2V]=\begin{bmatrix}5 & -2\\ -2 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & -2\\ -2 & 5\end{bmatrix}, \begin{bmatrix}5 & -2\\ -2 & 1\end{bmatrix}, \begin{bmatrix}3 & -2\\ -4 & 3\end{bmatrix} \pmod 6$
$\hat{\Gamma}^{3'}$	24	5	6	4	$(VP^2VP)^2=\begin{bmatrix}2 & 3\\ 3 & 5\end{bmatrix}$, $PVP^2(VP^2VP)^2PVP^2=\begin{bmatrix}-11 & -3\\ 15 & 4\end{bmatrix}$, $(VPVP^2)^2=\begin{bmatrix}5 & 3\\ 3 & 2\end{bmatrix}$, $V(P^2VP^2VP^2)^2V=\begin{bmatrix}1 & -3\\ 3 & -8\end{bmatrix}$, $(P^2VP^2VP^2)^2=\begin{bmatrix}8 & 3\\ -3 & -1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}2 & 3\\ 3 & 5\end{bmatrix}, \begin{bmatrix}5 & 3\\ 3 & 2\end{bmatrix} \pmod 6$
$\hat{\Gamma}(6)$	72	13	6	12		$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \pmod 6$

2.5. Further results. The method that we have used for obtaining fundamental regions for $\hat{\Gamma}(1)$ and its subgroups is convenient for the purposes we have in mind. However, if we had been considering the more general case of an arbitrary discontinuous subgroup $\hat{\Gamma}$ of $\hat{\Omega}$, we should probably have found it advantageous to introduce hyperbolic geometry and to adopt a slightly different definition of a fundamental region. A simple account of this more general theory can be found, for example, in the book by Lehner (1966), or the more advanced treatises by Fricke and Klein (1926) and Lehner (1964) may be consulted. We content ourselves here by sketching briefly how this general theory could be applied to a subgroup $\hat{\Gamma}$ of finite index μ in $\hat{\Gamma}(1)$; in this sketch we shall introduce a number of concepts that we do not define.

The *hyperbolic length* ds of an element of arc in $\mathbb{H}$ is defined by

$$(ds)^2 := y^{-2}\{(dx)^2+(dy)^2\}, \tag{2.5.1}$$

so that the hyperbolic length of a piecewise differentiable curve C in $\mathbb{H}$ is

$$L(C) := \int_C ds = \int_C \left\{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2\right\}^{\frac{1}{2}}\frac{dt}{y}. \tag{2.5.2}$$

In a similar way, the *hyperbolic area* $A(E)$ of a measurable subset E of $\mathbb{H}$ is defined to be

$$A(E) := \iint_E y^{-2}\,dx\,dy; \tag{2.5.3}$$

this can be infinite even when the ordinary Euclidean area of E is finite.

A curve in $\mathbb{H}$ is called a *hyperbolic straight line* if it is either a semicircle in $\mathbb{H}$ centred at a point of $\mathbb{R}$, or is an ordinary straight line in $\mathbb{H}$ that is orthogonal to $\mathbb{R}$. If z_1 and z_2 are distinct points of $\mathbb{H}$, there is a unique hyperbolic straight line passing through z_1 and z_2 and the arc of this hyperbolic straight line joining z_1 and z_2 is the curve C of smallest hyperbolic length joining z_1 to z_2. The *hyperbolic distance* $d(z_1, z_2)$ of z_1 from z_2 is defined to be $L(C)$; if $z_1 = z_2$, we put $d(z_1, z_2) = 0$.

It can be shown that $d(z_1, z_2)$ defines a metric on $\mathbb{H}$, and the corresponding metric topology on $\mathbb{H}$ is identical with the natural topology on $\mathbb{H}$. It is convenient to extend this topology from $\mathbb{H}$ to $\mathbb{H}'$

by giving each point of $\mathbb{P}$ a suitable base of neighbourhoods; see Lehner (1964), chapter 4.

The hyperbolic metric has the useful property that $L(TC) = L(C)$ and $A(TE) = A(E)$ for every $T \in \hat{\Omega}$. Thus every fundamental region for $\hat{\Gamma}$ has the same hyperbolic area. Further, this hyperbolic area is easy to evaluate by using the Gauss–Bonnet formula, which states that the area of a triangle bounded by hyperbolic line segments is $\pi-(\alpha+\beta+\gamma)$, where α, β and γ are the interior angles between the sides. This gives, for example, for a fundamental region $\mathbb{F}$ for $\hat{\Gamma}$

$$A(\mathbb{F}) = \mu A(\mathbb{F}_I) = \tfrac{1}{3}\mu\pi, \tag{2.5.4}$$

since $\mathbb{F}_I$ has interior angles 0, $\frac{1}{3}\pi$ and $\frac{1}{3}\pi$.

Different definitions of a fundamental region are given by different authors. What we have called a proper fundamental region is called a fundamental set by Lehner (1966) and Schoeneberg (1974). In his two books Lehner defines a subset $\mathbb{D}$ of $\mathbb{H}$ to be a fundamental region for $\hat{\Gamma}$ if (i) $\mathbb{D}$ is open in $\mathbb{H}$, (ii) no two distinct points of $\mathbb{D}$ are congruent modulo Γ and (iii) every point of $\mathbb{H}$ is congruent to a point of the closure $\bar{\mathbb{D}}$ of $\mathbb{D}$ in $\mathbb{H}$. On the other hand, Macbeath (1961) takes $\mathbb{D}$ to be closed and modifies (ii) accordingly. The interior of either of the fundamental regions $\mathbb{F}$ or $\hat{\mathbb{F}}$ of theorem 2.4.3 is a fundamental region according to Lehner.

Let $z_0 \in \mathbb{H}-\mathbb{E}(\Gamma)$ and define $\mathbb{D} = \mathbb{D}(z_0)$ to be the set of all points $z \in \mathbb{H}$ such that

$$d(z, z_0) < d(z, Tz_0)$$

for all $T \in \hat{\Gamma}$ except $T = \pm I$. It can be shown that $\mathbb{D}$ is a fundamental region in the sense of Lehner and that it is bounded by a finite number of segments of hyperbolic straight lines. This fundamental region is called a *normal polygon* or a *Dirichlet region*.

For theoretical purposes the normal polygon has many advantages as a fundamental region, but in individual cases it is not so easy to construct as the fundamental regions set up in theorem 2.4.3. An alternative method due to Ford (1929), which is also easily applicable in particular cases, defines $\mathbb{D}$ to be the region of $\mathbb{H}$ contained in $\mathbb{S}_{n_I}$ and outside all the *isometric circles* $|T:z| = 1$ $(T \in \Gamma, c \neq 0)$; see also Rankin (1954). This again yields a region bounded by segments of hyperbolic straight lines. When applied to $\hat{\Gamma}(1)$ the isometric circle method gives the interior of $\mathbb{F}_I$.

Each subgroup $\hat{\Gamma}$ has associated with it a Riemann surface, which is obtained in the following way. The set of orbits $\hat{\Gamma}z$ $(z \in \mathbb{H}')$ is denoted by $\mathbb{H}'/\Gamma$ and, when given the identification topology induced by congruence modulo Γ, becomes a connected Hausdorff space, which is, in fact, a Riemann surface. Its points are in one-to-one correspondence with those of a proper fundamental region $\mathbb{F}$ for $\hat{\Gamma}$. It is convenient to regard $\mathbb{H}'/\Gamma$ as being constructed from $\mathbb{F}$ by identifying pairs of congruent sides λ_i, λ_i' (see theorem 2.4.4). The Riemann surface $\mathbb{H}'/\Gamma$ has $\mathbb{H}'$ as a branched covering surface. At points of $\mathbb{H}-\mathbb{E}(\Gamma)$ the covering is unbranched; at a point of $\mathbb{E}_k(\Gamma)$ there is a branchpoint of order k; at points of $\mathbb{P}$ there are logarithmic winding points.

The surface $\mathbb{H}'/\Gamma$ is compact and has finite genus $g = g(\hat{\Gamma})$. The images of the triangles $\mathbb{F}^{(1)}$ and $\mathbb{F}^{(2)}$ (see (2.4.2, 3)) provide a natural triangulation of $\mathbb{H}'/\Gamma$ from which it is possible to show that

$$g = 1 + \tfrac{1}{2}(\mu - \lambda - \varepsilon_2 - \varepsilon_3). \tag{2.5.5}$$

Here, as in §2.4, λ is the number of incongruent cusps (mod Γ) and ε_k is the number of incongruent points of $\mathbb{E}_k$ (mod Γ) $(k = 2, 3)$; see, for example, Gunning (1962), §4, theorem 5 or Schoeneberg (1974), chapter 4, §7. It follows from (2.4.12) and (2.5.5) that we also have

$$g = 1 + \frac{1}{2}\left(\frac{\mu}{6} - \lambda - \frac{e_2}{2} - \frac{2e_3}{3}\right). \tag{2.5.6}$$

In particular, when $\hat{\Gamma}$ is a normal subgroup of $\hat{\Gamma}(1)$ with branch schema $\{n_2, n_2, n_\infty\}$, then, by (2.4.14) and (2.5.5),

$$g = 1 + \frac{1}{2}\mu\left(1 - \frac{1}{n_2} - \frac{1}{n_3} - \frac{1}{n_\infty}\right). \tag{2.5.7}$$

A knowledge of the genus of $\hat{\Gamma}$ is useful when applications of the Riemann–Roch theorem are made to find the number of linearly independent modular forms of different kinds for $\hat{\Gamma}$. However, it is possible in many cases to obtain exact results without using deep theorems of this kind, as we shall see. This happens, in particular, when the genus is zero. In this connexion we note that, by (2.5.7), $g = 0$ for all the groups in tables 3 and 4 except for $\hat{\Gamma}(1)$ and $\hat{\Gamma}(6)$, which both have genus 1.

Since, by theorem 2.4.5, $\hat{\Gamma}(n)$ has branch schema $\{2, 3, n\}$ for $n \geq 2$, it follows from (2.5.7) and theorem 1.4.1 that

$$g\{\Gamma(n)\} = 1 + \frac{n^2(n-6)}{24}\prod_{p|n}\left(1 - \frac{1}{p^2}\right) \quad (n \geq 3), \tag{2.5.8}$$

while $g\{\hat{\Gamma}(2)\} = 0$. It follows that the genus of $\hat{\Gamma}(n)$ is zero for $n \leq 5$.

We conclude by remarking that it can be shown that a *canonical fundamental region* for $\hat{\Gamma}$ can be constructed having $2n + 4g$ sides that are segments of hyperbolic straight lines. Here $n = \lambda + e_2 + e_3$ and the sides follow each other in the order

$$\lambda_1\lambda_1'\lambda_2\lambda_2' \ldots \lambda_n\lambda_n'\mu_1\nu_1\mu_1'\nu_1'\mu_2\nu_2\mu_2'\nu_2' \ldots \mu_g\nu_g\mu_g'\nu_g'.$$

Here

$$\lambda_i' = L_i\lambda_i, \quad \mu_j' = M_j\mu_j, \quad \nu_j' = N_j\nu_j \quad (1 \leq i \leq n,\ 1 \leq j \leq g)$$

where the mappings L_i, M_j, N_j belong to $\hat{\Gamma}$. Further the first e_2 of the L_i are elliptic transformations of order 2, the next e_3 are elliptic transformations of order 3 and the last λ are parabolic transformations; see Lehner (1964), chapter 7. The group $\hat{\Gamma}$ is generated by the $n + 2g$ mappings L_i, M_j, N_j $(1 \leq i \leq n,\ 1 \leq j \leq g)$. Each elliptic generator L_i satisfies a relation of the form $L_i^k = I$ $(k = 2$ or $3)$; apart from these relations we also have

$$L_1L_2 \ldots L_nM_1N_1M_1^{-1}N_1^{-1}M_2N_2M_2^{-1}N_2^{-1} \ldots M_gN_gM_g^{-1}N_g^{-1} = I, \tag{2.5.9}$$

and all other relations between the elements of $\hat{\Gamma}$ are consequences of the relations given.

Hsu, Identifying Congruence Subgroups of the Modular Group, Proc. AMS 124 (1996), 1351–1359

PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 124, Number 5, May 1996

IDENTIFYING CONGRUENCE SUBGROUPS OF THE MODULAR GROUP

TIM HSU

(Communicated by Ronald M. Solomon)

ABSTRACT. We exhibit a simple test (Theorem 2.4) for determining if a given (classical) modular subgroup is a congruence subgroup, and give a detailed description of its implementation (Theorem 3.1). In an appendix, we also describe a more "invariant" and arithmetic congruence test.

1. NOTATION

We describe (conjugacy classes of) subgroups $\Gamma \subset \mathbf{PSL}_2(\mathbf{Z})$ in terms of permutation representations of $\mathbf{PSL}_2(\mathbf{Z})$, following Millington [11, 12] and Atkin and Swinnerton-Dyer [1].

We recall that a conjugacy class of subgroups of $\mathbf{PSL}_2(\mathbf{Z})$ is equivalent to a transitive permutation represention of $\mathbf{PSL}_2(\mathbf{Z})$. Such a representation can be defined by transitive permutations E and V which satisfy the relations

$$1 = E^2 = V^3. \tag{1.1}$$

The relations (1.1) are fulfilled by

$$E = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}. \tag{1.2}$$

Alternately, such a representation can be defined by transitive permutations L and R which satisfy

$$1 = (LR^{-1}L)^2 = (R^{-1}L)^3, \tag{1.3}$$

with the relations being fulfilled by

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{1.4}$$

One can also use permuations E and L such that

$$1 = E^2 = (L^{-1}E)^3, \tag{1.5}$$

with E and L corresponding to the indicated matrices in (1.2) and (1.4), respectively.

Received by the editors September 1, 1994.
1991 *Mathematics Subject Classification.* Primary 20H05; Secondary 20F05.
Key words and phrases. Congruence subgroups, classical modular group.
The author was supported by an NSF graduate fellowship and DOE GAANN grant #P200A10022.A03.

The various notations can be translated using the following conversion table:

$$E = LR^{-1}L, \qquad V = R^{-1}L, \tag{1.6}$$

$$L = EV^{-1}, \qquad R = EV^{-2}, \tag{1.7}$$

$$R = E^{-1}L^{-1}E. \tag{1.8}$$

Example 1.1. The permutations

$$\begin{aligned} E &= (1\ 2)(3\ 4)(5\ 6)(7\ 8)(9\ 10), \\ V &= (1\ 3\ 5)(2\ 7\ 4)(6\ 8\ 9), \end{aligned} \tag{1.9}$$

or, alternately,

$$\begin{aligned} L &= (1\ 4)(2\ 5\ 9\ 10\ 8)(3\ 7\ 6), \\ R &= (1\ 7\ 9\ 10\ 6)(2\ 3)(4\ 5\ 8), \end{aligned} \tag{1.10}$$

describe a conjugacy class of subgroups of index 10 in $\mathbf{PSL}_2(\mathbf{Z})$.

Remark 1.2. Note that any concrete method of specifying a modular subgroup can easily be converted to permutation form. For instance, one way in which a modular subgroup Γ might be specified is by a list of generators. Such a list can be converted into permutations as follows: First, use the Euclidean algorithm to express each generator matrix as a product of L's and R's, where L and R are the elements in (1.4). Then enumerate the cosets of Γ in terms of these generators and presentation (1.3). This coset enumeration is easily converted into appropriate permutations L and R. Similarly, any reasonable membership test for Γ can be used to enumerate the cosets of Γ, with the same results as before.

2. Congruence subgroups and the level

We recall the following definitions.

Definition 2.1. $\Gamma(N)$ is defined to be the group

$$\{\gamma \in \mathbf{PSL}_2(\mathbf{Z}) \mid \gamma \equiv \pm I \pmod{N}\}. \tag{2.1}$$

$\Gamma(N)$ is the kernel of the natural projection from $\mathbf{PSL}_2(\mathbf{Z})$ to $\mathbf{SL}_2(\mathbf{Z}/N)/\{\pm I\}$. We say that a modular subgroup Γ is a *congruence subgroup* if Γ contains $\Gamma(N)$ for some integer N. Otherwise, we say Γ is a *non-congruence subgroup.*

An important invariant of (conjugacy classes of) modular subgroups is the following.

Definition 2.2. The *level* of a modular subgroup Γ, as specified by permutations L and R, is defined to be the order of L (or the order of R, since L is conjugate to R^{-1}).

We need the following result, sometimes known as Wohlfahrt's Theorem (Wohlfahrt [13]).

Theorem 2.3. *Let N be the level of a modular subgroup Γ. Γ is a congruence subgroup if and only if it contains $\Gamma(N)$.*

Proof. This amounts to proving that, for congruence subgroups, our definition of the level is the same as the classical definition of the level. See Wohlfahrt [13]. □

Theorem 2.4. *Let Γ be a modular subgroup of level N, and let*

$$\langle L,\, R \,|r_1,\, r_2,\, \ldots\rangle \tag{2.2}$$

be a presentation for $\mathbf{SL}_2(\mathbf{Z}/N)/\{\pm I\}$ which is compatible with (1.4). *Then Γ is a congruence subgroup if and only if the representation of $\mathbf{PSL}_2(\mathbf{Z})$ induced by Γ respects the relations $\{r_i\}$.*

Proof. From Theorem 2.3, we only need to check if Γ contains $\Gamma(N)$. Now, since $\Gamma(N)$ is normal in $\mathbf{PSL}_2(\mathbf{Z})$, Γ contains $\Gamma(N)$ if and only if the normal kernel of Γ contains $\Gamma(N)$. However, the normal kernel of Γ is exactly the kernel of the representation induced by Γ, and since the relations $\{r_i\}$ generate $\Gamma(N)$ as their normal closure, the theorem follows. □

Compare Magnus [9, Ch. III], Britto [4], Wohlfahrt [13], and Larcher [8]. Lang, Lim, and Tan [7] have also developed a congruence test; see the related paper Chan, Lang, Lim, and Tan [5].

Example 2.5. Suppose Γ is the conjugacy class of subgroups specified by (1.10). Since L has order 30, we need to use a presentation for $\mathbf{SL}_2(\mathbf{Z}/30)/\{\pm I\}$. We find that $\mathbf{SL}_2(\mathbf{Z}/30)/\{\pm I\}$ has a presentation with defining relations

$$1 = L^{30}, \tag{2.3}$$

$$1 = [L^2, R^{15}] = [L^3, R^{10}] = [L^5, R^6] \tag{2.4}$$

in addition to the relations in (1.3). (The commutator $[x, y]$ is defined to be $x^{-1}y^{-1}xy$, so $1 = [x, y]$ means "x commutes with y".) Only the commutator relations (2.4) need to be checked. However,

$$L^2 = (2\ 9\ 8\ 5\ 10)(3\ 6\ 7), \tag{2.5}$$

which does not commute with

$$R^{15} = (2\ 3), \tag{2.6}$$

so Γ is a non-congruence subgroup. (It is worth mentioning that Larcher's results also imply that Γ is non-congruence, since L does not contain a 30-cycle.)

Remark 2.6. The results in this section extend essentially verbatim to the *Bianchi groups* $\mathbf{SL}_2(O_d)$, where O_d is the ring of algebraic integers of an imaginary quadratic field $\mathbf{Q}\big[\sqrt{-d}\,\big]$ with class number 1. (See Fine [6] for more on the Bianchi groups.) However, for practical use, one needs a uniform presentation of $\mathbf{SL}_2(O_d)\,/\mathfrak{A}$ for $\mathfrak{A}$ any ideal of O_d.

3. Implementation

To assure the reader that the procedure described by Theorem 2.4 is practical, we provide the following detailed algorithm. Suppose we are given a subgroup Γ of finite index in $\mathbf{PSL}_2(\mathbf{Z})$.

1. Describe Γ in terms of permutations L and R. If necessary, use conversion (1.7), conversion (1.8), or another similar conversion. (See also Remark 1.2.)

2. Let N be the order of L, and let $N = em$, where e is a power of 2 and m is odd.

3. We have three cases:

(a) *N is odd:* Γ is a congruence subgroup if and only if the relation

$$1 = (R^2 L^{-\frac{1}{2}})^3 \tag{A}$$

is satisfied, where $\frac{1}{2}$ is the multiplicative inverse of 2 mod N.

(b) *N is a power of 2*: Let $S = L^{20}R^{\frac{1}{5}}L^{-4}R^{-1}$, where $\frac{1}{5}$ is the multiplicative inverse of 5 mod N. Γ is a congruence subgroup if and only if the relations

$$\begin{gathered}(LR^{-1}L)^{-1}S(LR^{-1}L) = S^{-1},\\ S^{-1}RS = R^{25},\\ 1 = (SR^5LR^{-1}L)^3\end{gathered} \tag{B}$$

are satisfied.

(c) *Both e and m are greater than* 1:

(i) Let $\frac{1}{2}$ be the multiplicative inverse of 2 mod m, and let $\frac{1}{5}$ be the multiplicative inverse of 5 mod e.

(ii) Let c be the unique integer mod N such that $c \equiv 0 \pmod{e}$ and $c \equiv 1 \pmod{m}$, and let d be the unique integer mod N such that $d \equiv 0 \pmod{m}$ and $d \equiv 1 \pmod{e}$.

(iii) Let $a = L^c$, $b = R^c$, $l = L^d$, $r = R^d$, and let $s = l^{20}r^{\frac{1}{5}}l^{-4}r^{-1}$.

(iv) Γ is a congruence subgroup if and only if the relations

$$\begin{gathered}1 = [a, r],\\ 1 = (ab^{-1}a)^4,\\ (ab^{-1}a)^2 = (b^{-1}a)^3,\\ (ab^{-1}a)^2 = (b^2a^{-\frac{1}{2}})^3,\\ (lr^{-1}l)^{-1}s(lr^{-1}l) = s^{-1},\\ s^{-1}rs = r^{25},\\ (lr^{-1}l)^2 = (sr^5lr^{-1}l)^3\end{gathered} \tag{C}$$

are satisfied.

Theorem 3.1. *The above procedure determines if Γ is a congruence subgroup.*

Before proving Theorem 3.1, we need an algebraic trick (Lemma 3.2) and some known results (Lemma 3.3, due to Behr and Mennicke [2]; and Lemma 3.4, due to Mennicke [10]).

Lemma 3.2 (Braid trick)**.** *Let x and y be elements which generate a group G and satisfy the relation*

$$(xyx)^2 = (yx)^3. \tag{3.1}$$

Then the element $(xyx)^2 = (yx)^3$ is central in G. Furthermore,

$$xyx = yxy \tag{3.2}$$

and

$$(xyx)^{-1}x(xyx) = y. \tag{3.3}$$

We call this the "braid trick" because (3.2) is the defining relation for the 3-string braid group.

Proof. The elements $X = xyx$ and $Y = yx$ also generate G, and the element $Z = (xyx)^2 = (yx)^3 = X^2 = Y^3$ commutes with both X and Y, so Z is central. (3.2) and (3.3) follow from cancellation in $xyxxyx = yxyxyx$. □

Lemma 3.3. *Let m be an odd integer, and let $\frac{1}{2}$ be the multiplicative inverse of 2 mod m. $\mathbf{SL}_2(\mathbf{Z}/m)$ is isomorphic to*

$$G = \left\langle a, b \;\middle|\; \right.$$

$$1 = a^m, \tag{3.4}$$

$$1 = (ab^{-1}a)^4, \tag{3.5}$$

$$(ab^{-1}a)^2 = (b^{-1}a)^3, \tag{3.6}$$

$$(ab^{-1}a)^2 = (b^2a^{-\frac{1}{2}})^3 \left.\right\rangle. \tag{3.7}$$

Relations (3.4)–(3.7) *are fulfilled by* $a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ *and* $b = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ *in* $\mathbf{SL}_2(\mathbf{Z}/m)$.

Proof. G is equivalent to Behr and Mennicke's presentation [2, (2.12)] by the following Tietze transformations. Add generators $A = b$ and $B = ab^{-1}a$. Applying the braid trick to (3.6), we get that B^2 is central, and from (3.2), we also get that

$$BA = b^{-1}a. \tag{3.8}$$

(3.8) implies that $a = ABA$, which means that we can eliminate a and b.

Using (3.3), (3.8), and the centrality of B^2, we see that (3.4)–(3.6) become

$$1 = A^m = B^4, \tag{3.9}$$

$$B^2 = (AB)^3, \tag{3.10}$$

so it remains to convert (3.7) to Behr and Mennicke's form. However, applying (3.3), we have

$$B^2 = (b^2a^{-\frac{1}{2}})^3 = (A^2B^{-1}A^{\frac{1}{2}}B)^3, \tag{3.11}$$

so, using $1 = B^8$ and the centrality of B^2,

$$1 = (A^2B^{-1}A^{\frac{1}{2}}B)^3B^6 = (A^2BA^{\frac{1}{2}}B)^3. \quad \square \tag{3.12}$$

Lemma 3.4. *Let $e = 2^n$, let $\frac{1}{5}$ be the multiplicative inverse of 5 mod e, and let $s = l^{20}r^{\frac{1}{5}}l^{-4}r^{-1}$. $\mathbf{SL}_2(\mathbf{Z}/e)$ is isomorphic to*

$$G = \left\langle l, r \;\middle|\; \right.$$

$$1 = l^e, \tag{3.13}$$

$$1 = (lr^{-1}l)^4, \tag{3.14}$$

$$(lr^{-1}l)^2 = (r^{-1}l)^3, \tag{3.15}$$

$$(lr^{-1}l)^{-1}s(lr^{-1}l) = s^{-1}, \tag{3.16}$$

$$s^{-1}rs = r^{25}, \tag{3.17}$$

$$(lr^{-1}l)^2 = (sr^5lr^{-1}l)^3 \left.\right\rangle. \tag{3.18}$$

Relations (3.13)–(3.18) *are fulfilled by* $l = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $r = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, *and* $s = \begin{pmatrix} 5 & 0 \\ 0 & \frac{1}{5} \end{pmatrix}$ *in* $\mathbf{SL}_2(\mathbf{Z}/e)$.

Proof. As the reader may verify, the relations (3.13)–(3.18) and $s = l^{20}r^{\frac{1}{5}}l^{-4}r^{-1}$ are satisfied in $\mathbf{SL}_2(\mathbf{Z}/e)$, so it suffices to show that G is a homomorphic image of Mennicke's presentation [10, p. 210]. Add generators $A = r$, $B = lr^{-1}l$, and $T = s$. Applying the braid trick to (3.15), we get that B^2 is central, $BA = r^{-1}l$, and l is conjugate to A^{-1}. As in the proof of the previous lemma, we can then eliminate generators l and r. Then (3.13), (3.14), (3.15), (3.16), (3.17), and (3.18) become Mennicke's relations (X), (Y), (P), (Z), (Q), and (R), respectively. □

For Lemma 3.5, we consider the following relations:

$$1 = L^N, \tag{3.19}$$

$$1 = [a, r], \tag{3.20}$$

$$1 = [b, l], \tag{3.21}$$

$$1 = (ab^{-1}a)^4, \tag{3.22}$$

$$(ab^{-1}a)^2 = (b^{-1}a)^3, \tag{3.23}$$

$$(ab^{-1}a)^2 = (b^2a^{-\frac{1}{2}})^3, \tag{3.24}$$

$$1 = (lr^{-1}l)^4, \tag{3.25}$$

$$(lr^{-1}l)^2 = (r^{-1}l)^3, \tag{3.26}$$

$$(lr^{-1}l)^{-1}s(lr^{-1}l) = s^{-1}, \tag{3.27}$$

$$s^{-1}rs = r^{25}, \tag{3.28}$$

$$(lr^{-1}l)^2 = (sr^5lr^{-1}l)^3. \tag{3.29}$$

All notation is as described in (2) and (3c)(i–iii) of the algorithm. Note that $1 = L^N$ implies that $L = al$ and $R = br$.

Lemma 3.5. $\mathbf{SL}_2(\mathbf{Z}/N)$ *has a presentation with generators* L *and* R*, and defining relations* (3.19)–(3.29). *The relations are fulfilled by* $L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ *and* $R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ *in* $\mathbf{SL}_2(\mathbf{Z}/N)$.

Proof. The Chinese Remainder Theorem implies that

$$\mathbf{SL}_2(\mathbf{Z}/N) \cong \mathbf{SL}_2(\mathbf{Z}/m) \times \mathbf{SL}_2(\mathbf{Z}/e)\,. \tag{3.30}$$

It also follows from the Chinese Remainder Theorem that, if $L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ in $\mathbf{SL}_2(\mathbf{Z}/N)$, the $\mathbf{SL}_2(\mathbf{Z}/m)$ factor is precisely $\langle a, b\rangle$ and the $\mathbf{SL}_2(\mathbf{Z}/e)$ factor is precisely $\langle l, r\rangle$. Therefore, the above relations are satisfied in $\mathbf{SL}_2(\mathbf{Z}/N)$.

On the other hand, since (3.19) implies (3.4) and (3.13), comparison with Lemmas 3.3 and 3.4 shows that the above presentation is the direct product of $\mathbf{SL}_2(\mathbf{Z}/m)$ and $\mathbf{SL}_2(\mathbf{Z}/e)$. The lemma follows. □

Proof of Theorem 3.1. After steps 1 and 2 of the procedure, we know that the relations

$$1 = L^N, \tag{3.31}$$

$$1 = (LR^{-1}L)^2, \tag{3.32}$$

$$1 = (R^{-1}L)^3 \tag{3.33}$$

must be satisfied. From Theorem 2.4, we see that if (3.31)–(3.33) and (A) (resp. (B), (C)) are defining relations for $\mathbf{SL}_2(\mathbf{Z}/N)/\{\pm I\}$ when N is odd (resp. N is a power of 2, and e and m are greater than 1), then Theorem 3.1 follows. Comparing (A) and Lemma 3.3, with $a = L$ and $b = R$, and comparing (B) and Lemma 3.4, with $l = L$ and $r = R$, the first two cases follow easily, so it remains to check the third.

Comparing (C) and (3.19)–(3.29), we see that it is enough to show that given (3.31)–(3.33) and (3.19)–(3.29), the relations (3.21), (3.25), and (3.26) are redundant. First, (3.31), (3.32), (3.20), and (3.21) give us

$$\begin{aligned} 1 &= (LR^{-1}L)^4 \\ &= (alr^{-1}b^{-1}al)^4 \\ &= (ab^{-1}a)^4(lr^{-1}l)^4, \end{aligned} \tag{3.34}$$

which means that (3.22) implies (3.25). Similarly, (3.31), (3.32), (3.33), (3.20), and (3.21) imply

$$\begin{gathered} (LR^{-1}L)^2 = (R^{-1}L)^3, \\ (ab^{-1}a)^2(lr^{-1}l)^2 = (b^{-1}a)^3(r^{-1}l)^3, \end{gathered} \tag{3.35}$$

which means that (3.23) implies (3.26). Finally, since (3.32), (3.33), and the braid trick (3.3) imply that L is conjugate to R^{-1}, we can eliminate (3.21), since it is implied by (3.20). □

For hand calculations, and for further study, we note the following relations which occur in $\mathbf{SL}_2(\mathbf{Z}/N)$:

$$(SL_2) \qquad Z = (LR^{-1}L)^2 = (R^{-1}L)^3,\ 1 = Z^2,$$

$$\text{(level)} \qquad 1 = L^N = R^N,$$

$$(ab \equiv 0 \pmod N) \qquad 1 = [L^a, R^b],$$

$$(ab \equiv -1 \pmod N) \qquad (L^aR^b)^3 = Z,$$

$$(ab \equiv -2 \pmod N) \qquad (L^aR^b)^2 = Z.$$

It has been verified by coset enumeration that the relations (SL_2), (level), and $(ab \equiv 0 \pmod N)$ are defining relations when $N \mid 360$. This means that if the level N divides 360, the congruence test reduces to checking that the relations $(ab \equiv 0 \pmod N)$ are satisfied.

Acknowledgements

The author would like to thank J. H. Conway and the referee for many helpful comments and suggestions.

Appendix A. An arithmetic congruence test

In this appendix, we present an arithmetic and "invariant" congruence test which uses the Ihara modular group $\mathbf{SL}_2\left(\mathbf{Z}\left[\frac{1}{p}\right]\right)$.

We begin by quoting the following result (Theorem A.1) of J. Mennicke [10]. (Note that Mennicke's Schur multiplier calculation and subsequent argument require the repairs described in F.R. Beyl [3, §5], but the main result still holds.) Let N be an integer, let p be a prime not dividing N, let R_N be the kernel in $\mathbf{SL}_2\Big(\mathbf{Z}\Big[\frac{1}{p}\Big]\Big)$ resulting from reduction mod N, and let Q_N be the normal closure of L^N in $\mathbf{SL}_2\Big(\mathbf{Z}\Big[\frac{1}{p}\Big]\Big)$.

Theorem A.1. $R_N = Q_N$. □

Let Γ be a modular subgroup of level N and index m in $\mathbf{SL}_2(\mathbf{Z})$. Consider the commutative diagram in Figure A.1.

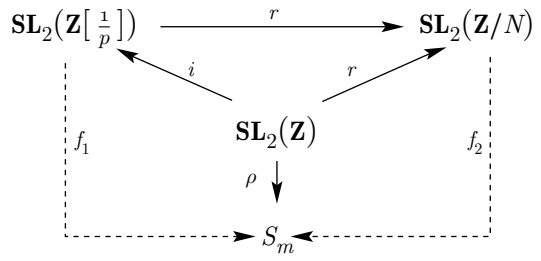


FIGURE A.1. Commutative diagram for Theorem A.2

Here, S_m is the symmetric group on m objects (the cosets of Γ in $\mathbf{SL}_2(\mathbf{Z})$), r is reduction mod N, i is inclusion, and ρ is the permutation representation of $\mathbf{SL}_2(\mathbf{Z})$ induced by Γ. Note that f_2 exists if and only if Γ is a congruence subgroup, and that such an f_2 is uniquely determined.

The setup in Figure A.1 provides us with an invariant congruence test.

Theorem A.2. *In the notation of Figure A.1, a map f_1 exists if and only if f_2 exists. In other words, Γ is congruence if and only if ρ can be factored through inclusion in $\mathbf{SL}_2\Big(\mathbf{Z}\Big[\frac{1}{p}\Big]\Big)$.*

Proof. If f_2 exists, let $f_1 = f_2 r$. Conversely, if f_1 exists, since L^N is in the kernel of ρ, L^N must be in the kernel of f_1, so in fact, f_1 is well defined on

$$\mathbf{SL}_2\Big(\mathbf{Z}\Big[\tfrac{1}{p}\Big]\Big)\Big/Q_N = \mathbf{SL}_2\Big(\mathbf{Z}\Big[\tfrac{1}{p}\Big]\Big)\Big/R_N \cong \mathbf{SL}_2(\mathbf{Z}/N)\,, \tag{A.1}$$

which means that f_1 defines an appropriate map f_2. □

Corollary A.3. *In Figure A.1, f_1 is determined uniquely if it exists.* □

One curious feature of Theorem A.2 is that if we know a given family of modular subgroups all have levels relatively prime to p, then we can handle all of them in a uniform manner. This is the principle behind Behr and Mennicke's presentation of $\mathbf{SL}_2(\mathbf{Z}/N)$ for N odd, as these cases can be handled in $\mathbf{SL}_2\big(\mathbf{Z}\big[\frac{1}{2}\big]\big)$.

We also note that if we fix the level N, then we can choose any p not dividing N to use in Theorem A.2. This leads to the following idea: For a given family of modular subgroups of level N, it seems plausible that one might be able to reduce the extensibility of ρ to the question of whether there exists a p which satisfies certain congruences mod N. Dirichlet's theorem might then be used to find a p which satisfies those congruences.

References

[1] A. O. L. Atkin and H. P. F. Swinnerton-Dyer, *Modular forms on noncongruence subgroups*, Proc. Symp. Pure Math., Combinatorics (Providence) (T. S. Motzkin, ed.), vol. 19, AMS, Providence, 1971, pp. 1–26. MR **49:**2550

[2] H. Behr and J. Mennicke, *A presentation of the groups* $PSL(2,p)$, Can. J. Math. **20** (1968), 1432–1438. MR **38:**4566

[3] F. R. Beyl, *The Schur multiplicator of* $SL(2,Z/mZ)$ *and the congruence subgroup property*, Math. Z. **191** (1986), 23–42. MR **87b:**20071

[4] J. Britto, *On the construction of non-congruence subgroups*, Acta Arith. **XXXIII** (1977), 261–267. MR **56:**12142

[5] S.-P. Chan, M.-L. Lang, C.-H. Lim, and S.-P. Tan, *Special polygons for subgroups of the modular group and applications*, Internat. J. Math. **4** (1993), no. 1, 11–34. MR **94j:**11045

[6] B. Fine, *Algebraic theory of the Bianchi groups*, Marcel Dekker, Inc., New York, 1989. MR **90h:**20002

[7] M.-L. Lang, C.-H. Lim, and S.-P. Tan, *An algorithm for determining if a subgroup of the modular group is congruence*, preprint, 1992.

[8] H. Larcher, *The cusp amplitudes of the congruence subgroups of the classical modular group*, Ill. J. Math. **26** (1982), no. 1, 164–172. MR **83a:**10040

[9] W. Magnus, *Noneuclidean tesselations and their groups*, Academic Press, 1974. MR **50:**4774

[10] J. Mennicke, *On Ihara's modular group*, Invent. Math. **4** (1967), 202–228. MR **37:**1485

[11] M. H. Millington, *On cycloidal subgroups of the modular group*, Proc. Lon. Math. Soc. **19** (1969), 164–176. MR **40:**1484

[12] ———, *Subgroups of the classical modular group*, J. Lon. Math. Soc. **1** (1969), 351–357. MR **39:**5477

[13] K. Wohlfahrt, *An extension of F. Klein's level concept*, Ill. J. Math. **8** (1964), 529–535. MR **29:**4805

Department of Mathematics, Princeton University, Princeton, New Jersey 08544

E-mail address: timhsu@math.princeton.edu

Current address: Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109

E-mail address: timhsu@math.lsa.umich.edu

Nils-Peter Skoruppa

HEIGHTS

Notes d'un cours de DEA

Ecole doctorale de Mathématiques — Université Bordeaux 1

Version: Id: heights.tex,v 1.3 2003-11-20 13:06:09 fenrir Exp

www.countnumber.de

Preface

These are the notes of a *cours de DEA avancé* held at Bordeaux in spring 1998. The aim of the course was to introduce the notion of height, one of the basic ingredients in Diophantine geometry, in an elementary and easy to understand manner, with the emphasis on results, open problems and 'highlights' instead of abstract theory.

Accordingly we start in Part 1 with the classical Lehmer conjecture and discuss the important theorems around and towards this conjecture. In particular, we discuss Langevin's theorem and Zhang's theorem. When I prepared the course, it came to my mind that the theorem of Langevin, the more recent theorem of Zhang and the long-known result of Schintzel on the absolute bound for heights of real-algebraic numbers seem to have some deep analogy. In the present notes I tried to work this out, and in the end I managed (at least) to give a sort of unified proof for these results.

In Part 2 we discuss, after a generalisation of Zhang's theorem to plane affine algebraic curves, heights on elliptic curves. We discuss in an explicit manner the method of infinite descent (and Mordell's theorem), and the local decomposition of the canonical height, i.e. the "local Green's functions" on an elliptic curve. In the Appendix (which is actually an examination given to the students at the end of the course) the reader finds a sketch of how to compute explicitly the canonical height function on an explicitly given cubic algebraic curve.

A logical step would have been a third part treating Green's functions an algebraic curves of arbitrary genus, and, in particular, to have very concrete examples, a part treating Green's functions on modular curves.

These notes are still preliminary: a section on elliptic curves is missing (section 2.4), more recent considerations on the relation of Mahler measures and special values of certain L-functions and Mahler measures as entropy of certain "algebraic dynamical systems" are missing. Also the list of references is not complete, and the Appendix is in French. Maybe we shall come back to this (and also the Part 3) at another occasion.

We finally note that parts of Part 1 are based on a course given by the author in the Max-Planck-Institute für Mathematik in spring 1993. Criticism, comments and pointers to typos are welcome.

Talence, March 8, 1999
Nils-Peter Skoruppa

Contents

Part 1

Heights of Algebraic Numbers

It is natural to try to associate to an algebraic solution of a Diophantine equation a measure of complexity. This is natural in view of the problem of computing and storing such a number, but it has also a theoretical significance if the measure of complexity can be chosen such that the number of measured objects in question below a given bound is always finite.

In this first part we shall consider the problem of finding such a measure for algebraic numbers. This will lead to the notion of height for numbers. We shall discuss various properties of the height function, and in particular we shall discuss the Lehmer conjecture.

1.1 The height of a rational number

Assume that $\alpha = \frac{x}{y}$ is a rational number, say $\gcd(x, y) = 1$. We define its height by

$$H(\alpha) := \max(|x|, |y|).$$

This clearly measures the complexity of α in the sense of how many information do we need to describe α. Indeed, $\log H(\alpha)$ is roughly the number of digits needed to write down the numerator or denominator of α. Moreover it is clear that the set

$$\{\alpha \in \mathbb{Q} : H(\alpha) < B\}$$

is finite for any real B.

There is one important property that one can already read off in this more or less trivial situation. The height function possesses a decomposition into local factors. We explain this in detail.

Recall that to each (rational) prime p we can associate the valuation $|\cdot|_p$ of $\mathbb{Q}$ defined by

$$|\alpha|_p = p^{-n},$$

where p^n is the exact power of p in the prime decomposition of α. A valuation of a field K is a function $v : K^* \to \mathbb{R}_{\geq 0}$ such that $v(\alpha) = 0$ if and only if $\alpha = 0$ and satisfying

$$v(\alpha\beta) = v(\alpha)v(\beta), \quad v(\alpha + \beta) \leq v(\alpha) + v(\beta)$$

for all $\alpha, \beta \in K^*$. Two valuations v and w are called equivalent if there is a real number $s > 0$ such that $v(\alpha) = w(\alpha)^s$ for all $\alpha \in K$. Any valuation of $\mathbb{Q}$ is either equivalent to a $|\cdot|_p$ or to the usual absolute value on $\mathbb{Q}$ which we denote by $|\cdot|_\infty$ [Neuk], p. 124. The latter writing suggest that the set of primes of $\mathbb{Q}$ should be completed by a "prime at infinity".

Theorem 1.1. *For each rational number $\alpha \neq 0$ one has*

$$H(\alpha) = \max(1, |\alpha|_\infty) \prod_p \max(1, |\alpha|_p).$$

Proof. As before let x and y denote the numerator and denominator of α. The contribution from the pth factor on the right equals p^{-n}, where p^n is the exact divisor of α, if n is negative, and it equals 1 otherwise. Thus the product over the primes equals $|y|$. The factor before the product is 1 if $|x| < |y|$, and it is $|x|/|y|$ otherwise. This proves the formula. □

In view of the theorem it is reasonable to call the function

$$H_p(\alpha) := \max(1, |\alpha|_p)$$

the local height of α at the prime p. The decomposition formula of H can then be rewritten in a more compact form as

$$H = \prod_p H_p,$$

where this time p runs through the finite primes and $p = \infty$.

1.2 The Mahler measure of a polynomial

An algebraic number α is, up to "equivalence", described by its unique normalised minimal polynomial f. By the last we understand the minimal polynomial whose coefficients are in $\mathbb{Z}$ and are relatively prime. Discussing the complexity of α is thus equivalent to discussing the complexity of f.

To measure the complexity of a polynomial

$$f = a_n X^n + a_{n-1} x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]$$

we may consider the number

$$\|f\|_1 := \sum_{j=0}^{n} |a_j|.$$

This is in essence the number of digits needed to write down f. However, one might find good arguments to consider

$$|f\|_\infty = \max_j |a_j|$$

or

$$|f\|_2 := \sqrt{a_d^2 + \cdots + a_0^2}$$

as complexity measure.

Obviously we would prefer a unique, canonical one instead of many. Now the above three examples all come from a norm on the real vector space $\mathbb{R}[x]_n$ of real polynomials of degree less or equal to n. All such norms are equivalent, i.e. if $\|\cdot\|$ is any norm on $\mathbb{R}[x]_n$, then there exist constants $A, B > 0$ such that

$$A\,\|f\|_\infty \leq \|f\| \leq B\,\|f\|_\infty$$

for all f (exercise). Suppose we could construct for any f in a canonical way a sequence of polynomials f_k of degree less or equal to n such that, for any norm $|||\cdot\|$, the measures $\|f_k\|^{\frac{1}{k}}$ are roughly $\|f\|$, and such that $\|f_k\|^{\frac{1}{k}}$ converges. By the last property the limit would not depend on the special choice of the norm as follows easily from the equivalence inequalities (see the proof of the next lemma for details). The limit can thus be considered as a good candidate for a canonical measure of complexity.

Such a sequence f_k can indeed be constructed. Let, for the following, f denote a polynomial with complex coefficients, say

$$f(x) = a_n x^n \cdots + a_0 = a_n \prod_{j=1}^{n} (x - \alpha_j).$$

We define

$$f_k(x) = a_n^k \prod_{j=1}^{n} (x - \alpha_j^k) = (-1)^{k(n+1)} \prod_{\zeta^n = 1} f(x^{\frac{1}{k}}\zeta).$$

Here the product is over all kth roots of unity. One might think of $\|f_k\|^{1/k}$ as being obtained by a sort of averaging over the roots of f. Note that f_k has rational coefficients if f has rational coefficients (since f_k is invariant

under the Galois group of the decomposition field of f), which are, moreover, integral if those of f are integral.

We define the Mahler measure of f by

$$\mu(f) = |a_n| \prod_{j=1}^{n} \max(1, |\alpha_j|).$$

Thus $\mu(f)$ is, up to the number $|a_n|$ the product of the roots of f outside the unit circle, where multiple roots are repeated. We shall need a formula expressing $\mu(f)$ without making explicit reference to the zeroes of f.

Theorem 1.2. *(Jensen's formula) For any $f \in \mathbb{C}[X]$, $f \neq 0$ one has*

$$\log \mu(f) = \int_0^{2\pi} \log |f(e^{it})| \, dt.$$

Proof. Since the logarithm is additive It suffices to consider the case $f(z) = z - \alpha$. If $|\alpha| > 1$ then $\log |f(z)|$ is a harmonic function in a neighbourhood of the unit circle, and hence the integral equals

$$\log |f(0)| = \log |\alpha|.$$

If $|\alpha| < 1$ then $g(z) = 1 - \overline{\alpha} z$ has no zeroes in the unit circle, and $\log |g(x)|$ is a harmonic function in a neighbourhood of the unit circle. Moreover, $|g(z)| = |f(z)|$ on the unit circle. The integral in question thus equals the same integral but with f replaced by g, i.e. it equals

$$\log |g(0)| = 0.$$

Finally, if $|\alpha| = 1$, then

$$\frac{1}{2\pi} \int_0^{2\pi} \log |\mathrm{e}^{it} - \alpha| \, dt = \frac{1}{2\pi} \int_0^{2\pi} \log |\mathrm{e}^{it} - 1| \, dt = 0$$

(exercise). □

The actual formula known as Jensen's formula in complex analysis applies to a slightly more general functions than only to polynomials as stated in the theorem [Ahlf], p. 205.

We need a second property of the mahler measure.

Theorem 1.3. *(Norm inequality) For all $0 \leq j < n$ one has*

$$|a_j| \leq \binom{n}{j} \mu(f).$$

Proof. This is an immediate consequence of

$$a_j = (-1)^{n-j} a_n \sum_{\{j_1,\dots,j_{n-j}\}\subset\{1,\dots,n\}} \alpha_{j_1} \cdots \alpha_{j_{n-j}}$$

and the very definition of the Mahler measure. □

We are now able to explain why the Mahler measure is the canonical complexity measure we are looking for.

Theorem 1.4. *Let $\|\cdot\|$ be a norm on the real vector space $\mathbb{C}[x]_n$. Then, for any polynomial $f \in \mathbb{C}[x]_n$, one has*

$$\lim_{k\to\infty} \|f_k\|^{1/k} = \mu(f).$$

Proof. From the equivalence inequality we obtain

$$A^{1/k} \|f_k\|_\infty^k \le \|f_k\|^{1/k} \le B^{1/k} \|f_k\|_\infty^{1/k}$$

Thus, if the theorem holds true for the $\|\cdot\|_\infty$-norm, then it holds true for all norms.

For the $\|\cdot\|_\infty$-norm we have

$$\mu(f) \le (n+1)\,\|f\|_\infty \le 2^n (n+1)\,\mu(f).$$

The first inequality follows from Jensen's formula for $\mu(f)$ on using

$$|f(x)| \le (n+1)\,\max|a_j|$$

for $|x| = 1$. The second one is an easy consequence of the norm inequality. Now by the very definition of $\mu(f)$ one has

$$\mu(f) = \mu(f_k)^{1/k}.$$

Combined with the above inequalities this gives

$$\mu(f) \le \left[(n+1)\,\|f_k\|_\infty\right]^{1/k} \le [2^n(n+1)]^{1/k}\mu(f).$$

Letting k tend to infinity we recognise the asserted formula. □

We note that there are other possibilities for defining the complexity of a polynomial f over $\mathbb{Z}$. One might consider for example $|f(1)|$, i.e. the absolute value of the sum of the coefficients of f. Again $|f_k(1)|^{\frac{1}{k}} \to \mu(f)$: indeed $\frac{1}{k}\log|f_k(1)|$ is just the $n-th$ Riemann sum approximating the integral defining $\log\mu(f)$.

In the computer algebra system Pari [Pari] one finds the function "polred" which finds for a given unitary integral f of degree n a new polynomial g which defines the same number field but which is (probably) minimal with respect to the function

$$l(f) = \sqrt{\alpha_1^d + \cdots \alpha_n^2}$$

[CoDi]. It is easy to verify that Again $l(f_k)^{\frac{1}{k}}$ tends to $\mu(f)$ for any f.

We note two simple but remarkable properties of the Mahler measure.

Theorem 1.5. *Let f and g be any complex polynomials. Then*

$$\mu(fg) = \mu(f)\mu(g), \quad \mu(f^*) = \mu(f).$$

Here f^* is the reciprocal polynomial of f, i.e. $f^*(x) = x^{\deg f} f(1/x)$.

Proof. The first identity is evident from the definition of μ. The second one is equivalent to

$$\frac{|a_0|}{|a_n|} = \prod_{j=1}^{n} |\alpha_j|.$$

□

By the norm inequality we see that, for any degree n and any bound B, there are only finitely many polynomials $f \in \mathbb{Z}[x]_n$ such that $\mu(f) \leq B$. In particular, for any real A the number

$$\inf\{\mu(f) : f \in \mathbb{Z}[x]_n,\ \mu(f) > A\}$$

is strictly greater 0 and is attained by a finite number of $f \in \mathbb{Z}[x]^n$. Obviously the Mahler measure of an integral polynomial is always greater or equal to 1. Thus it is natural to ask first of all for those polynomials with Mahler measure 1. This question is answered by a classical theorem.

Theorem 1.6. *(Kronecker) Let $f \in \mathbb{Z}[X]$. Then $\mu(f) = 1$ if and only if all roots of f are roots of unity or 0.*

Proof. Assume that $in\mathbb{Z}[x]$ has degree n and Mahler measure 1. For the j-th coefficient $a_j^{(k)}$ of f_k we have by the norm estimate and since $\mu(f_k) = 1$ the estimate

$$|a_j^k| \leq \binom{n}{j}.$$

Thus the set of all f_k is actually finite. Hence, the set of all roots of all f_n is finite. In particular, if α is a root of f then the α_n cannot be pairwise different. Consequently $\alpha^k = \alpha^l$ for some $k > l$, i.e. either $\alpha = 0$ or else $\alpha^{k-l} = 1$.

The inverse direction of the theorem is trivial. □

The theorem is usually cited in the form that an integral algebraic number whose conjugates are less or equal to 1 is necessarily a root of unity. Note that the statement in this form becomes false if one drops the integrality assumption; counter example: $\frac{3+4i}{5}$.

In view of the preceding theorem one is naturally interested in the numbers

$$\inf\{\mu(f) : f \in \mathbb{Z}[X]_n, \mu(f) > 1\}$$

and the polynomials realizing these Mahler measures. For a given degree n these minimizing polynomials are easy to calculate. In fact, one simply lists all polynomials $f \in \mathbb{Z}[x]_n$ with, say, $\mu(f) \leq 2$. This list is not empty since $\mu(x-2) = 2$, and it is contained in the finite set S_n of all integral f with

$$|a_j| \leq 2\binom{d}{j}$$

by the norm inequality. Thus this list can be compiled by searching S_n. However note that e.g. for $n = 4$ the set S_4 comprises already

$$(4\binom{4}{0}+1)^2(4\binom{4}{1}+1)^2(4\binom{4}{2}+1) = 180625$$

elements. This can of course be cut down by some factor on using $\mu(f^*) = \mu(f)$, $\pm\mu(f(\pm x)) = \mu(f)$ by rejecting all polynomials with leading and constant term different from ± 1. In Table 1.2 we listed the result of such a computational research for degrees $n \leq 5$.

n	f	$\mu(f)$	$\mathrm{disc}(f)$
1	$x-2$	2	2
2	x^2-x-1	1.618...	5
3	x^3-x+1	1.324...	-23
4	x^4-x^3-1	1.380...	-283
5	$x^5-x^4+x^3-x+1$	1.349...	$17 \cdot 97$

Table 1.1: This table gives, for a given degree $n \leq 5$, a polynomial f from the set T_n whose Mahler measure is minimal. Here T_n is the set of all irreducible polynomials in $\mathbb{Z}[x]$ of degree n with Mahler measure strictly greater than 1. The respective minimum is also attained by the polynomials $\pm f(\pm x)$ and $\pm x^n f(\pm 1/x)$, but by no other ones in T_n.

1933 Lehmer conjectured [Lehm] that even the number

$$\mu_1 := \{\mu(f) : f \in \mathbb{Z}[X],\ \mu(f) > 1\}$$

is strictly greater than one. He even conjectured that μ_∞ is assumed by the polynomial

$$f_L(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1.$$

Here

$$\mu(f_L) = 1.176\ldots$$

The conjecture is still unproven and Lehmer's lower bound is still not beaten. A huge amount of computations has been done [Boyd] giving evidence to Lehmer's conjecture.

1.3 Pisot and Salem numbers

In view of the Lehmer conjecture it is an amusing sport to find polynomials f in $\mathbb{Z}[x]$ with minimal Mahler measure $\mu(f)$. A first naive approach is to look at polynomials with small $\|f\|_\infty$ norm, say with coefficients equal to ± 1. Systematic searches in this direction have been done e.g. in [Boyd].

A more theoretic approach is to search for algebraic numbers who are not "too far away" from roots of unity. Indeed, since the Mahler measure is multiplicative and greater and equal to 1 it suffices to look at irreducible polynomials f. Moreover, since $\mu(f)$ is greater than or equal to the constant and the leading term of f, it suffices to look at polynomials where both are equal to 1, i.e. at minimal polynomials f of algebraic units α. Now, one might expect that the Mahler measure $\mu(f)$ is small if many of the conjugates of α lie in the unit disk or on the unit circle.

An integral algebraic number α is called a Pisot number if $\alpha > 1$ and all its conjugates α' satisfy $|\alpha'| < 1$. It is called a Salem number if $\alpha > 1$, if all its conjugates α' satisfy $|\alpha'| \leq 1$, but if at least one α' satisfies $|\alpha'| = 1$.

A Salem number satisfies actually a stricter condition.

Theorem 1.7. *Let f be the normalized minimal polynomial of a Salem number τ. Then $f^* = f$.*

Proof. Let τ' be a conjugate of τ on the unit circle. Then τ' is a root of f and of f^*. Thus $f^* = \pm f$. If $f^* = -f$ then $f(1) = 0$, which is impossible. □

Hence, for a Salem number τ, the set of its conjugates is stable under $z \mapsto 1/z$. In particular, a Salem number is a unit, and moreover we have:

Corollary 1.7.1. *An algebraic integer τ is a Salem number if and only if $1/\tau$ is conjugate to τ and all other conjugates of τ have absolute value 1.*

There are infinitely many Pisot numbers. Indeed, if α is a Pisot number, then so are the powers α^n $(n = 1, 2, \dots)$, respectively. Moreover, the integers are trivially Pisot numbers, and $\frac{1+\sqrt{5}}{2}$ is one. It is easy to construct others using the theorem of Rouché: If f and g are two polynomials such that $|f(z) - g(z)| < |g(z)|$ for all z on a circle $C : |z| = R$, then f and g have the same number of zeroes (counting multiplicities) inside the circle $|z| < R$. (Since the inequality implies that $F := f/g$ satisfies $|F(z) - 1| < 1$ on C, i.e. the curve $F \circ C$ is contained in the open disk $|w - 1| < 1$, and hence its winding number $\int_C d \log F$ around 0 equals 0. But this winding number is the number of zeros minus the number of poles of f/g contained in $|z| < R$.)

Theorem 1.8. *Let* $f = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x]$ *such that*

$$1 + |a_{n-2}| + |a_{n-3}| + \cdots + |a_0| < |a_{n-1}|.$$

Then exactly one root α *of* f *satisfies* $\alpha| > 1$*, and all other roots* α' *satisfy* $|\alpha'| < 1$*. In particular,* α *or* $-\alpha$ *is a Pisot number.*

Proof. The inequality implies $|f(z) - a_{n-1}z^{n-1}| < |a_{n-1}z^{n-1}|$ on the circle $|z| = 1$. By Rouchés theorem f has thus $n - 1$ roots inside the unit disk $|z| < 1$, and since $|a_0| \geq 1$, it has then exactly one root α outside the unit circle. Since $\overline{\alpha}$ is also a root of f we have $\overline{\alpha} = \alpha$, hence α or $-\alpha$ is real and greater that 1. □

The smallest Pisot number has been determined by Siegel [Sieg]. It is the real root of

$$f_S = x^3 - x - 1$$

(see section 1.8).

For Salem numbers there is a construction due to Salem, also based on Rouché's theorem.

Theorem 1.9. *Let* f *be the minimal polynomial of a Pisot number of degree greater or equal to 3, let* $\kappa = \pm 1$*, and set* $p_n = x^n f + \kappa f^*$*. Then there is an* n_0 *such that, for any* $n \geq n_0$*, one root of* p_n *is a Salem number.*

Proof. We leave it as an exercise to show that there is some n_0 such that for all sufficiently small $\varepsilon > 0$ and all $n \geq n_0$ one has $|z^n p(z)/p^*(z)| > 1$, i,e. $|p_n(z) - z^n p(z)| < |z^n p(z)|$, on the circle $|z| = 1 + \varepsilon$. Hence p_n has $n + \deg p = \deg p_n - 1$ zeroes on $|z| \leq 1$, and exactly one, say α, outside the unit circle. Since $p_n^* = \pm p_n$, the set of zeroes of p_n is invariant under $z \mapsto 1/z$. Hence all zeros different from α and $1/\alpha$ must lie on the unit circle $|z| = 1$. □

It is not yet known whether there is a smallest Salem number. A proof (or disproof) of this fact would be an important contribution towards deciding the Lehmer conjecture. The smallest*known* Salem number can be obtained by Salem's construction:

$$z^7 f_S - f_S^* = x^8(x^3 - x - 1) - (-x^3 - x^2 + 1) = (x - 1)f_L,$$

i.e. the unique root outside the unit circle of the polynomial f_L of Lehmer (see the end of last section), which has the so-far smallest known Mahler measure.

1.4 The height of an algebraic number

Before discussing further the Mahler measure and the known results in the direction of the Lehmer conjecture, we introduce its more number theoretic counter part, namely, the height of algebraic numbers. For an algebraic number α of degree n we define its "absolute" height by

$$H(\alpha) = \mu(f)^{1/n},$$

where f is the normalized minimal polynomial of α. The normalizing power $1/n$ is usually inserted to have a decomposition formula of the height function in local contributions which does depend on the field from which the valuations are taken. We shall explain this more precisely in a moment (see the proof of the next theorem).

If we set

$$f = a_n \prod_{j=1}^{n} (x - \alpha_j),$$

then

$$H(\alpha) = \Big[|a_d| \prod_{j=1}^{d} \max(1, |\alpha_j|) \Big]^{1/n}.$$

Note that this generalizes the height of a rational number defined in the first section. Indeed, if $\alpha = \frac{r}{s}$ with relative prime integers r, s, then $f = sx - r$, and hence $\mu(f) = |s|$ if $|r| \leq |s|$, and $\mu(f) = |r|$ otherwise.

As for the height of rational numbers one has a decomposition into local height contributions. We recall first of all the relevant facts about the valuations of an arbitrary number field K.

An equivalence class of valuations of K is called place of K or a prime of K. We always use P_K for the set of places of K, and we use P_K^∞ for the set

of archimedean places of K, i.e. the set of equivalence classes of valuations which extend the usual absolute value on $\mathbb{Q}$ (up to equivalence).

The representatives $|\cdot|_v$ for the places v of K can be chosen in a unique way that one has

$$\prod_{v\in P(K)_\infty} |\alpha|_v = |\mathrm{N}_{K/\mathbb{Q}}(\alpha)|$$

and

$$\prod_{v\in P_K} |\alpha|_v = 1.$$

for all $\alpha \in K$. We always assume that $|\cdot|_v$ is normalized in this way.

One can describe the $|\cdot|_v$ explicitly as follows. To each prime ideal $\mathfrak{p}$ of K one can associate a valuation by

$$|\alpha|_\mathfrak{p} = \mathrm{N}_{K/\mathbb{Q}}(\mathfrak{p})^{-k},$$

where $\mathfrak{p}^k$ is the exact divisor of α. This valuation satisfies the stronger triangle inequality

$$|\alpha+\beta|_\mathfrak{p} \le \min(|\alpha|_\mathfrak{p}, |\beta|_\mathfrak{p})$$

with equality if $|\alpha|_\mathfrak{p}$ and $|\beta|_\mathfrak{p}$ are different.

Let $\sigma_j : K \mapsto \mathbb{R}$ $(1 \le j \le r)$ be the real embeddings of K, and let $\sigma_j, \overline{\sigma_j} : K \mapsto \mathbb{C}$ $(r < j \le r+s+1)$ be the pairs of complex embeddings of K. Then, for each j we have the valuation

$$|\alpha|_j = |\sigma_j(\alpha)|^{e_j},$$

where the bars on the right indicate the usual absolute value in $\mathbb{R}$ or $\mathbb{C}$, and where $e_j = 1$ if σ_j is real, and $e_j = 2$ if σ_j is complex.

It is a known fact that for any place v of K the valuation $|\cdots|_v$ equals $|\cdot|_\mathfrak{p}$ for some $\mathfrak{p}$ if it is finite (i.e. non-archimedean), and that it equals $|\cdot|_j$ for some j otherwise.

We shall also need the following two facts. Let L be a finite extension of K. Then the compatibility formula holds true, i.e.

$$|\alpha|_v^{[L:K]} = \prod_{\substack{w\in P_L \\ w|v}} |\alpha|_w$$

for all α in the ground field K. Here $w|v$ means that $|\cdot|_w$ is an extension of $|\cdot|_v$ up to equivalence. Example: For $5 = (1+2i)(1-2i)$ and $K = \mathbb{Q}(i)$ one finds

$$|5|_{(5)}^2 = |5|_{(1+2i)}|5|_{(1-2i)},$$

where on the left we have the 5-adic valuation on $\mathbb{Q}$, and on the right the corresponding valuations on $\mathbb{Q}(i)$.

If L is galois over K then, for any place v of K the Galois group $\mathrm{Gal}(L/K)$ acts transitively on the places w of L dividing v.

We are now in the position of proving the following decomposition formula for the absolute height.

Theorem 1.10. *Let K be a number field and $\alpha \in K$. Then one has*

$$H(\alpha) = \prod_{v \in P_K} \max(1, |\alpha|_v)^{1/[K:\mathbb{Q}]}.$$

Proof. Note first of all that the value of the right hand side does not depend on the field K. This is an immediate consequence of the compatibility formula and the fact that $|\alpha|_v < 1$ if and only if $|\alpha|_w < 1$ for all $w|v$ in any extension L of K.

The formula is trivial in the case that $K = \mathbb{Q}(\alpha)$ and α is integral. Indeed, in this case $|\alpha|_v \leq 1$ for all finite places v, and hence the $[K : \mathbb{Q}]$th power over the local contributions equals

$$\prod_{j=1}^{r} \max(1, |\sigma_j(\alpha)|) \prod_{j=r+1}^{s} |\max(1, \sigma_j(\alpha)|^2),$$

with σ_j having the same meaning as before. But this is exactly $\mu(f)$.

In the general case one can proceed as with the case of an integral α to prove

$$H(\alpha) = |a_n|^{1/[\mathbb{Q}(\alpha):\mathbb{Q}]} \prod_{v \in P_K^\infty} \max(1, |\alpha|_v)^{1/[K:\mathbb{Q}]},$$

where $|a_n|$ is the leading term of the normalized minimal polynomial of f. Hence it remains to relate $|a_n|$, the leading term of the normalized minimal polynomial of f to the factors associated to the finite primes of K.

For this one uses Gauss's lemma [Heck], p. 105:

$$\mathrm{cont}_v(g_1 g_2) = \mathrm{cont}_v(g_1)\,\mathrm{cont}_v(g_2).$$

for all $g_1, g_2 \in K[X]$, and all finite places v of the number field K. Here

$$\mathrm{cont}_v(a_m x^m + \cdots + a_0) = \max_j |a_j|_v.$$

By enlarging K if necessary, we may assume that K is Galois, say with Galois group G, and contains all roots of f. We can write f in the form

$$f = a_n \prod_{\sigma \in G} \big(x - \sigma(\alpha)\big)^{1/[K:\mathbb{Q}(\alpha)]}.$$

Let p be a rational prime number. We then have

$$\begin{aligned} 1 = \mathrm{cont}_p(f)^{[K:\mathbb{Q}(\alpha)]} &= |a_n|_p^{[K:\mathbb{Q}(\alpha)]} \,\mathrm{cont}_p\Big(\prod_{\sigma\in G}(x-\sigma(\alpha))\Big) \\ &= |a_n|_p^{[K:\mathbb{Q}(\alpha)]} \prod_{\sigma\in G}\prod_{v|p} \mathrm{cont}_v(x-\sigma(\alpha))^{1/[K:\mathbb{Q}]} \\ &= |a_n|_p^{[K:\mathbb{Q}(\alpha)]} \prod_{v|p}\prod_{\sigma\in G} \max(1,|\sigma(\alpha)|_v)^{1/[K:\mathbb{Q}]} \\ &= |a_n|_p^{[K:\mathbb{Q}(\alpha)]} \prod_{v|p} \max(1,|\alpha|_v) \end{aligned}$$

Here the second identity follows from Gauss's lemma, the third one from the compatibility relation and Gauss lemma, and the last since G acts transitively on the places of K dividing p. Thus we find

$$|a_n| = \prod_{p \text{ finite}} \frac{1}{|a_n|_p} = \prod_{p \text{ finite}}\prod_{v|p} \max(1,|\alpha|_{\mathfrak{p}})^{1/[K:\mathbb{Q}(\alpha)]},$$

which implies the asserted formula. □

Sometimes one defines for a number field K the relative height function H_K on K by by

$$H_K(\alpha) = \prod_{v\in P_K} \max(1,|\alpha|_v).$$

In particular one has

$$H_{\mathbb{Q}(\alpha)}(\alpha) = \mu(f)$$

, where f is the normalized minimal polynomial of α.

Since the Mahler measure and the height of algebraic numbers represent essentially the same notion, one can easily deduce several properties of the height from the Mahler measure.

For instance, for any degree d and any bound B there is only a finite number of $\alpha \in \overline{\mathbb{Q}}$ of degree less or equal to d with $H(\alpha) \le B$. Moreover

$$H(\alpha) = H(1/\alpha)$$

for any $\alpha \neq 0$ (since $\pm f^*$ is the normalized minimal polynomial of $1/\alpha$ if f is the normalized minimal polynomial of α. Kronecker's theorem states states that, for any algebraic α one has $H(\alpha) = 1$ if and only if α is a root of unity.

Finally, Lehmer's conjecture is equivalent to the fact that for some constant $C > 1$ one has

$$H(\alpha) \ge C^{\deg \alpha}$$

for all algebraic α. Note the degree $\deg \alpha$ of α in this formula. Without this degree such an inequality cannot be true. Counter example:

$$H(2^{1/n}) = 2^{1/n} \mapsto 1.$$

1.5 Two easy Lehmer type theorems

In this section we prove a theorem of Schinzel from 1973 concerning an absolute lower bound for the height of totally real algebraic numbers, and a more recent one one of Zhang from 1992 about numbers simultaneously "close to 0 and 1". Both theorems admit surprisingly simple proofs ([HoSk] and [Zagi]) which are quite similar though they were found independently [1] . In this section we give the the proofs without any additional comment. A reinterpretation and two possible generalizations will follow in the two next sections.

Theorem 1.11. *(Schinzel [Schi]) Let $\alpha \neq 0, \pm 1$ be a totally real algebraic number. Then*

$$H(\alpha) \geq \sqrt{\frac{1+\sqrt{5}}{2}} = 1.2720\ldots,$$

with equality if and only if α equals one of the four numbers $\pm\frac{1\pm\sqrt{5}}{2}$.

Note that a theorem like this cannot be true in general, i.e. there is no absolute lower bound for the absolute height of all but a finite number of algebraic numbers. Indeed, $x^p - a$ is irreducible for all square-free positive integers a, and all rational primes p. Thus

$$H(\sqrt[p]{a}) = a^{1/p} \to 1.$$

Proof. (cf. [HoSk]) If, for x real, we set $\gamma(x) = |x|^{1/2}|x - 1/x|^{1/2\sqrt{5}}$, then we have

$$\max(1, |x|) \geq \sqrt{\frac{1+\sqrt{5}}{2}}\, \gamma(x),$$

with equality if and only if $x = \pm\frac{1\pm\sqrt{5}}{2}$. Indeed, since $|x|\gamma(1/x) = \gamma(x)$, since the same invariance property holds for the function $\max(1, |x|)$, and since both sides of the desired inequality are invariant under $x \mapsto -x$, it suffices

[1] The second proof differs slightly from the original version given in [Zagi]. When I prepared this manuscript I noticed that Zagier's proof could be presented in a form which makes it look much more similar to the proof of Schinzel's theorem in [HoSk].

to prove it for $0 \le x \le 1$. But in this interval maximum of $\gamma(x)$ occurs for $x = \frac{-1+\sqrt{5}}{2}$ with maximum value $\sqrt{\frac{-1+\sqrt{5}}{2}}$.

On the other hand

$$|a| \prod_j \gamma(\alpha_j) = |a|^{1/2-1/2\sqrt{5}} |f(0)|^{1/2-1/2\sqrt{5}} |f(1)f(-1)|^{1/2\sqrt{5}} \ge 1$$

where $f(x) = a\prod(x-\alpha_j)$ is the minimal polynomial of α. The result is now obvious. □

Theorem 1.12. *(Zhang [Zhan]) For all algebraic numbers $\alpha \ne 0, 1, \frac{1\pm\sqrt{-3}}{2}$, one has*

$$H(\alpha)H(1-\alpha) \ge \sqrt{\frac{1+\sqrt{5}}{2}}$$

with equality if and only if α or $1-\alpha$ is a primitive 10th root of unity.

Proof. (Cf. [Zagi]) Here, for complex z, we set

$$\gamma(z) = |z|^{1/2}|1-z|^{1/2}\Big(\frac{|z^2-z+1|}{|z^2-z|}\Big)^{1/2\sqrt{5}}.$$

It is straight-forward, though cumbersome, to prove

$$\max(1,|x|)\max(1,|1-x|) \ge \sqrt{\frac{1+\sqrt{5}}{2}}\,\gamma(x)$$

for all complex arguments, with equality if and only if x or $1-x$ equals $\mathrm{e}^{\pm\pi i/5}$ or $\mathrm{e}^{\pm 3\pi i/5}$. With the same notations as in the theorem before we find

$$|a| \prod_j \gamma(\alpha_j) = |f(0)f(1)|^{1/2-1/2\sqrt{5}} |f(\frac{1+\sqrt{-3}}{2})f(\frac{1-\sqrt{-3}}{2})|^{1/2\sqrt{5}} \ge 1,$$

which again implies the result. □

1.6 Numbers with conjugates outside a given set

The proofs of the two theorems of the preceding section have obviously very much in common. The both use a function $\gamma(z)$ with remarkable symmetry to bound $\max(1,|z|)$ to below. We formalize the construction of this bounding function.

Fix an arbitrary polynomial $p \neq 0$ with integral coefficients and a real number $s > 0$, and set

$$\gamma(z) := |z|^{1/2}|p(z)p(1/z)|^s.$$

Then

$$|z|\gamma(1/z) = \gamma(z).$$

Note that we also have

$$|z| \max(1, |1/z|) = \max(1, zx|).$$

Moreover, $\gamma(z)$ and $\max(1, |z|)$ are both invariant under $z \mapsto \overline{z}$. Thus if

$$\max(1, |z|) \geq \gamma(z)$$

for z in some subset E of $\mathbb{C}$, then we can assume without loss of generality that E is invariant under $z \mapsto 1/z$ and $z \mapsto \overline{z}$. By continuity we can furthermore assume that E is closed. The invariance of $z \mapsto 1/z$ implies that, for proving the desired estimate, we only have to look at arguments z in the intersection of E with the unit disk $|z| \leq 1$. Using that E is stable under complex conjugation it even suffices to look at the intersection of E with the unit circle. More precisely we have the following lemma.

Lemma 1.1. *Let E be a closed subset of $\mathbb{C}$ invariant under complex conjugation and under $z \mapsto 1/z$. Let $p \in \mathbb{C}[x]$, and suppose that*

$$\sup_{z \in E,\ |z|=1} |p(z)| < 1.$$

Then, for all sufficiently small $s > 0$ there exists a constant $C > 1$ such that

$$\max(1, |z|) \geq C\,|z|^{1/2}|p(z)p(1/z)|^s.$$

for all $z \in E$.

Proof. Denote by $\mathbb{D}$ the closed unit disk $|z| \leq 1$. We claim that there is a non-negative integer l such that

$$a := \sup_{z \in E \cap \mathbb{D}} |z^l p^*(z)p(z)| < 1.$$

Indeed, otherwise there would be a sequence z_k in $E \cap \mathbb{D}$ and of non-negative integers n_k such that $|z_k^{n_k} p^*(z_k)p(z_k)| \geq 1$. Since $E \cap D$ is compact, we may assume that z_k converges towards an $w \in E \cap D$. For this w we have by continuity $|p(w)p^*(w)| \geq 1$. If we had $|w| = 1$ then $|p(w)| < 1$, hence

$|p^*(w)| > 1$. But the latter is impossible since $|p^*(w)| = |p(\overline{w})|$ for $|w| = 1$, and since $\overline{w} \in E$ by the invariance of E under complex conjugation. But then $|w| < 1$, and hence $|p(z_k)p^*(z_k)| \geq |z_k|^{-n_k} \to \infty$, which is absurd.

Thus, for any $s \leq 1/2(l + \deg p)$ there is a $C > 1$ such that the desired estimate holds for all $z \in E \cap \mathbb{D}$. But this holds then true for all $z \in E$ by the transformation formulas of both sides of the inequality under $z \mapsto 1/z$. □

The lemma explains how to find a function $\gamma(x)$ as used for instance in the proof of Schinzel's theorem: There $E = \mathbb{R}$, thus any integral polynomial p with $|p(x)| < 1$ for $x \in \{\pm 1\}$, the intersection of $\mathbb{R}$ with the unit circle, would lead to a lower bound $C > 1$ for totally real numbers. The polynomial $p(x) = x^2 - 1$ is certainly the simplest solution, and it it exactly the one which one finds in our proof.

Now suppose finally that we have an integral $p \neq 0$ satisfying the hypothesis of the lemma. Then this yields immediately an absolute lower bound for the heights of algebraic numbers all of whose conjugates lie outside E.

Lemma 1.2. *Let E, p and $C > 1$ as in the preceding lemma. Suppose furthermore that $p \in \mathbb{Z}[x]$ and $p \neq 0$. Then*

$$H(\alpha) \geq C$$

for all α such that α and all its conjugates are contained in E and different from 0 and the roots of p and p^*..

Since $H(\alpha) = 1$ for roots of unity α, we see that the existence of a p satisfying the hypothesis of the lemma implies that, for each integer $n \geq 1$, either $\mathbb{C} \setminus E$ contains at least one primitive root of unity, or else p has all nth roots of unity as roots. In particular, $\mathbb{C} \setminus E$ must intersect the unit circle non-trivially.

Proof. By the preceding lemma we have the following estimate:

$$H(\alpha)^n C^{-n} \geq |a| \prod_{j=1}^{n} \gamma(\alpha_j) = |a|^{1/2-2ls} |f(0)|^{1/2-ls} a^{ls} \prod_{j=1}^{n} |p(\alpha_j)p^*(\alpha_j)|^s,$$

where l is the degree of p. But $|f(0)|$ and the factor after it are positive powers of positive integers. Hence, for $s < 1/4l$, we obtain

$$H(\alpha) \geq C.$$

□

It is certainly natural to start now with a set E, and to ask when we can find an integral $p \neq 0$ satisfying the assumptions of the first lemma. The answer will be found using the theory of transfinite diameters whose basics we shall develop in the next section. We shall find as answer:

Lemma 1.3. *Let E be a closed subset of $\mathbb{C}$ not containing the whole unit circle and stable under complex conjugation. Then there is polynomial $p \in \mathbb{Z}[x]$, $p \neq 0$, such that*

$$\sup_{z \in E,\ |z|=1} |p(z)| < 1.$$

Note that the condition that E does not contain the unit circle, is also necessary. Otherwise we would be able to prove that $H(\alpha)$ is absolutely bounded below by a constant greater than 1 for all α which are not roots of unity or 0, which is certainly false (see the counter example $H(\sqrt[n]{2})$ at the end of section 1.4. Postponing the proof of the preceding lemma to the next section we can summarise by saying:

Theorem 1.13. *(Langevin [Lvin]) Let G be an open region in $\mathbb{C}$ which intersect the unit circle $|z| = 1$. Then there exists a constant $C(G) > 1$ such that*

$$H(\alpha) \geq C(G)$$

for any $\alpha \in \overline{\mathbb{Q}}$ which has no conjugates in G, which is not a root of unity and different from 0.

Proof. This is a consequence of the preceding lemmas. For applying the lemma 1.1 we actually need that G, or equivalently $E := \mathbb{C} \setminus G$, is invariant under complex multiplication and $z \mapsto 1/z$. However, this can be assumed without loss of generality. If z_0 is on the unit circle and in G, then there is an open neighbourhood of z_0 and contained in G which is stable under $z \mapsto 1/\overline{z}$ (exercise). Clearly, we may replace G by this neighbourhood. Secondly, we may then replace G by the union of G and $G^* := \{\overline{z} : z \in G\}$, since all conjugates of α are outside G if and only if they are outside G^*. The resulting G has then the necessary invariance properties.

By the last lemma we find, for E, a p satisfying the hypothesis of Lemma 1.1. Thus Langevin's theorem is true for all $\alpha \neq 0$ having no conjugates in G apart from possibly those roots β of p or p^* which are not roots of unity. (Note in passing that any lth root of unity has a conjugate in G if $l \gg 0$, and that any root of unity for which this does not hold true must be a root of p). However, the heights of the β are strictly larger than 1. Hence by choosing a $C(G) > 1$ which is smaller than these finitely many heights and smaller than C, we obtain the desired theorem. □

There is a somewhat remarkable consequence of Langevin's theorem, which suggests that it is easier to believe Lehmer's conjecture than the contrary.

Corollary 1.13.1. *If there were a sequence f_n of polynomials in $\mathbb{Z}[x]$ with $\mu(f_n) > 1$ but $\lim \mu(f_n) = 1$ (i.e. if the Lehmer conjecture where false), then any point on the unit circle $|z| = 1$ would be an accumulation point of the roots of the f_n.*

1.7 Transfinite diameters

In this section we shall prove the main lemma 1.3, which we needed for the proof Langevin's theorem. For this we shall need the theory of transfinite diameters invented by Fekete and Tonelli [FeTo].

For a compact subset E of $\mathbb{C}$ we set

$$\rho_n(E) := \inf\{|p|_E : p \in \mathbb{C}[x]_n,\ p = x^n + \dots\}.$$

Here we use

$$|p|_E = \sup_{z \in E} |p(z)|.$$

One can show that the number $\rho_n(E)$ is even attained by a unitary polynomial $C_n \in \mathbb{C}[x]_n$, and that, even more, C_n is unique if E has more than n points [FeTo]. This C_n is called the nth Chebyshev polynomial of E.

Note that, in the definition of $\rho_n(E)$, we can restrict to real polynomials if E is stable under complex conjugation. Namely, in this case we have

$$|\frac{1}{2}(p + \overline{p})|_E \leq \frac{1}{2}(|p|_E + |\overline{p}|_E) = |p|_E$$

for any polynomial p, where $\overline{p}$ is obtained from p by taking the complex conjugates of the coefficients of p. In particular, C_n is real for such E.

By a similar argument we see that, in the case that E is the unit circle $\mathbb{S} : |z| = 1$, we can restrict to those unitary polynomials in $\mathbb{C}[x]_n$ which are invariant under $x \mapsto \zeta x$ for all nth roots of unity. Indeed,

$$\tilde{p}(x) = \frac{1}{n} \sum_{\zeta^n = 1} p(\zeta x)$$

is unitary of degree n and satisfies $|\tilde{p}|_{\mathbb{S}} \leq |p|_{\mathbb{S}}$. On the other hand, the only unitary polynomials in $\mathbb{C}[x]_n$ which are invariant under all $x \mapsto \zeta x$, are the

$x^n + c$, where c is a constant (exercise). Clearly x^n is thus the nth Chebyshev polynomial of the unit circle, and consequently

$$\rho_n(\mathbb{S}) = |x^n|_{\mathbb{S}} = 1.$$

For an integer $n \geq 0$ let $T_n(x)$ be the polynomial defined by

$$\cos(nt) = 2^{n-1} T_n(\cos t).$$

Thus T_n is a unitary polynomial of degree n. One has $T_1 = x$, $T_2 = x^2 - 1$, $T_3 = x^3 - \frac{3}{4}x$. The polynomial $2^{n-1}T_n$ is the polynomial which is usually called the nth Chebishev polynomial without making any reference to transfinite diameters. In fact, T_n is the nth Chebishev polynomial of the interval $I = [-1, +1]$ in the sense defined before.

Indeed, as is obvious from the definition, $|T_n(x)|_I = 1/2^{n-1}$. Furthermore $T_n(x)$ attains $n + 1$ times the critical values $\pm 2^{1-n}$ in the interval $[-1, +1]$, with alternating signs from left to right. Hence, if p were a unitary polynomial of degree n with $|p|_I < 2^{1-n}$, then $f - p$ would be a polynomial different from 0, of degree $\leq (n-1)$ and whose values changes $n + 1$ times the sign in I; thus it would have n zeroes, a contradiction.

The transfinite diameter (or Chebichev constant) of a compact subset E is defined as

$$\rho = \lim_n \rho_n(E)^{1/n}.$$

For the unit circle and closed intervals $[a, b]$ on the real axis we have by the preceding discussion

Theorem 1.14. *One has $\rho(\mathbb{S}) = 1$ and $\rho([a, b]) = \frac{b-a}{4}$ for all real $a \leq b$.*

Proof. The second formula follows from $\rho([-1, +1]) = \frac{1}{2}$ on using the general formulas $\rho(t + E) = \rho(E)$ and $\rho(\lambda E) = \lambda\rho(E)$ (which, in turn are an obvious consequence of $|f(x)|_{t+E} = |f(x + t)|_E$ and $|f(x)|_{\lambda E} = |\lambda||\lambda^{-1} f(\lambda x)|_E$). □

Theorem 1.15. *The limit $\rho(E)$ exists and is finite.*

Proof. Since E is compact it is contained in a disk $|z| \leq R$ for some R. Hence $\rho_n(E)^{1/n} \leq |x^n|_E^{1/n} = R$. In particular,

$$\alpha := \liminf \rho_n(E)^{1/n}, \qquad \beta := \limsup \rho_n(E)^{1/n}$$

are both finite. Let $\varepsilon > 0$, and let p a unitary polynomial, say of degree l, such that $|p|_E^{1/l} < \alpha + \varepsilon$. Then there exists a constant C such that

$$|z^r p^k|_E \leq C(\alpha + \varepsilon)^n \quad (n = kl + r,\ 0 \leq r < n)$$

for all n. Hence $\rho_n(E)^{1/n} \leq C^{1/n}(\alpha + \varepsilon) \to \alpha + \varepsilon$, and hence $\beta = \alpha$. □

We shall need

Lemma 1.4. *Let A be an arc of length $t \leq 2\pi$ on the unit circle. Then*

$$\rho(A) \leq \sin(t/4).$$

In fact one can show that one actually has equality [FeTo]

Proof. By applying a suitable rotation we may suppose that A is stable under complex conjugation, contains 1 and has end points $e^{it/2}$ and $e^{-it/2}$. Consider the map

$$R : A \to I := [\cos(t/2), 1], \quad z \mapsto \frac{1}{2}(z + \frac{1}{z}),$$

which is 2 to 1. If p is a unitary polynomial of degree n, then

$$|p|_I = |p \circ R|_A = 2^{-n}\big|(2x)^n p\big(\frac{1}{2}(x + \frac{1}{x})\big)\big|_A \geq 2^{-n}\rho_{2n}(A),$$

and hence

$$\frac{1 - \cos(t/2)}{4} = \rho(I) \geq \rho(A)^2,$$

i.e. $\sin^2(t/4) \geq \rho(A)^2$. □

Theorem 1.16. *(Kakeya) Let E be a compact subset such that $|p|_E < 1$ for some unitary polynomial $p \in \mathbb{R}[x]$. Then there exists a unitary polynomial $q \in \mathbb{Z}[x]$ with $|q|_E < 1$.*

Proof. Let p be real, unitary, say of degree n with $|p|_E < 1$. Clearly $n \geq 1$. For positive integral m write $m = qn + r$ with integral q and $0 \leq r < n$, and set

$$p_m(x) = x^r p(x)^q.$$

Then

$$|p_m| \leq b\, a^m \qquad (b = \max_{0 \leq r < n} |z^r|_E,\ a = |p|_E^{1/n}.)$$

Fix a positive integer s. For each r we can find scalars $|\lambda_j| < 1$ such that

$$L_r := p_r + \lambda_1 p_{r-1} + \cdots + \lambda_{r-s} p_s = G_r + H_r$$

with a unitary $G_r \in \mathbb{Z}[x]$ and an H of degree strictly smaller than s and whose coefficients have absolute value less than 1. One has

$$|L_r|_E \leq b(a^r + \cdots + a^s) \leq b\frac{a^s}{1 - a}.$$

Thus if s is sufficiently big, then $|L_r|_E < 1/3$ for all r. But by construction the sequence $|H_r|_E$ is bounded. Hence for some $r_1 > r_2$ we have $|H_{r_1} - H_{r_2}|_E < 1/3$. But then

$$|G_{r_1} - G_{r_2}|_E = |L_{r_1} - H_{r_1} - (L_{r_2} - H_{r_2})|_E \leq |L_{r_1}|_E + |L_{r_1}|_E + |H_{r_1} - H_{r_2}|_E < 1.$$

□

1.8 Heights of non-reciprocal numbers

We call a polynomial reciprocal if its set of roots is invariant under $z \mapsto 1/z$, and and we call an algebraic number $\alpha \neq 0$ reciprocal if the set of all conjugate numbers is invariant under $z \mapsto 1/z$. Clearly, a polynomial is reciprocal if and only if if and only if $f^* = af$ for some number a.

Theorem 1.17. *(Smyth) Let $f \in \mathbb{Z}[x]$, and assume that f is not reciprocal and $f(0) \neq 0$. Then*

$$\mu(f) \geq \theta = 1.3247\ldots,$$

where θ is the real solution of $\theta^3 - \theta - 1 = 0$.

Corollary 1.17.1. *If $f \in \mathbb{Z}[x]$ is irreducible and of odd degree, then*

$$\mu(f) \geq \theta.$$

Proof. Assume that $f = af^*$ for same integer a. Then the set of roots of f is invariant under the involution $\alpha \mapsto 1/\alpha$. Hence, if the degree of f were odd, then at least one root satisfies $\alpha = 1/\alpha$, i.e. $\alpha = \pm 1$. Hence any irreducible polynomial of odd degree is either equal to a multiple of $x+1$ or $x-1$, or else is not reciprocal. Hence Smyth theorem applies to all irreducible polynomials of odd degree. □

We may restate Smyth theorem and its corollary by saying: If α is an algebraic number of degree n such that n is odd degree or such that α is not reciprocal, then

$$H(\alpha) \geq \sqrt[n]{\theta}.$$

Corollary 1.17.2. *(Siegel) θ is the smallest Pisot number.*

Proof. The set of roots of a minimal polynomial f of a Pisot number can only be invariant under $\alpha \mapsto 1/\alpha$ if the degree of f is two. Thus Smyth theorem applies to f unless f is of degree 2. But in the latter case $\mu(f) \geq \frac{1+\sqrt{5}}{2}$ as we already saw before (see section 1.2. □

Corollary 1.17.3. *(Cassels) Assume that $f(x) = \prod_{j=1}^{n}(x - \alpha_j) \in \mathbb{Z}[x]$ satisfies $|\alpha_j| < 1 + \frac{\log\theta}{n}$ $(1 \leq j \leq n)$. Then $f = \pm f^*$.*

Proof. We remark that the original theorem of Cassels was stated with $\log\theta = 0.28\ldots$ replaced by 0.1.

For the proof we simply note that by assumption

$$\mu(f) < \left(1 + \frac{\log\theta}{n}\right)^n \leq \mathrm{e}^{\log\theta} = \theta.$$

Thus Smyth theorem cannot apply to f. □

1.9 Proof of Smyth's theorem

We follow in this section essentially the original proof in [Smy1]. For a complex number α set

$$B_\alpha(z) = \frac{z-\alpha}{1-\overline{\alpha}z}.$$

If α is inside the unit disk then B_α is holomorphic in an open neighborhood of the unit disk and satisfies $|B_\alpha(z)| = 1$ for $|z| = 1$. Let now $\alpha_1, \ldots, \alpha_r$ be complex numbers inside the unit disk, and let

$$B(z) = \prod_{j=1}^{r} B_{\alpha_j}(z) = c_0 + c_1 z + c_2 z^2 + \cdots .$$

We shall call B the Blaschke function associated to the family of the a_j.

Lemma 1.5. *One has* $1 = |c_0|^2 + |c_1|^2 + |c_2|^2 + \cdots$,

Proof. This follows from

$$1 = \frac{1}{2\pi}\int_0^{2\pi} |B(\mathrm{e}^{it})|^2\, dt = \frac{1}{2\pi}\sum_{k,l} c_k c_l \int_0^{2\pi} \mathrm{e}^{i(k-l)}\, dt.$$

□

Assume now that f is a real polynomial without zeroes on the unit circle and such that $f(0) \neq 0$. Let B and $\hat{B}$ be the Blaschke functions associated to the zeroes of f and f^* inside the unit circle, respectively (with repeated multiple roots). Then B/f has no zeroes, and its poles are the roots of f outside the unit circle and the $1/\overline{\alpha}$, where α runs through the roots of f inside the unit circle; and the same holds true for $\hat{B}/f^*$ since the set of roots of f is invariant under $z \mapsto \overline{z}$. In fact, one easily checks

$$\frac{B}{f} = \frac{\hat{B}}{f^*}.$$

Let c_k, d_k and a_k denote the Taylor coefficients of B, $\hat{B}$ and f/f^* at $z = 0$, respectively. Assume that the constant and leading term of f are equal to ± 1. Then $c := |c_0| = |d_0| = 1/\mu(f)$ and $a_0 = \pm 1$. If, furthermore, f/f^* is not constant, then there exists a smallest $k \geq 1$ such that $a_k \neq 0$, and consequently,

$$c_k - a_0 d_k = a_k d_0.$$

From this (and $|a_0| = 1$) we see that $|d_k| \geq |a_k d_0|/2 = |a_k|c/2$ or $|c_k| \geq |a_k|c/2$. Hence from the preceding lemma

$$1 \geq c^2 + \frac{|a_k|^2}{4}c^2,$$

and thus

$$\mu(f) \geq \left(1 + \frac{|a_k|^2}{4}\right)^{\frac{1}{2}}.$$

Assume now that f is integral, irreducible and not reciprocal and $f(0) \neq 0$. Then it has no roots on the unit circle (since such a root would be a root of f^* too), and f/f^* is not constant. For the proof of Smyth theorem we may moreover assume that the leading term and constant term of f is 1 (since otherwise $\mu(f) \geq 2$). Thus f satisfies the hypothesis used in the last paragraph, and accordingly the last estimate for $\mu(f)$ holds true. However, here f/f^* has integral Taylor coefficients, in particular $|a_k| \geq 1$. Hence

$$\mu(f) \geq \sqrt{\frac{5}{4}} = 1.118\cdots.$$

This is already a weak version of Smyth's theorem. His sharp bound is obtained, essentially by the same method, However, by a more subtle investigation of the coefficients of the Blaschke function than in our lemma above.

Lemma 1.6. *Let $n \geq 1$.*

1. *For all real $x_0, \ldots, x_n$ one has*

$$(c_0x_0)^2 + (c_0x_1 + c_1x_n)^2 + \cdots + (c_0x_n + \cdots + c_nx_0)^2 \leq x_0^2 + x_1^2 + \cdots + x_n^2. \quad (1.1)$$

2. *Set*

$$A = \begin{pmatrix} c_n & c_{n-1} & c_{n-2} & \cdots & c_0 \\ c_{n-1} & c_{n-2} & \cdots & c_0 & 0 \\ c_{n-2} & \cdots & c_0 & 0 & 0 \\ \vdots & & & & \\ c_0 & & & & \end{pmatrix}.$$

Then $1 + A$ and $1 - A$ are symmetric, positive definite matrices.

3. In particular, one has

$$1 \geq c_0^2 + |c_n|, \tag{1.2}$$

$$-(1 - c_0^2 - \frac{c_n^2}{1 + c_0}) \leq c_{2n} \leq 1 - c_0^2 - \frac{c_n^2}{1 - c_0}. \tag{1.3}$$

and the same inequalities with c_k replaced by d_k.

Proof. In fact, the above inequalities hold true for the Taylor coefficients at 0 of any function which is holomorphic in an open neighborhood of the unit disk $|z| \leq 1$, satisfies $|f(z)| \leq 1$ for $|z| = 1$ and has real Taylor coefficients c_j. Indeed, setting $p(z) = x_0 + x_1 z + \cdots + x_n z^n$, we have

$$\sum_{j=0}^{n} (c_0 x_j + \cdots + c_j x_0)^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(z)p(z)|^2 \, dt \qquad (z = \mathrm{e}^{it})$$

$$\leq \frac{1}{2\pi} \int_0^{2\pi} |p(z)|^2 \, dt = \sum_{j=0}^{n} x_j^2.$$

The second assertion is obtained using the Cauchy-Schwartz inequality and the first one:

$$\pm x^t A x \leq |x| \, |Ax| \leq |x|^2.$$

Let $\varepsilon = \pm 1$. Since $1 + \varepsilon A \geq 0$ we obtain in particular

$$\det \begin{pmatrix} 1 + \varepsilon c_n & \varepsilon c_0 \\ \varepsilon c_0 & 1 \end{pmatrix}, \det \begin{pmatrix} 1 + \varepsilon c_{2n} & \varepsilon c_n & \varepsilon c_0 \\ \varepsilon c_n & 1 + \varepsilon c_0 & 0 \\ \varepsilon c_0 & 0 & 1 \end{pmatrix} \geq 0,$$

which implies the last two inequalities. □

We saw above that $\max(|c_n|, |d_n|) \geq |c|/2$. Together with the third inequality this implies already $1 - c^2 \geq c/2$, $\mu(f)^2 - \mu(f)/2 - 1 \geq 0$, and hence

$$\mu(f) \geq \frac{1 + \sqrt{17}}{4} = 1.280\ldots.$$

We can assume without loss of generality that $\mu(f) \leq \frac{4}{3}$. We set

$$f(0)\frac{f}{f^*} = 1 + a_k z^k + a_l z^l + O(z^{l+1}),$$

where $k < l$ and $a_k, a_l \neq 0$. By multiplying B by $\pm f(0)$ and $\widehat{B}$ by ± 1 we can then assume that $B(0), \widehat{B}(0) > 0$ and that

$$B = (1 + a_k z^k + a_l z^l + \cdots)\widehat{B}.$$

In particular, we have

$$c := c_0 = d_0 = \mu(f)^{-1},$$

and furthermore

$$c_j = d_j \qquad (0 \le j < k) \tag{1.4}$$
$$c_k = d_k + c_0 a_k \tag{1.5}$$
$$c_{k+1} = d_{k+1} + d_1 a_k \tag{1.6}$$
$$c_{l-1} = d_{l-1} + d_{l-k-1} a_k \tag{1.7}$$
$$c_l = d_l + d_{l-k} a_k + c_0 a_l \tag{1.8}$$

As a first consequence we note

$$|a_k| = 1 \tag{1.9}$$
$$|a_l| = 1 \tag{1.10}$$
$$|c_k| + |d_k| = c \tag{1.11}$$

Indeed, if $|a_k| \ge 2$, then, by (1.5), we would have $\max(|c_k|, |d_k|) \ge c$. Hence, by the lemma $1 - c^2 \ge c$, which contradicts $c \ge \frac{4}{3}$.

Similarly, if $|a_l| \ge 2$, then, by (1.8), $\max(|c_l|, |d_l|, |d_{l-k}|) \ge \frac{2}{3}c$, and hence, by the lemma, $1 - c^2 \ge \frac{2}{3}c$. Again this contradicts $c \ge \frac{4}{3}$.

Finally, by (1.5) $|c_k| + |d_k| \ge |c_k - d_k| \ge c$. If the inequality were strict, then $c = |c_k| - |d_k|$ or $c = |d_k| - |c_k|$, in any case, $\max(|c_k|, |d_k|) \ge c$, which is impossible as we have already seen.

Case $2k \le l$: We can assume that $2k < l$. Otherwise we interchange f and f^*. Namely, using

$$(1 + a_k x^k + a_{2k} x^{2k} + \cdots)^{-1} = 1 - a_k x^k + (a_k^2 - a_{2k}) x^{2k} + \cdots,$$

we see that then $a_k^2 - a_{2k} = 0$ (since otherwise this would be 2 by (1.9), (1.10)).

We now apply (1.3) to obtain

$$-(1 - c^2 - \frac{c_k^2}{1 + c_0}) \le c_{2k} \le 1 - c^2 - \frac{c_k^2}{1 - c_0}$$
$$-(1 - d^2 - \frac{d_k^2}{1 - d_0}) \le -d_{2k} \le 1 - d^2 - \frac{d_k^2}{1 + d_0}).$$

Adding both inequalities gives (on using also (1.5))

$$-2(1 - c^2) + \frac{d_k^2}{1 - d_0} + \frac{c_k^2}{1 + c_0} \le c_{2k} - d_{2k} = d_k a_k \tag{1.12}$$
$$\le 2(1 - c^2) - \frac{d_k^2}{1 + d_0} - \frac{c_k^2}{1 - c_0}. \tag{1.13}$$

Using (1.11) this gives

$$|d_k| \le \max(H(|d_k|), H(|c_k|)),$$

where we use

$$H(x) := 2(1-c^2) - \Big(\frac{x^2}{1+c} + \frac{(c-x)^2}{1-c}\Big).$$

But $1-c^2 \ge |d_k| \ge c - |c_k| \ge c + c^2 - 1$ by (1.3),(1.7) and (1.2), respectively, and the same holds true with c_k and d_k interchanged. Hence if we set $I = [c^2+c-1, 1-c^2]$, then we find

$$c^2 + c - 1 \le \max_{x\in I} H(x).$$

But $H(x)$ takes its maximum in $x = \frac{1+c}{2}$. Since $c \ge \frac{3}{4}$ we have $\frac{1+c}{2} \ge 1-c^2$ (since the latter is equivalent to $c \notin]-1, \frac{1}{2}[$). Hence $H(x)$ is increasing on I, and thus

$$c^2 + c - 1 \le H(1-c^2) = 2(1-c^2) - \frac{(1-c^2)^2}{1+c} - \frac{(c^2+c-1)^2}{1-c},$$

i.e. $-c^3 - c^2 + 1 \ge 0$. This gives finally $\mu(f)^3 - \mu(f) - 1 \ge 0$, which means that $\mu(f)$ is to the right of the real root θ of $x^3 - x - 1 = 0$.
Case $l < 2k$**:** We may assume $a_k = \pm 1$ (otherwise interchange f and f^*). By (1.1) we have, for all $\beta, \gamma \in \mathbb{R}$,

$$\begin{gathered} c^2 + (c_{l-k} + \gamma c)^2 + (c_k + \gamma c_{2k-l} - c)^2 + (c_l + \gamma c_k - c_{l-k} + \beta c)^2, \\ c^2 + (-d_{l-k} - \gamma c)^2 + (-d_k - \gamma d_{2k-l} - c)^2 + (-d_l - \gamma d_k - d_{l-k} + \beta c)^2 \\ \le 2 + \gamma^2 + \beta^2. \end{gathered}$$

We add these two inequalities, use $\frac{a^2+b^2}{2} \ge \left(\frac{a+b}{2}\right)^2$ and $c_j = d_j$ for $1 \le j < k$, and set $x = c_{l-k} = d_{l-k}$ to obtain

$$c^2 + \big(x+\gamma c\big)^2 + \big(\frac{c_k - d_k}{2} - c\big)^2 + \big(\frac{c_l - d_l}{2} + \gamma\frac{c_k - d_k}{2} - x + \beta c\big)^2 \le 2 + \gamma^2 + \beta^2.$$

By (1.5) and (1.8), using $a_k = +1$, this can be rewritten as

$$\frac{5}{4}c^2 + (x+\gamma c)^2 + \big(\frac{x + ca_l}{2} + \gamma\frac{c}{2} - x + \beta c\big)^2 \le 2 + \gamma^2 + \beta^2.$$

Replacing x by $-a_l x$, β by $a_l\beta$ and γ by $-a_l\gamma$ we get

$$\frac{5}{4}c^2 + (x+\gamma c)^2 + \big(\frac{c + x - \gamma c}{2} + \beta c\big)^2 \le 2 + \gamma^2 + \beta^2.$$

If we view the difference of the right hand side and the left hand side as quadratic polynomial in β, then the inequality states that its discriminant is ≤ 0, i.e. (using $1-c^2>0$) that

$$\frac{5}{4}c^2+(x+\gamma c)^2+\frac{(c+x-\gamma c)^2}{4(1-c^2)}\leq 2+\gamma^2.$$

Again, viewing the difference of both sides as polynomial in γ, we obtain that its discriminant is ≤ 0. Thus (using that the coefficient of γ^2 is positive since $c<4/(1+\sqrt{17})$, as follows from $1-c^2\geq c/2$) we have

$$\frac{5}{4}c^2+x^2+\frac{(c-x)^2}{4(1-c^2)}+\frac{(2xc-\frac{c(c+x)}{2(1-c^2)})^2}{4(1-c^2-\frac{c^2}{4(1-c^2)})}\leq 2$$

Now, again, since $c<4/(1+\sqrt{17})$, the left hand side minus 2 viewed as polynomial in x has positive leading term. Since it is ≤ 0 for at least one x it has a real root, hence non negative discriminant. By a straight forward calculation this yields $40c^4-93c^2+40\geq 0$, or, in terms of $\mu(f)$, finally

$$\mu(f)^4-\frac{93}{40}\mu(f)+1\geq 0.$$

This implies

$$\mu(f)\geq 1.3248\cdots>\theta=1.3247\ldots,$$

and proves thus Smyth's theorem.

1.10 Remarks

Parts of the proof of Smyth theorem can already be found in [Sieg], where it was proved that the real root of $x^3-x-1=0$ is the smallest Pisot number. The sharpest result in the direction of the general Lehmer conjecture is due to Dobrowolski, Cantor and Straus and Louboutin [Dobr], [Loub] which states that there exists a constant $\gamma>0$ such that

$$H(\alpha)^n\geq 1+\gamma\left(\frac{\log\log n}{\log n}\right)^3$$

for all $\alpha\neq 0$ of degree n which are not equal to a root of unity. The presentation chosen in this chapter, which led from the easy proof of Schinzel's and Zhang's theorem to Langevin's theorem, does not correspond to the correct chronological order of their discovery. Indeed, Langevin's theorem was

published in 1985, and the former proofs were found almost ten years later. However, they are all three based on what is sometimes called the resultant method, which is already more or less explicitly used by Schinzel [Schi]. Zhang's theorem (along the lines of Zagier's proof) has been generalized by Beukers, Schieckewei, Schmidt, Wirsing, Zagier for obtaining absolutely lower bounds for heights along hypersurfaces; see [BeZa] and the references therein, and see the next section for a theorem of this kind. In particular, as corollary of the main result in [BeZa] one obtains a part of Smyth theorem: If the trace of α is integral and different from $1/\alpha$ (and α is thus not self-reciprocal), then $H(\alpha)^n \geq \sqrt{\frac{1+\sqrt{5}}{2}}$, where n is the degree of α. Another possible generalization of Zhang's theorem was investigated in [Smy2]. Here Zhang's theorem is interpreted as giving an absolute lower bound for the Mahler measure of polynomials in $X(X-1)$, and the paper generalizes this result to polynomials of the form $p(T(x))$, where $T(X) \in \mathbb{Z}[X]$ is of degree $n \geq 2$, divisible by X, but $\neq \pm X^n$. This point of view is also taken up in [Doch]

Part 2

Heights on Elliptic Curves

So far we have discussed heights of algebraic numbers. One may view this theory as theory of heights on the curve $\mathbb{P}^1$. Indeed, for a point $P = [x : y] \in \mathbb{P}^1(K)$, where K is a number field, define

$$H(P) = \prod_{v \in P_K} \max(|x|_v, |y|_v)^{1/[K:\mathbb{Q}]}.$$

By the product formula this does not depend on the choice of projective coordinates of P, and if we identify $\alpha \in K$ with the point $P := [\alpha : 1] \in \mathbb{P}^1(K)$, then $H(P) = H(\alpha)$. In this section we now discuss heights on curves of genus 1, which may be viewed as a natural step after the genus 0 case discussed before.

However, before going into this theory, we shall reinterprete Zhang's theorem. This theorem is in a sense on the boundary between the theory of heights of algebraic numbers and heights on general curves. Next, we have to discuss shortly heights on projective space, since some of the general results about such heights are needed for the theory of heights on elliptic curves.

2.1 Heights on affine plane curves

In this section we generalize the proof of Zhang's theorem as given in [Zagi]. For this we restate Zhang's theorem as a theorem about heights on affine, possibly reducible, plane algebraic curves defined over $\mathbb{Q}$. By such a curve we understand the set C of solutions (x, y) of an equation $F(x, y) = 0$, where $F \in \mathbb{Q}[x, y]$, and F is not constant. We use C^* for the curve defined by

$$F^*(x, y) := x^m y^n F(1/x, 1/y),$$

where m and n are the degrees of $F(x, y)$ in x and y respectively.

Zhang's theorem may be restated by saying that $H(\alpha)H(\beta) > C$ for all (α, β) on the curve $x + y = 1$. This suggests of thinking of $H(\alpha)H(\beta)$ as height of the point $P = (\alpha, \beta)$, and then Zhang's theorem says that the heights of the points on the line $x + y = 1$ are bounded to below. Or it may also be thought of saying that the heights of two algebraic numbers satisfying an algebraic (here linear) relation can not be both arbitrary small. It is not hard to generalize Zhang's theorem as follows:

Theorem 2.1. *Let C be an affine plane curve defined over $\mathbb{Q}$ such that C intersects C^* in only finitely many points. Then there is a constant $A > 1$ such that*

$$H(\alpha)H(\beta) \geq A$$

for all pairs of algebraic numbers (α, β) on C such that $\alpha, \beta \neq 0$ and (α, β) is not an intersection point of C with C^*.

Proof. Let $G(x, y)$ be a polynomial which vanishes at the intersection points of C with C^*. For real $s > 0$ set

$$\gamma_s(z, w) = |z|^{\frac{1}{2}}|w|^{\frac{1}{2}}|G(z, w)G(1/z, 1/w)|^s.$$

We show that for every sufficiently small $s > 0$ there is a constant $A = A_s > 1$ (depending on s) such that

$$\max(1, |z|)\max(1, |w|) \geq A_s\gamma_s(z, w)$$

for all (z, w) on the truncated curve $D := C \cup C^*$, which is defined by $FF^* = 0$, if, say C is defined by $F = 0$.

Since both sides of the desired inequality have the same invariance under $z \mapsto 1/z$ and under $w \mapsto 1/w$, it suffices to prove the estimate for all points on the curve $D_0 := D \cap (\mathbb{D} \times \mathbb{D})$, where $\mathbb{D}$ is the disk $|z| \leq 1$. Hence we have to show that for all $l \gg 0$

$$|z|^l|w|^l|G(z, w)G^*(z, w)| < 1$$

on D_0.

For proving this note that the number of points (z, w) of D_0 with $|zw| = 1$ is finite, and that, for any such point, one has $G(z, w) = 0$. Indeed, if (z, w) is such a point, then $(1/z, 1/w) = (\overline{z}, \overline{w})$, and hence, using that F and F^* have real coefficients, $F(z, w) = 0$ implies $F^*(z, w) = 0$ and vice versa, i.e. (z, w) is an intersection point of C and C^*. Hence, there is an open neighborhood U of all these points such that the last inequality holds true on U. Since $D_0 \setminus U$ is compact there exists a constant $R < 1$ such that $|zw| \leq R$ on

$D_0 \setminus U$. Moreover, $|GG^*| < a$ on $D_0 \setminus U$ with a suitable constant a. Thus, if l satisfies $R^{l'}a < 1$, where $l' = (l + \max(m, n))/2$ with m and n denoting the degree of G in x and y respectively, then the desired inequality holds true on all of D_0.

To finish the proof we proceed exactly as in the proof of Zhang's theorem. Let (α, β) is a pair of algebraic numbers on C, say of degree d and e and with normalized minimal polynomials $f = ax^d + \cdots$ and $g = bx^e + \cdots$, respectively. Then, for all sufficiently small s, we have

$$\begin{aligned} H(\alpha)^d H(\beta)^e &= |ab| \prod_{\alpha',\beta'} \max(1, |\alpha'|) \max(1, |\beta'|) \\ &\geq A_s^{d+e} |a|^{\frac{1}{2}-sm} |b|^{\frac{1}{2}-sn} |f(0)|^{\frac{1}{2}-sm^*} |g(0)|^{\frac{1}{2}-sn^*} \cdot \\ &\qquad \cdot \prod_{\alpha',\beta'} |a^{m+m^*} b^{n+n*} (GG^*)(\alpha', \beta')|^s, \end{aligned}$$

where m^* and n^* are the degrees of G^* in x and y, respectively, and where α' and β' are running through the conjugates of α and β. If

$$s \max(m, m^*, n, n^*) < 1/2$$

and if G has integral coefficients, then the right hand side is A^{d+e} times positive powers of nonnegative integers. Hence it is bounded to below by $\geq A^{m+n}$, unless $\alpha\beta(GG^*)(\alpha, \beta) = 0$.

We finally assume that we have chosen G such that the curves $D : GG^* = 0$ and C intersect in only finitely many points. If (α, β) is on C, but not on C^*, then α and β are not both roots of unity, and hence $H(\alpha)H(\beta) > 1$ by Kronecker's theorem. Thus, replacing A_s by the minimum of A_s and the $H(\alpha)H(\beta)$, where (α, β) runs through the finitely many points of C and $D : GG^* = 0$, but not on C^*, finally gives the desired estimate.

It remains to ensure the existence of a G with integral coefficients, vanishing on $C \cap C^*$, but such that $D : GG^* = 0$ and C have only finitely many points in common. Indeed, such polynomials exist. We can e.g. choose through each intersection P point of C and C^* a line $L_P(x, y) = 0$, such that neither this line, nor one of its finitely many conjugate lines $L_{P,j}(x, y) = 0$ lie on C or C^*. (A conjugate line is one whose defining equation is obtained by applying to the coefficients of L_P a Galois substitution of $\overline{\mathbb{Q}}$.) Then $G := \prod_{P,j} L_{P,j}$ has the desired properties. □

As already mentioned before, we had $C : x + y = 1$ in Zhang's theorem. Thus $C^* : x + y = xy$. The intersection points of C and C^* are $\rho = \frac{1+\sqrt{-3}}{2}$ and its complex conjugate. If we take for the G used in the preceding proof

$$G = (\rho x - \overline{\rho} y)(\overline{\rho} x - \rho y) = x^2 - xy + y^2,$$

then $G^* = x^2 - xy + y^2$, and

$$\gamma_s(1, 1-z) = |z|^{\frac{1}{2}}|1-z|^{\frac{1}{2}}\Big(\frac{|z^2-z+1|}{|z||1-z|}\Big)^{2s}.$$

This is the function we actually used in our original proof with $s = 1/4\sqrt{5}$.

2.2 Heights on projective space

For a point P in $\mathbb{P}^n$, say with projective coordinates $[x_0 : \cdots : x_n]$ in a number field K, we define its height $H_K(P)$ relative to K and its absolute height $H(P)$ by

$$H_K(P) = \prod_{v \in P_K} \max_{0 \le j \le n} |x_j|_v, \qquad H(P) = H_K(P)^{1/[K:\mathbb{Q}]}.$$

By the product formula $\prod_v |t|_v = 1$ ($t \in K$) this is well defined (see the proof of 1.10), and by the compatibility relations $H(P)$ does not depend on the choice of the field K.

If $P \in \mathbb{P}^n(\mathbb{Q})$, then we may choose the projective coordinates x_j in $\mathbb{Z}$ and such that $\gcd(x_0, \ldots, x_n) = 1$. But then, for each non-archimedian v, we have $|x_j|_v \le 1$ for all j and $|x_j|_v = 1$ for at least one j, and hence $H(P)$ is given by the more intuitive formula

$$H(P) = \max_j |x_j|$$

with the usual archimedean absolute values $|x_j|$.

If $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n(\mathbb{Q})$ and, say, $x_j \neq 0$, then the minimal field of definition $\mathbb{Q}(P)$ if P is defined as

$$\mathbb{Q}(P) = \mathbb{Q}\big(\frac{x_0}{x_j}, \ldots, \frac{x_n}{x_j}\big).$$

This does not depend on the choice of x_j.

We shall need two basic properties of the absolute height.

Theorem 2.2. *For each constant C and for each integer d the set*

$$\{P \in \mathbb{P}^n \mid H(P) \le C,\ [\mathbb{Q}(P) : \mathbb{Q}] \le d\}$$

is finite.

Proof. Indeed, one has for any $P \in \mathbb{P}^n$, say $P = [x_0 : \cdots : x_n]$ with at least one $x_j = 1$ and with $K = \mathbb{Q}(P)$,

$$H_K(P) = \prod_{v \in P_K} \max_j |x_j|_v \geq \max_j \prod_v \max(1, |x_j|_v) = \max_j H_K(x_j).$$

If $[\mathbb{Q}(P) : \mathbb{Q}] \leq d$ then we also have $[\mathbb{Q}(x_j) : \mathbb{Q}] \leq d$ for all j. Thus the theorem follows from the special case $n = 1$, which we proved in section 1.4. □

By a morphism

$$F : \mathbb{P}^n \to \mathbb{P}^m$$

of degree d we understand a map of the form

$$F(P) = [f_0(x_0, \ldots, x_n), \ldots, f_n(x_0, \ldots, x_n)], \qquad (P = [x_0 : \cdots : x_n]),$$

where the f_j are homogeneous polynomials of degree d and with coefficients in $\overline{\mathbb{Q}}$. In particular, for such a set of polynomials f_j, one has $f_j(x_0, \ldots, x_n) = 0$ for all $0 \leq j \leq m$ if and only if $x_0 = x_1 = \cdots = x_m = 0$.

Theorem 2.3. *Let $F : \mathbb{P}^n \to \mathbb{P}^m$ be a morphism of degree d. Then there exist constants $C_1, C_2 > 0$ such that*

$$C_1 H(P)^d \leq H(F(P)) \leq C_2 H(P)^d$$

for all $P \in \mathbb{P}^n$.

Proof. Let $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n(K)$. For a place $v \in P_K$ we set $\varepsilon(v) = 1$ if v is archimedean, and $\varepsilon(v) = 0$ otherwise. Using this symbol we have, for all v and all points $a_j \in K$,

$$|a_1 + \cdots + a_r|_v \leq r^{\varepsilon(v)} \max_{1 \leq j \leq r} |a_j|_v.$$

Moreover, we use $H_v(P) = \max_j |x_j|_v$, thus $H_K = \prod_v H_v$.

Accordingly we have (using $f_{j,k}$ for the $\binom{n+d}{d}$ coefficients of f_j)

$$\begin{aligned} H_v(F(P)) &= \max_j |f_j([x_0 : \cdots : x_n])|_v \\ &\leq \binom{n+d}{d}^{\varepsilon(v)} \big(\max_{j,k} |f_{j,k}|_v\big)\big(\max_j |x_j|_v^d\big). \end{aligned}$$

This yields the second inequality.

For the first one we need the Hilbert Nullstellensatz (see any text book an algebraic geometry). In our case it asserts that, for any polynomial f which

vanishes at the common zeroes of all the f_j, some positive integral power f^r lies in the ideal I generated by the f_j in the ring $\overline{\mathbb{Q}}[X_0, \ldots, X_n]$. Now, the is the only common zero of the f_j is the point 0, and hence, for a suitable integer r the polynomials $X_0^r, \ldots, X_n^r$ lie in I. In other words $X^r = \sum_j P_{k,j} f_j$ with suitable $P_{k,j} \in \overline{\mathbb{Q}}[X_0, \ldots, X_n]$. These identities remain valid if we replace the $P_{k,j}$ by their $r-d$th homogeneous component, and hence we may assume that the $P_{k,j}$ are homogeneous of degree $r-d$. Enlarging K if necessary, we may furthermore assume that the $P_{k,j}$ have coefficients in K. Then, similar to the reasoning above, we have

$$\begin{aligned} H_v(P)^r &= \max_k |\sum_j (P_{k,j} f_j)(x_0, \ldots, x_n)|_v \\ &\le (m+1)^{\varepsilon(v)} \big(\max_{k,j} |P_{k,j}(x_0, \ldots, x_n)|_v \big) \big(\max_j |f_j(x_0, \ldots, x_n)|_v \big) \\ &\le (m+1)^{\varepsilon(v)} \binom{n+r-d}{r-d}^{\varepsilon(v)} C\, H_v(P)^{r-d} H_v(F(P)), \end{aligned}$$

where C is the maximum of the v-adic valuations of the coefficients of all the $P_{k,j}$. This implies the first estimate. □

2.3 Plane curves as diophantine equations

Everybody knows how to compute $L(\mathbb{Q})$ for a line $L/\mathbb{Q}$ in the projective plane $\mathbb{P}^2$. It is also not difficult to compute $C(\mathbb{Q})$ for a projective plane curve $C/\mathbb{Q}$ of degree 2. Let us consider, to have a concrete example, the circle C which is given in affine coordinates by $C : x^2 + y^2 = 1$. We fix a point $O \in C(\mathbb{Q})$. Then, for $P \in C(\mathbb{Q})$, the line L_P through O and P is defined over $\mathbb{Q}$. If $P = (x_1, y_1)$, then L_P is given by the equation

$$y = \frac{y_1}{x_1 - 1}(x - 1).$$

Conversely, if L is a line through O, then L intersects C in exactly two points, in O and in a second point $P = (x_1, y_1)$. (If $P = O$ then L_P is the tangent to C at O and vice versa.) If L is defined over $\mathbb{Q}$, then so is P. Indeed, if L is given by $y = \lambda(x-1)$, then x_1 is a solution of the quadratic equation over $\mathbb{Q}$ obtained by replacing y in $x^2 + y^2 = 1$ by $\lambda x + \mu$. Since $x = 1$ is also a solution, x_1 is necessarily rational, and so is $y_1 = \lambda x_1 + \mu$. Working out the details one finds $x_1^2 + \lambda^2 (x_1 - 1)^2 = 1$, i.e.

$$x_1 = \frac{\lambda^2 - 1}{\lambda^2 + 1}, \quad y_1 = \frac{-2\lambda}{\lambda^2 + 1}.$$

In general, if $C/\mathbb{Q}$ is an irreducible smooth projective plane curve of degree 2, and $O = (x_0, y_0) \in C(\mathbb{Q})$, then one can easily verify that the map

$$C \to \mathbb{P}^1, \quad P = (x_1, y_1) \mapsto \frac{y_1 - y_0}{x_1 - x_0} = \text{slope of the line through } O \text{ and } P$$

is an isomorphism defined over $\mathbb{Q}$ and mapping $C(\mathbb{Q})$ onto $\mathbb{P}^1(\mathbb{Q})$. The above method of determining $C(\mathbb{Q})$ is effective, apart from the fact that we have to find at least one $O \in C(\mathbb{Q})$ to start with.

We now turn to cubic curves defined over $\mathbb{Q}$. Here the situation has still some similarities with the quadratic case, though there are also much more complications. Again we start with the idea of reducing to algebraic equations in one variable by intersecting with lines. However, if we intersect a cubic curve $C/\mathbb{Q}$ with a line, then there will be in general three intersection points. But still, if the line is defined over $\mathbb{Q}$ and two of the intersection points are in $C(\mathbb{Q})$, then the third one belongs to $C(\mathbb{Q})$ too. However, one can make an even stronger statement.

To explain this we restrict for the following to elliptic curves in Weierstrass form defined over a number field K. By such a curve we understand a cubic curve E which, in affine coordinates, is given by an equation of the form

$$E : y^2 = x^3 + Ax + B,$$

where A, B are elements of K, and where we assume that the the polynomial in x on the right has no multiple roots, i.e. that its discriminant

$$\Delta_E := -4A^3 - 27B^2 \neq 0.$$

Such a curve has exactly one point O on the line at infinity, which in homogeneous coordinates is given by

$$O = [0 : 1 : 0].$$

The condition $\Delta_E \neq 0$ ensures that E is a non-singular curve. The restriction to such curves is not a serious one, since any non-singular plane cubic curve is isomorphic to a curve in Weierstrass form (see the next section for details).

If for $P = (x, y) \in E$ we set $-P := (x, -y)$, and if we define a binary operation $+$ on E by letting $P_1 + P_2$ the unique point P such that P_1, P_2 and $-P$ (counting multiplicities) are the intersection points of E with a line, then E becomes a group (for details and a proof of this see the next section). Clearly, the point 0 at infinity is the neutral element of E (it is an inflection point), and if α is a root of $X^3 + AX + B$, then $(\alpha, 0)$ is a point of order 2. Finally, if E is defined over K, then $E(K)$ is a subgroup of E. This follows

easily by looking at the equations expressing the affine coordinates of P_1+P_2 in terms of those of P_1, P_2 (again, see the next section for details).

Assume now, to come back to diophantine equations over $\mathbb{Q}$ and to show the idea for the general theory developed in a moment, that E is of the special form

$$E : y^2 = (x-a)(x-b)(x-c)$$

with pairwise different integers a, b, c. Clearly the question is when, for a rational number x, the product $(x-a)(x-b)(x-c)$ is a square in $\mathbb{Q}$. To analyze this we introduce the map

$$\phi : E(\mathbb{Q}) \to G := (\mathbb{Q}^*/\mathbb{Q}^{*2})^2,$$

$$P \mapsto \begin{cases} \big((x-a)\mathbb{Q}^{*2}, (x-b)\mathbb{Q}^{*2}\big) & \text{if } x \neq a, b,\ P \neq 0 \\ 1 & \text{if } P = 0 \\ \big((x-b)(x-c)\mathbb{Q}^{*2}, (x-b)\mathbb{Q}^{*2}\big) & \text{if } x = a \\ \big((x-a)\mathbb{Q}^{*2}, (x-a)(x-c)\mathbb{Q}^{*2}\big) & \text{if } x = b \end{cases},$$

where (x, y) are the affine coordinates of P if $P \neq 0$.

Lemma 2.1. *The map ϕ is a group homomorphism with kernel $2E(\mathbb{Q})$.*

Proof. For showing that ϕ is a group homomorphism it clearly suffices to show that $\phi(P_1)\phi(P_2)\phi(P_3) = 1$ if $P_1 + P_2 + P_3 = 0$. This is trivial if one of he P_j is 0. Otherwise the P_j lie on a line $y = \lambda x + \mu$ with $\lambda, \mu \in\in \mathbb{Q}$, $\lambda \neq 0$. Hence, if we set $P_j = (x_j, y_j)$, then the x_j are the solutions of

$$(x-a)(x-b)(x-c) - (\lambda x + \mu)^2 = 0,$$

Hence we have

$$(x-a)(x-b)(x-c) - (\lambda x + \mu)^2 = (x-x_1)(x-x_2)(x-x_3).$$

In particular, considering this equation for $x = a$, $x = b$ and $x = c$, respectively, we observe that

$$(a-x_1)(a-x_2)(a-x_3), (b-x_1)(b-x_2)(b-x_3), (c-x_1)(c-x_2)(c-x_3) \in \mathbb{Q}^2.$$

From this one easily obtains $\phi(P_1)\phi(P_2)\phi(P_3) = 1$.

Clearly $2E(\mathbb{Q})$ is mapped to 1 since $\mathbb{Q}^*/\mathbb{Q}^{*2}$ has exponent 2. Conversely, assume that $\phi(P) = 1$. ... to be completed later. □

For $v \in P_{\mathbb{Q}}$, let G_v denote the subgroup (of order 2) in $\mathbb{Q}^*/\mathbb{Q}^*$ generated by $p\mathbb{Q}^{*2}$ if v is non-archimedean belonging to the prime number p, and by $(-1)\mathbb{Q}^{*2}$, if v is archimedean. Clearly, $\mathbb{Q}^*/\mathbb{Q}^{*2} = \sum_{v \in P_{\mathbb{Q}}} G_v$.

Lemma 2.2. *The image of ϕ is contained in*

$$\bigoplus_{p|\Delta_E \text{ or } p=\infty} G_p^2,$$

where $\Delta_E = (a-b)^2(a-c)^2(b-c)^2$ is the discriminant of E. In particular, it is finite.

Proof. Let $P \in E(\mathbb{Q})$, $P \neq 0$, say $P = (x, y)$. Let p be a prime number, and let $\phi(P) = (u\mathbb{Q}^{*2}, v\mathbb{Q}^{*2})$. We have to show that $\mathrm{ord}_p(u)$ and $\mathrm{ord}_p(v)$ are both even.

For this let p^n be the exact power of p in the prime decomposition of x.

If $n < 0$, then $x \neq 0, a, b$ and we can take $u = x - a$ and $v = x - b$. Since a, b, c are integral p^n is also the exact power of p in $x - a$, $x - b$ and $x - c$. We have accordingly $\mathrm{ord}_p(y^2) = 3n$. On the other hand $\mathrm{ord}_p(y^2)$ is even. It follows that $n = \mathrm{ord}_p(u) = \mathrm{ord}_p(v)$ is even.

If $n \geq 0$, then the order at p of each of the three numbers $x - a$, $x - b$ and $x - c$ is nonnegative. At most one of them has positive order since the difference of two of any of these divides Δ. Again, since their product is a square in $\mathbb{Q}$, this implies that the orders at p of these numbers are even. Hence if $x \neq a, b$ then u and v have even order.

The case $n \geq 0$ and $x = a$ or $x = b$ is left to the reader. □

Let R be a set of representatives for $E(\mathbb{Q})/2E(\mathbb{Q})$. By the preceding lemma R is a finite set. The set R (and possibly a finite number of additional points in $E(\mathbb{Q})$ to be explained in a moment) play the role of the point O in the case of quadrics considered above. Namely, let $P_0 \in E(\mathbb{Q})$. Then we can find an $Q \in R$ such that $P_0 = Q_0 + 2P_1$ for some $P_1 \in E(\mathbb{Q})$. Again, we find a $Q_1 \in R$ such that $P_1 = Q_1 + 2P_2$ with a suitable $P_2 \in E(\mathbb{Q})$, and so forth. Suppose that in each step the point P_j is of less complexity, say needs less digits to be described, than its predecessor P_{j-1}. Then we may hope that our descent procedure will end in the sense that P_n for some n is in a finite set S of very simple points. Hence P_0 is a linear combination of the points in $R \cup S$, which solves the problem of determining $E(\mathbb{Q})$. That the group $E(\mathbb{Q})$ is finitely generated is indeed the case for any elliptic curve over $\mathbb{Q}$; this is Mordell's theorem which we shall prove in the next sections following exactly the ideas sketched in this paragraph. The complexity of points in $E(\mathbb{Q})$ will of course be measured using a height function.

For curves of genus strictly greater than 1 the situation is completely different from the genus 0 and 1 case. Here one has Mordell's conjecture, which was proved by Faltings (for another proof, based on Faltings', but shorter, more self contained and using the theory of heights instead of arithmetic intersection theory, see [Bomb]).

Theorem 2.4. *(Mordell-Faltings) For a projective curve $C/\mathbb{Q}$ with genus ≥ 2 the set $C(\mathbb{Q})$ of its rational points is finite.*

Thus, for curves of genus 2 the problem is too find good a priori upper bounds for the height (to be properly defined in some sense) of its rational points.

2.4 Basic facts about elliptic curves

This section is still incomplete. To complete the logical thread of this second part the following topics would have to be reviewed: group law — E(K) — K(P)— Weierstrass form — action of Galois [m] is surjective — affine and projective form — K(E) = maps onto $\mathbb{P}^1$ — deg(f) — $\widetilde{E}$ — E as Jacobian of itself

2.5 Heights on elliptic curves

We fix for this section an elliptic curve E defined over a number field K, which we suppose always to be given in Weierstrass form

$$E : y^2 = x^3 + Ax + B, \qquad (A, B \in K)$$

As height $H_0(P)$ of a point $P \in E$, say with homogeneous coordinates $[x : y : z]$ in a number field L, we may consider the height of P considered as point of the projective plane $\mathbb{P}^2$, i.e.

$$H_0(P) = \prod_{v \in P_L} \max(|x|_v, |y|_v, |z|_v)^{1/[L:\mathbb{Q}]}.$$

Another possibility would be to view x as a function from E onto $\mathbb{P}^1$, and to take $H_x(P) := H(x(P))$ as the height of P, where $H(x(P))$ is the height of $x(P)$ as point of $\mathbb{P}^1$. Or, more generally, we could take any nonconstant function $f \in K(E)$, consider it as function onto $\mathbb{P}^1$ and take $H_f(P) := H(f(P)$ as height function.

However, as it turns out, all these possibilities are essentially equivalent. Also, notations become more natural if one uses additive notation, i.e. if one uses the logarithmic heights

$$h_f(P) := \frac{1}{\deg f} \log H(f(P)).$$

The reason for normalizing be the factor $1/\deg f$ will become clear in a moment.

Theorem 2.5. *Let $f, g \in K(E)$ be nonconstant functions on E. Then, for every $\varepsilon > 0$, there are constants $C_1, C_2 > 0$ such that*

$$C_1 H_f(P)^{-\varepsilon} \leq \frac{H_f(P)^{1/\deg f}}{H_g(P)^{1/\deg g}} \leq C_2 H_f(P)^{+\varepsilon}$$

for all P. Or, using logarithmic heights, for every $\varepsilon > 0$, there is a constant C such that

$$|h_f(P) - h_g(P)| \leq C + \varepsilon h_f(P)$$

for all $P \in E$.

Proof. It is easy to check that the last inequality defines an equivalence relation on the set of all functions h_f with f running through the non constant elements of $K(E)$. Hence it suffice to prove the last inequality for some specific choice of g and arbitrary f. We choose $g = x$. Moreover, we assume also that f is even. For the general case we refer the reader to [Weil] (or [Lan1], Ch. 4, Cor. 3.5). Here we call f even if $f(-P) = f(P)$. For even f the desired inequality is in fact true even for $\varepsilon = 0$.

Now f is a rational function in x and y, say $f = p(x, y)/q(x, y)$, with two polynomials $p, q \in \overline{\mathbb{Q}}[X, Y]$. Since y^2 is a polynomial in x we can even write $f = (p_1(x) + yp_2(x))/(q_1(x) + yq_2(x))$ with polynomials $p_j, q_j \in \overline{\mathbb{Q}}[X]$. Also, we may assume that the numerator and denominator are relatively prime. Then they are unique up to multiplication by constants. But then we observe, on using that y is an odd function, i.e. $y(-P) = -y(P)$, that f can only be even if $p_2 = q_2 = 0$.

Hence $f = r \circ x$, where r is the rational function $r : \mathbb{P}^1 \to \mathbb{P}^1$ given by $r(t) = p_1(t)/q_2(t)$. Since any such rational function is a morphism (in the sense explained in section 2.2), the theorem for f and $g = x$ now follows from Theorem 2.3: there exists constants $C_1, C_2 > 0$ such that

$$C_1 H(x(P))^{\deg r} \leq H((r \circ x)(P))^{\deg r} \leq C_2 H(x(P))^{\deg r}.$$

Using $\deg f = 2 \deg r$ we obtain the desired inequality. □

The heights h_f possess a striking property, which we shall use to derive a canonical height from them by a procedure analogous to the one which led us to the definition of the Mahler measure.

Theorem 2.6. *Let $f \in K(E)$. Then there is a constant C such that*

$$|h_f(P + Q) + h_f(P - Q) - (2h_f(P) + 2h_f(Q))| \leq C$$

for all $P, Q \in E$.

Proof. It suffices to prove this identity for some particular function f. The general result follows then from the preceding theorem. For f we choose the coordinate function x.

For the proof we look at the following diagram:

$$\begin{array}{ccc} E \times E & \xrightarrow{\phi} & E \times E \\ {\scriptstyle x\times x}\downarrow & & {\scriptstyle x\times x}\downarrow \\ \mathbb{P}^1 \times \mathbb{P}^1 & \longrightarrow & \mathbb{P}^1 \times \mathbb{P}^1 \\ {\scriptstyle \iota}\downarrow & & {\scriptstyle \iota}\downarrow \\ \mathbb{P}^2 & \xrightarrow{\underline{\phi}} & \mathbb{P}^2 \end{array}$$

Here we use

$$\begin{gathered} \phi : (P, Q) \mapsto (P + Q, P - Q), \\ \iota : ([x : y], [x'y']) \mapsto [yy', xy' + x'y, xx''], \\ \underline{\phi} : [a : b : c] \mapsto [b^2 - 4ac : 2b(Aa + c) + 4Ba^2 : (c - Aa)^2 - 4Bab]. \end{gathered}$$

(Here A, B are the coefficients of the Weierstrass equation defining E.) It is not completely obvious, though straightforward, to check that the diagram is commutative and that $\underline{\phi}$ is a morphism (see section 2.2).

Moreover, we leave it as an exercise to verify that there exist constants $C_1, C_2 > 0$ such that

$$C_1 \leq \frac{H(A)H(B)}{H(\iota(A, B))} \leq C_2$$

for all $P, Q \in \mathbb{P}^1$.

We use $h(A) := \log H(A)$ for $A \in \mathbb{P}^n$ and H denoting the height on $\mathbb{P}^n$. Finally, for any two real valued functions α, β on $E \times E$ we write $\alpha \approx \beta$ if $|\alpha\beta|$ is bounded on $E \times E$. We then have

$$\begin{aligned} h_x(P + Q) + h_x(P - Q) &= h(x(P + Q)) + h(x(P - Q)) \\ &\approx h(\iota(x(P + Q), x(P - Q))) \\ &= h(\underline{\phi} \circ i(x(P), x(Q))) \\ &\approx 2h(\iota(x(P), y(Q))) \approx 2h(x(P)) + 2h(x(Q)). \end{aligned}$$

Here, for the last but not least identity we used Theorem 2.3 and that the degree of $\underline{\phi}$ is 2. This proves the desired estimates. □

We now define the canonical height (or Néron-Tate) height of a point P on E by

$$h(P) = \lim_k \frac{1}{4^k} h_f(2^k P).$$

If we right n for 2^k and if we use that $E[n]$ consists of exactly n^2 points, then $h(P)$ can be viewed more suggestively as the limit of the sequence

$$\frac{1}{n^2} h_f\Big(\sum_{\substack{Q\in E\\ nQ=P}} Q\Big).$$

This is exactly the kind of formula (written additively) which we used to define the Mahler measure. In fact, it could be shown that, instead of powers of 2, we can take powers of any arbitrary nonnegative integer for obtaining the same limit.

Theorem 2.7. *The limit defining $h(P)$ converges uniformly in P. It does not depend on the choice of f. There is a constant C such that*

$$|h(P) - h_f(P)| \le C$$

for all $P \in E$.

Proof. By the last theorem, setting $Q = P$, we obtain that

$$|h_f(2P) - 4h_f(P)| \le C$$

for all P with a constant independent of P. We use this to show that $4^{-k}h(2^kP)$ is a Cauchy sequence uniformly in P. Indeed, if $m > n$ then, using the above estimate, we obtain

$$\begin{aligned} |4^{-m}h_f(2^mP) - 4^{-n}h_f(2^nP)| &= \sum_{k=n}^{m-1} |4^{-(k+1)}h_f(2^{k+1}P) - 4^{-k}h_f(2^kP)| \\ &\le \sum_{k=n}^{m-1} \frac{C}{4^{k+1}} < \frac{4C}{3\cdot 4^{n+1}}, \end{aligned}$$

which tends to zero, independent of P, for $m, n \to \infty$.

The last assertion of the theorem follows similarly by writing

$$h(P) - h_f(P) = \sum_{k=0}^{\infty} 4^{-(k+1)}h_f(2^{k+1}P) - 4^{-k}h_f(2^kP).$$

If g is another nonconstant function on E, then, for each $\varepsilon > 0$, we have $|h_f(P) - h_g(P)| \le \varepsilon h_f(P) + C$ with a constant independent of P. Replacing here P by 2^kP, dividing by 4^k and letting k tend to infinity shows that the difference of the limits of $4^{-k}h_g(2^kP)$ and $4^{-k}h_f(2^kP)$ is bounded by ε times the second limit. Since this is true for all $\varepsilon > 0$ the two limits must be equal. □

Immediately from the definition we obtain that h is an even function and that $h(0) = 0$, as follows easily on taking x for f in the definition of h. Similarly, one obtains

Theorem 2.8. *For each $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ and each $P \in E$ one has $h(P^\sigma) = h(P)$.*

Proof. This follows on writing $h(P)$ as limit of $\log H(x(nP))^{1/n}$ $(n = 2^k)$, and using $H(x(P^\sigma)) = H(x(P)^\sigma) = H(x(P))$, where the last identity is obvious from the very definition of the height H on $\mathbb{P}^2$. □

Theorem 2.9. *For each constant C and each integer d, the set*

$$\{P \in E \,|\, h(P) \leq C,\ [\mathbb{Q}(P) : \mathbb{Q}] \leq d\}$$

is finite.

Proof. Since $h_x(P) \leq h(P) + C$ for some constant C it suffices to prove the theorem with h replaced by h_x. But this is an immediate consequence of the fact that the map $P \mapsto x(P)$ is two-to-one, and the fact that there is only a finite number of algebraic numbers with height and degree below a fixed bound (see section 1.4). □

An important property is that the height is a quadratic form as is already suggested by the quasi-parallelogram law for the h_f as stated in Theorem 2.6

Theorem 2.10. *The map*

$$\langle\ ,\ \rangle : E \times E \to \mathbb{R},\ \langle P, Q\rangle := h(P + Q) - h(P) - h(Q)$$

is $\mathbb{Z}$-bilinear. In particular, one has $h(nP) = n^2 h(P)$ for all integers n and all P.

Proof. By writing in Theorem 2.6 nP and nQ for P and Q, dividing by n^2 and letting n tend to infinity we obtain the so-called parallelogram law

$$h(P + Q) + h(P - Q) = 2h(P) + 2h(Q).$$

From this the bilinearity follows by a simple algebraic manipulation. Since the pairing $\langle\ ,\ \rangle$ is symmetric it suffices to prove

$$\langle P + Q, R\rangle = \langle P, R\rangle + \langle Q, R\rangle.$$

It is straightforward to check that this is equivalent to

$$h(P+Q+R) - h(P+Q) - h(P+R+) - h(Q+R) + h(P) + h(Q) + h(R) = 0.$$

But this follows indeed from the parallelogram law (and using the evenness of h) as follows. Applying four times the parallelogram law gives

$$\begin{aligned}
&h(P+Q+R)+h(P+Q-R)-2h(P+Q)-2h(R)=0\\
&h(P-Q+R)+h(P+Q-R)-2h(Q-R)-2h(P)=0\\
&h(P-Q+R)+h(P+Q+R)-2h(P+R)-2h(Q)=0\\
&2h(Q+R)+2h(Q-R)-2h(Q)-2h(R)=0,
\end{aligned}$$

and taking the alternate sum of these four equations is exactly the desired identity..

The second assertion follows from $2h(P)=h(P)+h(-P)-h(P-P)=-\langle P,-P\rangle=\langle P,P\rangle$. □

As direct generalization of Kronecker's theorem one has

Theorem 2.11. *One has $h(P)=0$ if and only if P is a torsion point.*

Proof. By the preceding theorem we clearly have $\langle P,P\rangle=0$ if $nP=0$ for some integer $n\geq 1$. Conversely, if $h(P)=0$, then $h(nP)=0$ for all P. If L/K is a number field such that $P\in E(L)$, then $nP\in E(L)$ for all n. But the set of all $Q\in E(L)$ with $h(Q)=0$ is finite as we saw above. Hence P must have finite order. □

Since the set of points on E with height below a given bound affine coordinates in a given number field L is finite, we see that in particular $E(K)_{\rm tor}$ is finite. However, one can say much more. The theorem of Mazur [] says that, for an E defined over $\mathbb{Q}$ the subgroup $E(\mathbb{Q})_{\rm tor}$ is always isomorphic to one of a given list of fifteen abelian groups. It is conjectured that this is true for all number fields K in the following sense: For each number field K there is a constant N such that $E(K)_{\rm tor}$, for any elliptic curve E defined over K, has not more than N points. By a theorem of Manin [Mani] one knows at least that for any K and any prime number p there exists a constant N such that the p-part of $E(K)_{\rm tor}$, for any E/K, is bounded to above by N.

From the last theorem we also obtain

Theorem 2.12. *The height pairing $\langle\ ,\ \rangle$ on E factors to a non-degenerate pairing $E/E_{tor}\times E/E_{tor}\to\mathbb{R}$.*

Proof. Clearly $\langle P,Q\rangle=0$ for all Q if $nP=0$ for some $n\geq 1$. Conversely $\langle P,Q\rangle=0$ for all Q implies $h(P)=0$, and hence that P is a torsion point. □

We conclude this section with another result showing that the canonical height deserves its name.

Theorem 2.13. *Let h' be a real valued function on E which satisfies the two following properties:*

1. *There exists an integer $n \geq 2$ such that $h'(nP) = n^2 h'(P)$ for all $P \in E$.*
2. *There exists a function $f \in E(K)$ and a constant C such that $|h(P) - h_f(P)| \leq C$ for all $P \in E$.*

Then $h' = h$.

Proof. From the second assumption we see that $|h'(P) - h(P)| \leq C'$ for all P with a suitable constant C' (not depending on P). But then from the first assumption $h'(n^k P) = n^{2k} h'(P)$ for all $k \geq 0$, and hence

$$|h'(P) - h(P)| = \frac{1}{n^{2k}} |h'(n^k P) - h(n^k P)| =\leq \frac{C'}{n^{2k}}$$

for all k, whence, for $k \to \infty$, we obtain $h'(P) = h(P)$. □

2.6 Infinite descent on elliptic curves

In this section, using the theory of heights on elliptic curves, we can finally make precise the infinite descent procedure described at the the end of section 2.3. For this let E be a given elliptic curve defined over the number field K. We shall prove in the next section, that $E(K)/mE(K)$ is a finite group for each integer $m \geq 2$. As already indicated before this, together with the infinite descent procedure, implies that $E(K)$ is a finitely generated group. The descent procedure is effective, i.e. it shows how to calculate generators for $E(K)$ (provided we we can compute a set of representatives for the quotient $E(K)/mE(K)$).

Let $\mathfrak{R}$ be a system of representatives for $E(K)/mE(K)$ for a fixed $m \geq 1$. For this set of representatives let

$$C := 2 \max\{h(P) \mid P \in \mathfrak{R}\}.$$

We then have, for all $P \in E$ and all $R \in \mathfrak{R}$.

$$h(P + R) = 2h(P) - h(P - R) + 2h(R) \leq 2h(P) + C.$$

Let now $P \in E(K)$. We define a sequence of points $P_l \in E$ and $Q_l \in \mathfrak{R}$ by $P_0 = P$ and for $l \geq 1$

$$mP_l = P_{l-1} - Q_{l-1}.$$

Then

$$\begin{aligned} h(P_l) &\leq \frac{1}{m^2}(2h(P_{l-1}) + C) \\ &\leq \frac{2^l}{m^{2l}}h(P) + C(\frac{1}{m^2} + \frac{2}{m^4} + \frac{4}{m^6} + \cdots + \frac{2^{l-1}}{m^{2(l-1)}}) \\ &\leq \frac{2^l}{n^{2l}}h(P) + \frac{C}{m^2-2}. \end{aligned}$$

Finally, let $\mathfrak{R}_0$ be the set of all $Q \in E(K)$ with $h(Q) \leq C/(m^2-2)$. This is a finite set. Since the set of all Q with $h(Q) \leq C/(m^2-2) + .1$ is also finite, we can find a $\varepsilon > 0$ such that $\mathfrak{R}_0$ coincides with the set of all $Q \in E(K)$ with $h(Q) \leq C/(m^2-2) + \varepsilon$.

But then we conclude that $P_l \in \mathfrak{R}_0$, if l is large enough. In other words the set $\mathfrak{R} \cup \mathfrak{R}_0$ is a set of generators for $E(K)$. The set $\mathfrak{R}_0$ can be calculated by a systematic search.

2.7 The Mordell-Weil theorem

Again, throughout this section, E denotes an elliptic curve defined over a number field K. Moreover we fix an integer $m > 0$. The purpose of this section is to prove

Theorem 2.14. *(Weak Mordell-Weil theorem) The group $E(K)/mE(K)$ is finite.*

Together with the infinite descent procedure of the last section this implies then strong Mordell-Weil theorem

Theorem 2.15. *The group $E(K)$ is finitely generated.*

The proof of the so-called weak Mordell-Weil theorem has actually nothing to do with heights, but uses what is called Kummer theory for elliptic curves. However, we include it here for the sake of completeness. The Mordell-Weil theorem was actually first proved by Mordell for the case of an elliptic curve over $\mathbb{Q}$, was before already conjectured by Poincaré, and later generalized to arbitrary K (and arbitrary abelian varieties) by Weil, based on work of Siegel who introduced the powerful tool of heights into the study of diophantine problems. The proof uses the two fundamental finiteness theorem of algebraic number theory, the finiteness of class numbers and Dirichlet's unit theorem.

We shall show first that we can enlarge K without restriction of generality.

Lemma 2.3. *Let L/K be a finite extension. If $E(L)/mE(L)$ is finite, then so is $E(K)/mE(K)$.*

Proof. Let N be the kernel of the natural map

$$E(K)/mE(K) \to E(L)/mE(L);$$

thus $N = (E(K) \cap mE(L))/mE(K)$. We have to show that N is finite.

For each $C \in N$ pick a $P \in C$, and then a $Q \in E$ such that $P = mQ$. we set

$$\lambda_C : \mathrm{Gal}(L/K) \to E[m], \quad \lambda_C(\sigma) = Q^\sigma - Q.$$

Note that indeed $\lambda_C(\sigma) \in E[m]$ since $mQ^\sigma = P^\sigma = P = mQ$. If $\lambda_C = \lambda_{C'}$, say $C' = P' + mE(K)$ with associated $mQ' = P'$, then $Q - Q'$ is invariant under all $\sigma \in \mathrm{Gal}(L/K)$, and is hence in $E(K)$. But this means $P - P' \in mE(K)$, i.e. $C = C'$. Thus the map $C \mapsto \lambda_C$ is injective; its image being finite implies the lemma. □

The proof, being a little bit puzzling at the first glance, has a very natural explication in term of Galois cohomology. We shall explain this below (see section 2.8).

In the following we can hence assume, by enlarging K if necessary, that

$$E[m] \subset E(K).$$

Note that this implies in particular the following: If $Q \in E$ is such that $mQ \in E(K)$, then $L := K(Q)$ is a Galois extension of K. Indeed, if $\sigma : L \to \mathbb{C}$ is an embedding leaving K invariant, then $L^\sigma = K(Q^\sigma)$. But $Q^\sigma \in Q + E[m]$ (since $m(Q^\sigma) = (mQ)^\sigma) = mQ$), and hence $Q^\sigma \in L$, i.e. $L^\sigma = L$.

We set

$$L := K(Q \,|\, QP \in E(K))/qquadG := \mathrm{Gal}(L/K).$$

Then L is a Galois extension of K (a priori possibly infinite). We have a map

$$E(K) \times G \to E[m],$$

given by

$$(P, \sigma) \mapsto Q^\sigma - Q,$$

where Q is any point of E such that $mQ = P$. (We recall that such a point Q always exists since multiplication by m is a nonconstant morphism.)

Note that this definition does not depend on a particular choice of Q since any two inverse images of P under multiplication by m differ by an element

of $E[m]$, which, as subset of $E(K)$, is invariant under G. The map is actually bilinear. It is linear in the right argument since

$$Q^{\sigma\tau} = (Q^\sigma - Q)^\tau + (Q^\tau - Q) = (Q^\sigma - Q) + (Q^\tau - Q),$$

where we used that $Q^\sigma - Q$ is in $E[m]$ and hence stable under G. It is obviously linear in the first argument.

The left kernel of the pairing (i.e. the subgroup of $P \in E(K)$ such that $\langle P, G\rangle = 0$) clearly contains $mE(K)$; in fact, it equals $mE(K)$. Indeed, if a $Q \in E$ with $P := mQ \in E(K)$ satisfies $Q^\sigma = Q$ for all $\sigma \in G$, then $Q \in E(K)$, i.e. $P = mQ \in mE(K)$. Thus the above pairing factors through a pairing

$$E(K)/mE(K) \times \mathrm{Gal}(L/K) \to E[m],$$

the so-called Kummer pairing, which is left non-degenerate. Or, to state this differently, the associated homomorphism

$$E(K)/mE(K) \to \mathrm{Hom}(G, E[m])$$

is injective. For proving the weak Mordell theorem it thus suffices to show that L is a finite extension of K. Hence, we start now to investigate more closely the field L.

First of all we note that the Kummer pairing is even perfect. Namely, for a fixed σ, let $Q^\sigma = Q$ for all Q with $mQ \in E(K)$. This means that σ leaves invariant L, and hence equals 1. Hence G embeds injectively into $\mathrm{Hom}(E(K)/mE(K), E[m])$, In particular, L is abelien with exponent m.

We now assume that E is given by a Weierstrass equation of the form $y^2 = x^3 + Ax + B$ with A and B being integral algebraic integers (in K). This is no restriction of generality since for each pair $A, B \in K$ we can find an integer $N > 0$ such that N^4A and N^6B are integral, and we may the consider $y^2 = x^3 + N^4AX + N^6B$, which is isomorphic to E via $(x, y) \mapsto (N^2x, N^3y)$. We use Δ for the discriminant of E, i.e.

$$\Delta = -4A^3 - 27B^2.$$

Under this assumption we then have

Lemma 2.4. *Let $\mathfrak{p}$ be a prime ideal of K not dividing the discriminant Δ of E. Then L is not ramified at $\mathfrak{p}$.*

Proof. For $P \in E(K)$ let $M = K(Q \in E \,|\, mQ = P)$. It suffices to show that M is unramified at $\mathfrak{p}$ (since L is is the compositum of all such M).

Indeed let $D_\mathfrak{p}$ be the decomposition group of $\mathfrak{p}$ i.e. the subgroup of all $\sigma \in G$ leaving invariant one prime ideal (and hence all prime ideals) $\mathfrak{P}$ of M

above $\mathfrak{p}$. Let $I_\mathfrak{p}$ be the inertia group at $\mathfrak{p}$, i.e. the subgroup of $\sigma \in D_\mathfrak{p}$ such that $x^\sigma \equiv x \bmod \mathfrak{P}$ for all $x \in O$, where O is the ring of integers of M. That M is not ramified at $\mathfrak{p}$ is equivalent to the statement that $I_\mathfrak{p}$ is trivial.

For proving this we consider, $\widetilde{E}$, the curve obtained from E by reducing modulo $\mathfrak{P}$. More precisely we consider the following: If $P = [x : y : z]$ is a point of $E(M)$, then we may assume that x, y, z are in O, and at least one homogeneous coordinate is not divisible by $\mathfrak{P}$ (indeed take any homogeneous coordinates of P in M and divide by the the one with smallest $\mathfrak{P}$-order; since the new homogeneous coordinates are $\mathfrak{P}$-integral, we can find an integer $N \neq 0$ and not divisible by $\mathfrak{P}$ such that multiplication by N yields homogeneous coordinates in O). We then set $\rho(P) := [\widetilde{x} : \widetilde{y} : \widetilde{z}]$, where the tilde denotes the class modulo $\mathfrak{P}$. This does not depend on the special choice of homogeneous coordinates. The association $P \mapsto \tilde{P}$ thus defines a map

$$E(L_P) \to \widetilde{E}(O/\mathfrak{P}) = \{[\widetilde{x} : \widetilde{y} : \widetilde{z}] \,|\, y^2z \equiv x^3 + Axz^2 + z^3 \bmod \mathfrak{P}\}.$$

It is easy to see that $E(O/\mathfrak{P})$ is a group (defined analogous to the group structure on $E(K)$), and that the reduction map is a group homomorphism. Moreover, it is a fundamental fact that the restriction of the reduction map to

$$E[m] \to \widetilde{E}(O/\mathfrak{P})$$

is injective if the discriminant of E is not divisible by $\mathfrak{P}$ (or, equivalently, not divisible by $\mathfrak{p}$). This is obvious for $m = 2$ (the case, which actually suffices to deduce the Mordell-Weil theorem). In this case $[0 : 1 : 0]$ and $[\alpha_i : 0 : 1]$ $(i = 1, 2, 3)$, with α_i denoting the roots of $f(x) := x^3 + Ax + B = 0$, are the points of $E[2]$ (recall that $\alpha_i \in K$ since $E[2] \subset E(K)$). Obviously they are in fact incongruent modulo $\mathfrak{P}$ if and only if $\mathfrak{P}$ does not divide the discriminant $\Delta = \prod_{i \neq j}(\alpha_i - \alpha_j)^2$. For general m see e.g. [Sil1], VII Proposition 3.1(b).

Let now $\sigma \in I_\mathfrak{p}$. Then

$$\rho(Q^\sigma - Q) = \rho(Q^\sigma) - \rho(Q) = 0$$

for all $Q \in E(M)$. On the other hand side, $Q^\sigma - Q \in E[m]$ for $mQ = P$. By the injectivity of the last map hence $Q^\sigma - Q = 0$. Thus σ is the identity on M, showing that $I_\mathfrak{p}$ is trivial and thus proving the theorem. □

Our information about L obtained so far suffices to prove that is is finite over K. One has the following general theorem:

Theorem 2.16. *Let L be an abelien extension of K with exponent m, and which is ramified only at a finite number of primes. Then L is a finite extension of K.*

Proof. Let S be the set of prime ideals of K, where L is ramified. By enlarging S we can assume that all prime ideals dividing m are contained in S. Moreover, by again enlarging if necessary, we can even more assume that the ring R of S-integers in K is a principal ideal domain. Indeed, let h be the class number of K, pick prime ideals $\mathfrak{p}_j$ $(1 \leq j \leq h)$ which represent the ideal classes of the class group of K, and adjoin to S all prime ideals conjugate to one of these prime ideals; clearly, $\mathfrak{p}_j^n R = R$ for all integers n (if $p \in \mathfrak{p}_j$ is a rational prime then $p^{-1} \in R$). But then, if $M \subseteq R$ is an ideal of R (and hence $M \cap O$ is an ideal of the ring of integers O of K), then, on writing $M \cap O$ as $M \cap O = \alpha \prod_j \mathfrak{p}_j^{n_j}$ with suitable integers n_j and suitable $\alpha \in O$, shows $(M \cap O)R = \alpha R$. But $(M \cap O)R = M$ (since, for each $\alpha \in M$, we can find a rational integer $N \in R$, only divisible by prime ideals in S, such that $N\alpha \in O$; but then $\alpha \in (M \cap O)R$ since $1/n \in R$).

Finally, we leave it to the reader to verify that, by adjoining mth roots of unity to K and L, one can assume without loss of generality that K contains all mth roots of unity. Or else the reader can restrict to the case of $m = 2$, where this is automatically satisfied, and which suffices for the proof Mordell-Weil theorem (and which in turn implies the weak version for arbitrary $m \geq 2$ anyway).

The main theorem of Kummer theory states that L is then a subfield of $K(\sqrt[m]{a} \,|\, a \in K)$ ([Lan2], VIII, §8) or any other reasonable text book including sections on Galois theory). Again, it is a straight-forward exercise in Galois theory to verify this statement for $m = 2$.

To begin with the proper proof of the desired theorem, we remark first of all that, for $a \in K$, the field $K(\sqrt[m]{a})$ is unramified at a prime $\mathfrak{p} \not| m$ if and only if $m | \operatorname{ord}_{\mathfrak{p}}(a)$ (Exercise).

Thus, if we let T be the set of classes $a(K^*)^m$ in $K^*/(K^*)^m$ such that $m | \operatorname{ord}_{\mathfrak{p}}(a)$ for all $\mathfrak{p} \notin S$, then

$$L \subseteq K(\sqrt[m]{a} \,|\, (aK^*)^m \in T).$$

We thus want have to show that T is finite. For this let R^* be the group of units of R. Clearly, $\operatorname{ord}_{\mathfrak{p}}(a) = 0$ for all $\mathfrak{p} \notin S$. Hence we have the natural map

$$R^* \to T.$$

We claim that it is surjective (for our special choice of R). Indeed, let $a \in K^*$ represent an element of T. Then the (fractional) R-ideal aR is the m th power of an R-ideal (consider the O-prime ideal decomposition of aO, multiply by R, and use that $\mathfrak{p}R = R$ for any $\mathfrak{p} \in S$). But R is a principal ideal domain, and hence $aR = b^m R$ for some $b \in K$, whence $a = b^m e$ for some unit $e \in R^*$,

proving the surjectivity of our map. This map factorizes then to a surjective map $R^*/(R^*)^m \to T$.

By Dirichlet's S-unit theorem R^* is finitely generated (see [Lan3] V§1), hence $R^*/(R^*)^m$, and thus T too, is finite. □

2.8 Supplements

The Kummer pairing $E(K)/mE(K) \times G \to E[m]$ can be interpreted as injection

$$\delta_E : E(K)/mE(K) \to \mathrm{H}^1(G, E[m]).$$

Here $\mathrm{H}^1(G, E[m])$ is the first cohomology group of $G := \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ acting on $E[m]$. Recall that, for any abelian group M which is a G-right module, this is the group

$$\mathrm{H}^1(G, M) = \frac{\{c : G \to M \mid c(\sigma\tau) = c(\sigma)^\tau + c(\tau)\}}{\{c : G \to M \mid \exists m \in M\, \forall \sigma \in G\ : c(\sigma) = m^\sigma - m\}}.$$

If $E[m] \subset E(K)$, as we assumed, then $\mathrm{H}^1(G, E[m])$ is nothing else than the group of homomorphisms $G \to E[m]$. Moreover, the map δ_E is nothing else as the map induced by the first connecting homomorphism, usually denoted δ, in the long exact sequence of homology groups

$$0 \to E[m](K) \to E(K) @> \times m >> E(K) @> \delta >> \mathrm{H}^1(G, E[m])$$

associated to the short exact sequence of G-modules

$$0 \to E[m] \to E @> \times m >> E \to 0.$$

Note that the map δ_E exists for arbitrary E defined over K, not just for those with $E[m] \subset E(K)$. Along these lines the given proof of the Mordell theorem may be reinterpreted and reanalyzed in terms of Galois cohomology.

The approach to the weak Mordell theorem in section 2.3 using the map $E(\mathbb{Q})/2E(\mathbb{Q}) \to (\mathbb{Q}^*/\mathbb{Q}^*)^2$, can easily be generalized to arbitrary number fields (see [Lan1], V, §1), and it can also be generalized to arbitrary m (see e.g. [Sil1], X, Theorem 1.1). It is related to the second proof as follows.

By Hilbert's theorem 90 (which states $\mathrm{H}^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \overline{Q}^*) = 0$) we know that any homomorphism

$$\alpha : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \{\pm 1\}$$

is of the form $\alpha(\sigma) = \sqrt{a}^\sigma/\sqrt{a}$ with a suitable $a \in \mathbb{Q}^*$. Hence we have an isomorphism

$$\delta_K : \mathbb{Q}^*/\mathbb{Q}^{*2} \to \mathrm{Hom}(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), \{\pm 1\}).$$

Suppose, we have a perfect pairing $e_2 : E[2] \times E[2] \to \{\pm 1\}$. Then we can define a unique map ν such that the following diagram is commutative:

$$\begin{array}{ccc} E(\mathbb{Q})/2E(\mathbb{Q}) \times E[2] & \xrightarrow{\nu} & \mathbb{Q}^*/\mathbb{Q}^{*2} \\ \downarrow{\scriptstyle \delta_R \times 1} & & \downarrow{\scriptstyle \delta_K} \\ \mathrm{Hom}(G, E[2]) \times E[2] & \xrightarrow[e_2']{} & \mathrm{Hom}(G, \{\pm 1\}) \end{array}$$

Here ($G = \mathrm{Gal}(\overline{Q}/\mathbb{Q})$,) and e_2' is the map induced by e_2, i.e. $e_2'(c, Q)(P) = e_2(c(P), Q)$ for all $P, Q \in E[2]$. Choosing a basis P_1, P_2 for $E[2]$, we then obtain an injection

$$\gamma : E(\mathbb{Q})/2E(\mathbb{Q}) \to (\mathbb{Q}^*/\mathbb{Q}^{*2})^2, \quad P \mapsto \big(\nu(P, P_1), \nu(P, P_2)\big).$$

Now, for e_2 one may take the so-called Weil pairing, which is defined as

$$e_2(P, Q) = g_Q(X + S)/g_Q(X),$$

where $g_Q \in K(E)$ is any function with divisor

$$\mathrm{div}(g) = \sum_{2R=Q} (R) - 4(O),$$

and where X is any point of E such that $g_Q(X + S)$ and $g_Q(X)$ are both different from 0 and ∞ (see any text book on (algebraic) elliptic curves) If $Q = (\alpha, 0)$ in affine coordinates, then it is not hard to check that $g_Q^2(V) = x(2V) - \alpha$ for all $V \in E$ (after suitably normalizing g_Q). Using this one can finally verify that γ is the map used in section 2.3.

2.9 Local decomposition

As in the case of algebraic numbers the canonical height on an elliptic curve has a decomposition into local contributions. In this section we describe the corresponding formulas. Again, we assume throughout that E is given by an equation of the form

$$E : y^2 = x^3 + Ax + B \qquad (A, B \in K),$$

where, as usual, K denotes a number field.

2.9.1 The Green's function of an elliptic curve

We start by describing the archimedian contributions. It is a well-known and classical fact that there exist a lattice in $\mathbb{C}$ of the form $L = \mathbb{Z}\tau + \mathbb{Z}$ with $\mathrm{Im}(\tau) > 0$ and a complex number $\lambda \neq 0$ such that the map

$$z \mapsto \begin{cases} [\wp(\tau, z) : \frac{1}{2}\wp'(\tau, z) : 1] & \text{if } z \notin L; \\ [0 : 1 : 0] & \text{if } z \in L \end{cases}$$

defines a surjective group homomorphism

$$\exp : \mathbb{C} \to E'(\mathbb{C})$$

with kernel L, where E' is the elliptic curve

$$E' : y^2 = x^3 + \lambda^4 Ax + \lambda^6 B.$$

Here $\wp(\tau, z)$, for fixed τ, as function of z, is the classical Weierstrass $\wp$ function associated to the lattice L, and $\wp'(\tau, z)$ is its derivative with respect to z. Thus, $\wp(\tau, z)$ is meromorphic in $\mathbb{C}$ with poles only in L, periodic with respect to L, and

$$\wp(\tau, z) = \frac{1}{z^2} + O(z) \qquad (z \to 0).$$

These three properties uniquely determine $\wp(\tau, z)$ (since the difference of any two such functions would be holomorphic on all of $\mathbb{C}$, periodic under L, hence bounded on $\mathbb{C}$, hence constant by the maximum principle, and finally equal to 0 because its Taylor development at $z = 0$ starts with positive powers of z). We use here the name exp because this is natural when viewing $E(\mathbb{C})$ as Lie group. Note that *exp* is continuous, when we equip $E'(\mathbb{C})$ with the natural topology (inherited from the natural quotient topology of $\mathbb{P}^2(\mathbb{C}) = (\mathbb{C}^3 \setminus \{0\})/\mathbb{C}^*$). To check this at points in L write

$$[\wp(\tau, z) : \frac{1}{2}\wp'(\tau, z) : 1] = [\frac{\wp(\tau, z)}{\frac{1}{2}\wp'(\tau, z)} : 1 : \frac{1}{\frac{1}{2}\wp'(\tau, z)}],$$

an let z tend towards a point in L.

Clearly E' and E are isomorphic (as elliptic curves over $\mathbb{C}$) via the map $(x, y) \mapsto (\lambda^2 x, \lambda^3 y)$. For the following we assume that $E = E'$ (and hence $\lambda = 1$). Of course, then A, B are not necessarily algebraic numbers.

One can even more introduce a natural structure of Riemann surface on $\mathbb{C}/L$ and on $E(\mathbb{C})$ so that the map $\mathbb{C}/L \to E(\mathbb{C})$ becomes an isomorphism of Riemann surfaces. The map exp induces an isomorphism of fields

$$\exp^* : K(E)_{\mathbb{C}} \to \mathcal{M}(L),$$

where $K(E)_{\mathbb{C}}$ is the field if algebraic functions on E, considered as algebraic curve over $\mathbb{C}$, and where $\mathcal{M}(L)$ is the field of meromorphic functions on $\mathbb{C}$ which are periodic with respect to L.

We use $\sigma(\tau, z)$ for the Weierstrass σ function associated to L. It is uniquely characterized by the fact that, as function in z, it is odd and holomorphic on $\mathbb{C}$, satisfies $\sigma(\tau, z) = z + O(z^2)$ $(z \to 0)$, and its second logarithmic derivative equals $\wp(\tau, z)$. Setting

$$q = \mathrm{e}^{2\pi i \tau}, \quad \zeta = \mathrm{e}^{2\pi i z},$$

one has the following explicit formula ([Skor], Appendix 1)

$$\sigma(\tau, z) = \mathrm{e}^{-2\pi i \frac{\eta'}{\eta}(\tau) z^2} \frac{\zeta^{1/2} - \zeta^{-1/2}}{2\pi i} \prod_{n \geq 1} \frac{(1 - q^n \zeta)(1 - q^n \zeta^{-1})}{(1 - q^n)^2},$$
$$\eta(\tau) = q^{1/24} \prod_{n \geq 1} (1 - q^n).$$

(Here η' is the ordinary derivative of η with respect to τ.) It is straightforward that the right hand side of this formula satisfies in fact all the listed properties, which proves the existence of $\sigma(\tau, z)$ (and $\wp(\tau, z)$) and, by the uniqueness, the identity in question. We leave the details to the reader (or see [Skor], Appendix 1). We cite without proof the following lemma (see)

Lemma 2.5.

$$(2\pi i)^{12} \eta^{24}(\tau) = \mathrm{disc}(x^3 + Ax + B) = -(4A^3 + 27B^2).$$

Instead of in $\sigma(\tau, z)$, we are more interested in the so-called Siegel function

$$S(z) = q^{\frac{1}{12}} \zeta^{-\frac{1}{2}} (\zeta - 1) \prod_{n \geq 1} (1 - q^n \zeta)(1 - q^n \zeta^{-1}).$$

We suppress the dependence of τ. Note that $S(z)$, considered as function of z is nothing else but $\sigma(\tau, z)$, up to multiplication by trivial factors. The important point is that $S(z)$ has a nicer transformation law under L than $\sigma(\tau, z)$. Namely, one has

Lemma 2.6.

$$S(z + 1) = -S(z), \qquad S(z + \tau) = -q^{-\frac{1}{2}} \zeta^{-1} S(z).$$

Proof. This can be verified by a straight-forward calculation. □

From this we deduce that

$$G(z) := \mathrm{e}^{-\pi \frac{y^2}{v}} |S(z)|$$

is periodic with respect to L. Here y and v are the imaginary parts of z and τ, respectively.

factors through a function on $\mathbb{C}/L$. This function is Green's function associated to E. Its important property is

Theorem 2.17. *Let $f \in K(E)_{\mathbb{C}}$, let $D = \sum_{j=1}^{r} n_j(P_j)$ ($n_j \in \mathbb{Z}$, $P_j \in E(\mathbb{C})$) its divisor, and let $P_j = \exp(z_j)$ with suitable $z_j \in \mathbb{C}$. Then there exists a constant c such that*

$$|f(\exp(z))| = c \prod_{j=1}^{r} G(z - z_j)$$

for all $z \in \mathbb{C}$.

Proof. The function

$$g(z) := f(\exp(z)) / \prod_{j=1}^{r} S(z - z_j)^{n_j}$$

is holomorphic on $\mathbb{C}$ and has no zeroes. From this it is easy to verify that

$$\tilde{g}(z) := \log g(z) + \pi \frac{1}{v} \sum_{j=1}^{r} n_j \operatorname{Im}(z - z_j)^2$$

is harmonic (though $G(z)$ itself, because of the factor $\mathrm{e}^{-\pi y/v}$, is not harmonic). Note that

$$\sum_{j=1}^{r} n_j \operatorname{Im}(z - z_j)^2$$

is harmonic since D, as divisor of a function on $K(E)_{\mathbb{C}}$, satisfies $\deg D = \sum_{j=1}^{r} n_j = 0$.

But $\tilde{g}$ is periodic with respect to L, hence bounded on $\mathbb{C}$, and thus constant by the maximum principle. □

As corollary we obtain

Corollary 2.17.1.

$$|\wp(z) - \wp(a)| = |\Delta|^{\frac{1}{6}} \frac{G(z-a)G(z+a)}{G(z)^2 G(a)^2}.$$

Proof. By the foregoing theorem we have, for fixed a and all z

$$|\wp(z) - \wp(a)| = c\frac{G(z-a)G(z+a)}{G(z)^2G(a)^2}$$

with a suitable constant c. Now, if we multiply by $|z|^2$ and let z tend to 0, then the left hand side tends to 1. For the right hand side the limit is

$$c \cdot \lim_{z\to 0} \frac{|z|^2}{G(z)^2} = c/(2\pi|\eta|^2)^2,$$

which proves the lemma. □

We finally introduce the so-called Néron function on $E(\mathbb{C})\setminus\{0\}$ by setting

$$\lambda(P) := -\log G(z),$$

where $P = \exp(z)$ (this does not depend on a particular choice of z since $G(z)$ is periodic with respect to L.)

Theorem 2.18. *The Néron function satisfies the following three conditions:*

1. *λ is continuous and is bounded on the complement of every open neighbourhood of 0.*
2. *The limit $\lim_{P\to 0}\left(\lambda(P) + \frac{1}{2}\log|x(P)|\right)$ exists and is finite.*
3. *For all $P, Q \in E(\mathbb{C})$ such that $P, Q, P+Q, P-Q \neq 0$ one has*

$$\lambda(P+Q) + \lambda(P-Q) = 2\lambda(P) + 2\lambda(Q) - \log|x(P) - x(Q)| + \frac{1}{6}\log|\Delta|.$$

Moreover, λ is the only function on $E(\mathbb{C})\setminus\{0\}$ satisfying these conditions.

Proof. Property (i) to (iii) follow immediately from the corresponding properties for $-\log G(z)$ on setting $P = \exp(z)$ and $Q = \exp(a)$, so that, in particular $x(P) = \wp(\tau, z)$.

For proving the uniqueness statement we note that the difference f of any two functions satisfying the three properties can be continuously extended to 0 (by (ii)), is hence bounded on $E(\mathbb{C})$ (by (i)), and satisfies the parallelogram law

$$f(P+Q) + f(P-Q) = 2f(P) + 2f(Q)$$

(by (iii)), by continuity even for all P, Q. In particular, $f(0) = 0$ (set $P = Q = 0$), hence $f(2P) = 4f(P)$ (set $P = Q$), and then $f(2^nP) = 4^nf(P)$ for all P and n. Letting n tend to infinity and observing that $f(2^nP)$ remains bounded it follows $f(P) = 0$. □

If $E' : y^2 = x^3 + A'x + B'$ is an elliptic curve isomorphic to E, say via $\alpha : (x, y) \mapsto (a^2x, a^3y)$, we transfer λ to a function λ' on E' by setting $\lambda' = \lambda \circ \alpha$. Note that the conditions (i) and (iii) remain literally valid for the new function λ' on E'. Indeed, if we write $x'(P)$ for the first coordinate function on E', then we have $(x \circ \alpha)(P) = a^2x'(P)$, whereas the discriminant Δ' of E' is

$$\Delta' = -(4A'^3 + 27B'^2) = -a^{12}(4A^3 + 27B^2) = a^{12}\Delta$$

(since $A' = a^4A$ and $B' = a^6B$) Hence

$$\log|x(\alpha(P)) - x(\alpha(Q))| - \frac{1}{6}|\Delta| = \log|x'(P) - x'(Q)| - \frac{1}{6}|\Delta'|.$$

Hence we can summarise by saying that on each elliptic curve E defined over $\mathbb{C}$, given by a Weierstrass equation with discriminant Δ, there is a unique function $\lambda : E(\mathbb{C}) \setminus \{0\} \to \mathbb{R}$ which satisfies properties (i) to (iii).

The condition (iii) can be replaced by another one, which is technically simpler to verify.

Theorem 2.19. *Let $E : y^2 = x^3 + Ax + B$ an elliptic curve defined over $\mathbb{C}$. Then the Néron function λ is the unique function $\lambda : E(\mathbb{C}) \setminus \{0\} \to \mathbb{R}$ which satisfies conditions (i), (ii) of Theorem 2.18 and the condition:*

(iii)' For all $P \in E(\mathbb{C})$ such that $2P \neq 0$ one has

$$\lambda(2P) = 4\lambda(P) - \log|2y(P)| + \frac{1}{4}\log|\Delta|.$$

Proof. The proof that λ is uniquely determined by (i),(ii) and (iii)' is exactly the same as the uniqueness proof of the preceding theorem. In fact, all we used from (iii) is that the difference f of any two Néron functions satisfies $f(2P) = 4f(P)$, which is already implied by (iii)'.

For proving (iii)' we assume first of all as before that $E(\mathbb{C})$ is the homomorphic image under the exponential map exp with respect to a suitable lattice $L := \mathbb{Z}\tau + \mathbb{Z}$. Then, setting $P = \exp(z)$ (so that $2P = \exp(2z)$ and $\frac{1}{2}\wp'(\tau, z) = y(P)$) we have to prove

$$|\wp'(\tau, z)| = |\Delta|^{\frac{1}{4}}\frac{G(2z)}{G(z)^4}.$$

But this follows immediately from Theorem 2.17 on comparing divisors on both sides (note that $G(2z) = 0$ if and only if $z \in \frac{1}{2}L$), which proves the identity up to multiplication by a constant, and by multiplying by $|z|^3$ and letting z tend to 0.

Finally one proves as for condition (iii) that (iii)' remains literally valid for the Néron function on an arbitrary elliptic curve (over $\mathbb{C}$) in Weierstrass normal form. □

2.9.2 The Néron functions associated to places

In this section we return again to an elliptic curve E defined over a number field K, say

$$E : y^2 = x^3 + Ax + B, \quad \Delta = -(4A^3 + 27B^2), \qquad (A, B \in K)$$

If v is a place of (i.e. equivalence class of valuations on) K, then we use $\|.\|_v$ for that valuation in v, whose restriction to $\mathbb{Q}$ equals the ordinary p-adic valuation $|\cdot|_p$ for some prime number p or the usual archimedean absolute value on $\mathbb{Q}$. We then have, with a suitable integer $n_v \geq 1$, the identity

$$|\alpha|_v = \|\alpha\|_v^{n_v}$$

for all $\alpha \in K$. We use K_v for the v-adic completion of K, and we the same symbol for the extension of $\|\cdot\|_v$ to K_v.

Generalising the theorem of the last section one can prove:

Theorem 2.20. *Let v be a place of K. Then there exists a unique function $\lambda_v : E(K_v)\setminus\{0\} \to \mathbb{R}$ satisfying properties (i) to (iii) of Theorem 2.18 with $\mathbb{C}$ replaced by K_v and the complex absolute value replaced by $\|\cdot\|_v$. The function λ_v can also be characterised as the unique real-valued function on $E(K_v)\setminus\{0\}$ which satisfies conditions (i), (ii) and condition (iii)' (of Theorem 2.19 with the same replacements as before). Assume that A and B are integral. Then, for all but finitely many v one has*

$$\lambda_v(P) = \frac{1}{2}\max(0, \log\|x(P)\|_v)$$

for all $P \in E(K_v) \setminus \{0\}$.

The function λ_v is called the local Néron function on E associated to v. The uniqueness of λ_v follows literally as in the proof of Theorem 2.18. If L is an extension of K, and if w is a place of L over v, then, since the restriction of λ_w to $E(K_v) \setminus \{0\}$ satisfies (i) to (iii), whence $\lambda_w(P) = \lambda_v(P)$ for $P \in E(K) \setminus \{0\}$.

If v is archimedean, i.e. if $K_v = \mathbb{C}$ or $K_v = \mathbb{R}$, then the existence of λ_v is ensured by Theorem 2.18. We shall not give the complete proof of the preceding theorem in the case of a non-archimedean v, but refer to the literature (cf. [Sil2]).

Here we content ourselves to prove the following theorem, which implies a part the preceding one for non-archimedean v where E has good reduction (and a little bit more). To state this theorem we need some notation.

Let $v \in P_K$ non-archimedean and assume that A and B are v-integral (i.e. $\|A\|_v, \|B\|_v \leq 1$). Let

$$O_v = \{x \in K_v \,|\, \|x\|_v \leq 1\}, \quad \mathfrak{m}_v = \{x \in K_v \,|\, \|x\|_v < 1\}.$$

Denote by $\widetilde{E}$ the curve over the field $O_v/\mathfrak{m}_v$ obtained from E by reducing A and B modulo the maximal ideal $\mathfrak{m}_v$. We have the map (in fact a homomorphism)

$$E \to \widetilde{E}, \qquad P \mapsto \widetilde{P}$$

obtained by reducing modulo $\mathfrak{m}_v$ (as explained in the proof of Lemma 2.4). We set

$$E_0(K_v) = \{P \in E(K_v) \,|\, \widetilde{P} \text{ is a nonsingular point of } \widetilde{E}\}.$$

It can be proved that this is a subgroup of $E(K_v)$ (see e.g. [Sil1], VII §2).

Theorem 2.21. *Let $v \in P_K$ non-archimedean, and assume that A and B are v-integral. Then*

$$\lambda_v(P) = \frac{1}{2}\max(\log \|x(P)\|_v, 0) - \frac{1}{12}\log \|\Delta\|_v$$

for all $P \in E_0(K_v) \setminus \{0\}$.

Proof. Denote the function on $E(K_v) \setminus \{0\}$ defined by the right hand side of the desired formula by λ. Clearly, λ satisfies properties (i) and (ii) of the local Néron function. Writing $|x|$ for $\|x\|_v$ we shall show the duplication formula

$$\lambda(2P) = 4\lambda(P) - \log|2y(P)| + \frac{1}{4}\log|\Delta|$$

for all $P \in E_0(K_v) \setminus \{0\}$.

This than implies that the restriction of λ_v to $E_0(K_v) \setminus \{0\}$ equals λ by the usual argument. Indeed, the difference $f = \lambda - \lambda_v$ extends to a continuous and bounded function on all of the subgroup $E_0(K_v)$ of $E(K_v)$. One has $f(2P) = 4f(P)$, by continuity even if $P = 0$ or $2P = 0$. But then $f(P) = 4^{-n}f(2^nP)$ for all n, which implies $f(P) = 0$ since f is bounded.

To prove the duplication formula for λ we note first of all (writing $x_1 = x(P)$ and $y_1 = y(P)$) that

$$x(2P) = -2x_1 + \frac{F_x(P)^2}{F_y(P)^2} = \frac{x_1^4 - 2Ax_1^2 - 8Bx_1 + A^2}{4y_1^2} =: \frac{\phi}{\psi},$$

where $F(x, y) = y^2 - (x^3 + Ax + B)$ and F_x, F_y denote partial derivatives.

Hence, the duplication formula is equivalent to

$$\frac{1}{2}\max(\log|\phi| - \log|\psi|, 0) = 2\max(\log|x_1|, 0) - \log|2y_1|,$$

which, using $|\psi|^2 = |2y|$, can be written as

$$\max(|\phi|, |\psi|) = \max(|x_1|^4, 1)$$

Assume, first of all that $|x_1| > 1$. Then, using that A, B are v-integral, we have $|\phi| = |x_1|^4$ and $|\psi| = |4y_1^2| = |4(x_1^3 + Ax_1 + B)| = |4x_1^3| < |x_1|^4 = |\phi|$. Hence the desired identity is true.

Now assume that x_1 is v-integral. Since A, B are v-integral y_1 is then v-integral too, in particular, we have $\widetilde{P} = [x_1 + \mathfrak{m}_v : y_1 + \mathfrak{m}_v : 1]$. We shall now use that $\widetilde{P}$ is a non-singular point of the reduced curve $\widetilde{E}$. This is equivalent to $|F_x(x_1, y_1)| = 1$ or $|F_y(x_1, y_1)| = 1$. Since

$$\phi = F_x(P)^2 - 2x_1F_y(P)^2 \quad \psi = F_y(P)^2$$

this implies that indeed $\max(|\phi|, |\psi|) = 1$. □

Note that in the case of good reduction, i.e. if $\|\Delta\|_v = 1$, we have the explicit formula

$$\lambda_v(P) = \frac{1}{2}\max(\log\|x(P)\|_v, 0),$$

and that we have actually proved that the right hand side satisfies the defining conditions (i), (ii) and (iii)' of the local Néron function at v.

2.9.3 The decomposition formula

Using the local Néron functions λ_v we can finally give the desired local decomposition of the canonical height h.

Theorem 2.22. *Let E be an elliptic curve defined over the number field K, let h be the canonical height on E, and, for each $v \in P_K$ let λ_v be the local Néron height function associated to v. Then*

$$h(P) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in P_K}\lambda_v(P)$$

for all $P \in E(K) \setminus \{0\}$.

Proof. Note that by Theorem 2.20 for each $P \in E(K)$, $P \neq 0$ we have $\lambda_v(P) = \frac{1}{2}\max(\log \|x(P)\|_v, 0)$ for almost all v. Hence the sum on the right hand side of the desired formula is actually finite (and hence well-defined). Denote by $h'(P)$ the function on $E(K)$ defined by the right hand side of the desired formula if $P \neq 0$, and such that $h'(0) = 0$.

To prove $h = h'$ it suffices to prove that $|h'(P) - \frac{1}{2}h_x(P)|$ is bounded and that $h'(2P) = 4h'(P)$ (see Theorem 2.13); here the bars denote the ordinary absolute value on $\mathbb{R}$.

The latter follows, for $2P \neq 0$, immediately from

$$\lambda_v(2P) = 4\lambda_v(P) - \log \|2y(P)\|_v + \frac{1}{4}\log \|\Delta\|_v$$

and the product formula (here written additively)

$$\sum_{v \in P_K} n_v \log \|x\|_v = 0,$$

valid for all $x \in K$, $x \neq 0$. For $P = 0$ its is trivially true since $h'(0) = 0$ by definition. For $2P = 0$ and $P \neq 0$ we have to show $h'(P) = 0$. This can be done e.g. by the triplication formula $\lambda_v(3P) = \lambda_v(P) + \log \|f(P)\| + \frac{2}{3}\log \|\Delta\|$ valid for all P with $3P \neq 0$ (cf. [Sil2], Exercise 6.4 (e); here $f \in K(E)$ independent of P).

From property (i) and (ii) of the Néron function we deduce the existence of constants c_v such that

$$-c_v \leq \lambda_v(P) - \frac{1}{2}\log \max(\|x(P\|_v, 1) \leq c_v$$

for all $v \in P_K$ and all $P \in E(K) \setminus \{0\}$. Even more, by the last theorem we can and will choose $c_v = 0$ for all but a finite number of v. Multiplying by $n_v/[K : \mathbb{Q}]$ and summing over all v then yields

$$|h'(P) - \frac{1}{2}h_x(P)| \leq \frac{1}{[K : \mathbb{Q}]} \sum_{v \in P_K} n_v c_v,$$

and hence the desired inequality. □

Part 3

Appendix: Exercises

The following exercises were given to the student at the end of the course as a written examination (in French). However, they supplement some of the threads of these notes and may hence be of independent interest.

3.1 Mesure de Mahler de polynômes en plusieurs variables

Pour un polynôme $P \in \mathbb{C}[X_1, \dots, X_n]$, $P \neq 0$, on pose

$$\mu(P) := \exp\Big(\int_0^1 \dots \int_0^1 \log |P(\mathrm{e}^{2\pi i t_1}, \dots, \mathrm{e}^{2\pi i t_n})| \, dt_1 \cdots dt_n \Big),$$

et on pose $\mu(0) = 0$. Dans l'exercice suivant la formule du cours

$$\int_0^1 \log |\alpha - \mathrm{e}^{2\pi i t}| \, dt = \log_+ |\alpha|$$

sera utile[1].

(i) En utilisant que $\mu(f) \geq |a_d|$ pour tout polynôme $f(x) = a_d X^d + \dots + a_0$ en une variable, montrer par récurrence sur n que $\mu(P) \geq 1$ si P a des coefficients entiers.

(ii) Montrer : Si $|\alpha_k| \geq \sum_{j=0,\, j\neq k}^{n} |\alpha_j|$ pour un $0 \leq k \leq n$, alors

$$\mu(a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n) = |a_k|.$$

En déduire $\mu(X_1 + X_2 + k) = |k|$ pour $|k| \geq 2$.

[1] Nous utilisons la notation $\log_+ x = \log \max(x, 1)$ $(x \in \mathbb{R},\, x > 0)$.

(iii) Calculer $\mu(X_1 + X_2)$.

(iv) Montrer d'abord que $\mu(X_1+X_2+1) = \int_{-1/3}^{1/3} \log|1+\mathrm{e}^{2\pi it}|\,dt$. Devélopper $\log|1+\mathrm{e}^{2\pi it}| = \mathrm{Re}\log(1+\mathrm{e}^{2\pi it})$ comme série en puissance de $\mathrm{e}^{2\pi it}$, échanger l'intégration et sommation (on admets la justification), et en déduire que

$$\log \mu(X_1 + X_2 + 1) = \frac{3\sqrt{3}}{4\pi} L(2, \left(\frac{\cdot}{3}\right)),$$
$$\text{où} \qquad L(s, \left(\frac{\cdot}{3}\right)) := \sum_{n=1}^{\infty} \left(\frac{n}{3}\right) n^{-s} \quad (s > 1).$$

On utilisera $\sum_n \left(\frac{n}{3}\right) n^{-s} = -4^{-s} \sum_n \left(\frac{n}{3}\right) n^{-s} + \sum_{n \text{ impair}} \left(\frac{n}{3}\right) n^{-s}$.

3.2 Calcul rapide de l'hauteur canonique

Soit $E : y^2 = x^3 + Ax + B$ une courbe elliptique définie sur $\mathbb{Q}$. Dans cet exercice on se propose de démontrer une formule pour l'hauteur canonique h sur E, qui peut être utile pour un calcul rapide. Pour simplifier nous supposons le suivant :

1. $A, B \in \mathbb{Z}$.
2. On a $f(x) := x^3 + Ax + B = (x-\alpha)(x-\overline{\alpha})(x-\beta)$ avec $\alpha \notin \mathbb{R}$ et $\beta > 0$.

Soit $\phi(x)$ le polynôme (de degré 4) tel que

$$x(2P) = \frac{\phi(x(P))}{4f(x(P))}$$

pour tout $P \in E(\mathbb{R})$, $P \neq 0$. Nous posons $h'(0) = 0$, et pour $P \in E(\mathbb{Q})$, $P \neq 0$, $x(P) = \frac{a}{b}$ avec $a, b \in \mathbb{Z}$ tels que $\mathrm{pgcd}(a, b) = 1$ nous posons

$$h'(P) = \log|a| + \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \log|\phi(x_n)/x_n^4|$$
$$\text{où} \quad x_0 = x(P), \quad x_{n+1} = \frac{\phi(x_n)}{4f(x_n)} \ (n \geq 0).$$

(i) Montrer que, pour $x \in \mathbb{R}$, $x > \beta$, on a $f(x) \neq 0$ et $\phi(x)/4f(x) \geq \beta$. Calculer $\phi(x)$ et montrer que $\phi(x)/x^4 \to 1$ pour $t \to \infty$ et $\phi(\beta)/\beta^4 > 0$. En déduire qu'il existe des constantes $c_1 > 0$ et c_2 telles que l'on a $c_1 \leq \phi(x)/x^4 \leq c_2$ pour tout $x \geq \beta$.

(ii) Déduire de (i) que la somme qui définit $h'(P)$ est bien-définie et converge absolument (en fait très rapidement).

(iii) En utilisant sans preuve le fait que $\mathrm{pgcd}(\phi(a/b)b^4, 4f(a/b)b^4) = 1$, montrer que $h'(2P) = 4h'(P)$.

(iv) Montrer : Il existe une constante c tel que $|h'(P) - \log\max(|a|,|b|)| \leq c$. (Ici l'estimation de (i) sera encore utile).

(v) Déduire de (iii) et (iv) que $h(P) = \frac{1}{2}h'(P)$.

3.3 Fonctions de Néron

Soit $E : y^2 = x^3 + Ax + B$ une courbe elliptique avec discriminant Δ définie sur le corps de nombre K_0, soit $|\cdot|$ une valuation de K_0 et K la complétion de K_0 par rapport à $|\cdot|$. Nous allons montrer dans cet exercice l'existence de la fonction de Néron en $|\cdot|$. Plus précisemment, nous nous proposons de montrer qu'il existe une fonction $\lambda : E(K) \setminus \{0\} \to \mathbb{R}$ tel que

1. λ est continu est borné sur le complément de tout voisinage de 0.
2. $\lim_{P\to 0}(\lambda(P) - \frac{1}{2}\log|x(P)|)$ existe (et est fini).
3. $\lambda(2P) = 4\lambda(P) - \log|2y(P)| + \frac{1}{4}\log|\Delta|$ pour tout $P \in E(K)$ tel que $2P \neq 0$.

(0) Montrer que $x(2P) = \frac{\phi(x(P))}{4f(x(P))}$ pour tout $P \in E$, où $f(x) = x^3 + Ax + B$ et $\phi(x) = -8xf(x) + f'(x)^2$.

(i) Pour $P \in E(K)$, $2P \neq 0$ on pose

$$f(P) := \frac{1}{2}\log_+|x(2P)| - 2\log_+|x(P)| + \log|2y(P)| - \frac{1}{4}\log|\Delta|.$$

Montrer que $g(P) := \exp(f(P))$ peut être prolongé à une fonction continue sur $E(K)$. Calculer $g(0)$ et en déduire qu'il existe un $c > 0$ tel que $g(P) > 0$ pour $|x(P)| > c$.

(ii) Montrer que les polynômes $\phi(x)$ et $4f(x)$ sont relativement premiers, et qu'ils existent donc des polynômes $a(x)$, $b(x)$ tel que $1 = a\phi + 4bf$, En déduire que $g(P) > 0$ pour $x(P) \leq c$ (avec le c de (i)).

(iii) Déduire de (i) et (ii) que $f(P)$ peut être prolongé uniquement à une fonction continue et bornée sur tout $E(K)$.

(iv) Montrer, en utilisant (iii), que la somme

$$\mu(P) := \sum_{n=0}^{\infty} 4^{-n} f(2^n P)$$

converge pour tout $P \in E(K)$ et définit une fonction continue et bornée $\mu : E(K) \to \mathbb{R}$ qui satisfait $f(P) = 4\mu(P) - \mu(2P)$ pour tout $P \in E(K)$.

(v) Montrer, en résumant, que la fonction $\lambda(P) := \lambda_1(P) + \mu(P)$, définie pour $P \in E(K)$, $P \neq 0$, satisfait aux propriétés 1. à 3.

Bibliography

[Ahlf] Lars V. Ahlfors, Complex Analysis. 2nd edition, McGraw-Hill Kogakusha, Tokyo 1966. 4

[BeZa] F. Beukers and D. Zagier, Lower bounds of heights of points on hypersurfaces, Acta Arith. LXXIX (1997), 103–111. 29

[Bomb] E. Bombieri, The Mordell conjecture revisited, preprint *???* 39

[Boyd] D.W. Boyd, reciprocal numbers having small measure I, II, Comp. Math. 35 (1980), 1361–1377 and 53 (1989), 355–357, S1–S6. 8

[CoDi] H. Cohen and F. Diaz y Diaz, A polynomial reduction algorithm Sém. Théor. Nombres Bordeaux (Sér. 2) 3 (1991), 351-360. 6

[Dobr] E. Dobrowolski, On a question of Lehmer and the number of irreducible factors of a polynomial, Acta Arith. XXXIV (1979). 28

[Doch] C. Doche, Thèse de troisième cycle en préparation, Bordeaux 1998. 29

[FeTo] M. Fekete and G. Szegö, On algebraic equations with integral coefficients whose roots belong to a given point set, Math. Zeitschr. 63 (1955), 158–172. 19, 21

[Heck] E. Hecke, Vorlesungen über die Theorie der algebraischen Zahlen, Chelsea, New York 1970. 12

[HoSk] G. Hoehn and N-P. Skoruppa, Un résultat de Schinzel, Journ. Théor. Nombres Bordeaux 5 (1993), 185. 14

[Lan1] S. Lang, Fundamentals of Diophantine Geometry, Springer, New York 1983. 41, 52

[Lan2] S. Lang, Algebra, Addison-Wesley, Reading 1978. 51

[Lan3] S. Lang, Algebraic numbers, Addison Wesley, Reading 1964. 52

[Lvin] M. Langevin, *???* (ref. to Langevin's theorem) 18

[Lehm] D.H. Lehmer, Factorization of certain cyclotomic functions, Ann. Math. 34 (1933), 461–479. 7

[Loub] R. Louboutin, Sur la mesure de Mahler d'un nombre algébrique, C.R.Acad.Sci. Paris 296 (1983), 707–708. 28

[Mani] *???* (ref. to "For any K and any prime number p there exists a constant N such that the p-part of $E(K)_{\text{tor}}$, for any E/K, is bounded to above by N.") 45

[Mazu] *???* (ref. to the classification of $E(\mathbb{Q})_{\text{tor}}$.)

[Neuk] J. Neukirch, Algebraische Zahlentheorie. Springer, Berlin 1992. 2

[Pari] BaBeBeCoOl, Computer algebra package for number theorists, Bordeaux 1989-1998. 6

[Schi] A. Schinzel, On the product of the conjugates outside the unit circle of an algebraic number, Acta Arithmetica 24 (1973), 385–399. 14, 29

[Sieg] C.L. Siegel, Algebraic integers whose conjugates lie in the unit circle, Duke M. J. 11 (1944), 597–602 or No. 46 in gesammelte Abhandlungen. 9, 28

[Sil1] J.H. Silverman, The Arithmetic of Elliptic Curves, Springer, New-York 1986. 50, 52, 60

[Sil2] J.H. Silverman, Advanved Topics in the Arithmetic of Elliptic Curves, Spinger, New-York 1994. 59, 62

[Skor] N-P. Skoruppa, Modular forms *in* Hirzebruch, Berger and Jung, Manifolds and Modular Forms, Vieweg, Braunschweig 1992. 55

[Smy1] C.J. Smyth, On the product of the conjugates outside the unit circle of an algebraic integer, Bull. London Math. Soc. 3 (1971), 169–175. 23

[Smy2] C.J. Smyth, On the Mahler measure of the composition of two polynomials, Acta Arith. LXXIX (1997), 239–247. 29

[Weil] A, Weil, *???* (ref. to Weil's study of "divisor → line bundle → projective embedding → height".) 41

[Zagi] D. Zagier, Algebraic numbers close to both 0 and 1, Math. Comp. 61 (1993), 485–491. 14, 15, 31

[Zhan] S. Zhang, Positive line bundles on arithmetic surfaces, preprint, Princeton 1992. *???* 15