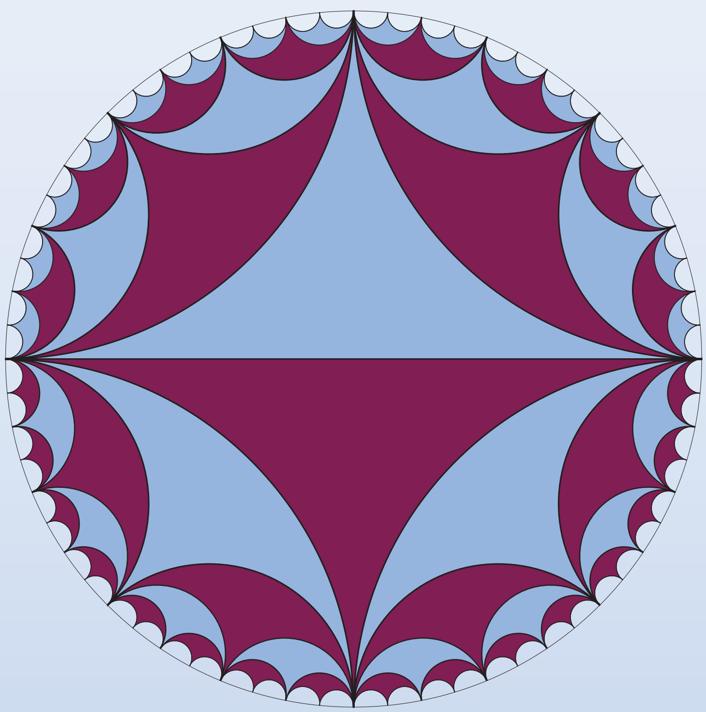
Topology of Numbers

Allen Hatcher



Topology of Numbers

Allen Hatcher

Revised Version
December 2023

AMERICAN MATHEMATICAL SOCIETY

2020 *Mathematics Subject Classification.* Primary 11-XX; Secondary 11-01.

The publisher's webpage for this book is:
www.ams.org/bookpages/mbk-145

Library of Congress Cataloging-in-Publication Data
Names: Hatcher, Allen, author.
Title: Topology of numbers / Allen Hatcher
Description: Providence, Rhode Island : American Mathematical Society, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2022021573 | ISBN 9781470456115 (paperback)
Subjects: LCSH: Geometry of numbers. | Algebraic topology. | Number theory. | AMS: Number theory. | Number theory – Instructional exposition (textbooks, tutorial papers, etc.).
Classification: LCC QA241.5 .H38 2022 | DDC 512.7/5-dc23/eng20220723
LC record available at https://lccn.loc.gov/2022021573

Copying and reprinting. Individual readers of this publication, and nonprofit libraries acting for them, are permitted to make fair use of the material, such as to copy select pages for use in teaching or research. Permission is granted to quote brief passages from this publication in reviews, provided the customary acknowledgment of the source is given.

Republication, systematic copying, or multiple reproduction of any material in this publication is permitted only under license from the American Mathematical Society. Requests for permission to reuse portions of AMS publication content are handled by the Copyright Clearance. Center. For more information, please visit **www.ams.org/publications/pubpermissions**.

Send requests for translation rights and licensed reprints to **reprint-permission@ams.org**.

©2022 by Allen Hatcher. All rights reserved.
Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1   27 26 25 24 23 22

Table of Contents

Preface

This book is an introduction to Number Theory from a more geometric point of view than is usual for the subject, inspired by the idea that pictures are often a great aid to understanding. The title of the book, *Topology of Numbers*, is intended to express this visual slant, where we are using the term "Topology" with its general meaning of "the spatial arrangement and interlinking of the components of a system".

The other unusual aspect of the book is that, rather than giving a broad introduction to all the basic tools of Number Theory without going too deeply into any one, it focuses on a single topic, quadratic forms $Q(x,y) = ax^2 + bxy + cy^2$ with integer coefficients. Here there is a very rich theory that one can really immerse oneself into to get a deeper sense of the beauty and subtlety of Number Theory. Along the way we do in fact encounter many standard number-theoretic tools, with some context to show how useful they can be.

A central geometric theme of the book is a certain two-dimensional figure known as the *Farey diagram*, discovered by Adolf Hurwitz in 1894, which displays certain relationships between rational numbers beyond just their usual distribution along the one-dimensional real number line. Among the many things the diagram elucidates that will be explored in the book are Pythagorean triples, the Euclidean algorithm, Pell's equation, continued fractions, Farey sequences, and two-by-two matrices with integer entries and determinant ± 1.

But most importantly for this book, the Farey diagram can be used to study quadratic forms $Q(x,y) = ax^2 + bxy + cy^2$ via John Conway's marvelous idea of the *topograph* of such a form. The origins of the wonderfully subtle theory of quadratic forms can be traced back to ancient times. In the 1600s interest was reawakened by numerous discoveries of Fermat, but it was only in the period 1750-1800 that Euler, Lagrange, Legendre, and especially Gauss were able to uncover the main features of the theory.

The principal goal of the book is to present an accessible introduction to this theory from a geometric viewpoint that complements the usual purely algebraic approach. Prerequisites for reading the book are fairly minimal, hardly going beyond high school mathematics for the most part. One topic that often forms a significant part of elementary number theory courses is congruences modulo an integer n. It would be helpful if the reader has already seen and used these a little since we will not develop congruence theory as a separate topic and will instead just use congruences as the need arises, proving whatever nontrivial facts are required including several of the basic ones that form part of a standard introductory number theory course. Among these is quadratic reciprocity, where we give Eisenstein's classical proof since it involves some geometry.

The high point of the basic theory of quadratic forms $Q(x,y)$ is the *class group* first constructed by Gauss. This can be defined purely in terms of quadratic forms, which is how it was first presented, or by means of Kronecker's notion of ideals introduced some 75 years after Gauss's work. For subsequent developments and generalizations the viewpoint of ideals has proven to be central to all of modern algebra. In this book we present both approaches to the class group, first the older version just in terms of forms, then the later version using ideals.

Here is how the book is organized. A preliminary Chapter 0 gives a sample of some of the sorts of questions studied in Number Theory, in particular motivating the study of quadratic forms by seeing how they arise in understanding Pythagorean triples, the integer side-lengths of right triangles, such as 3,4,5 and 5,12,13.

After this introduction the next three chapters lay the groundwork for our approach to quadratic forms by introducing the Farey diagram and its first applications to visualizing the Euclidean algorithm and continued fractions, both finite and infinite.

The next four chapters are the heart of the book. Chapter 4 introduces the topograph of a quadratic form, which displays all its values visually in a convenient and effective picture. A variety of examples are given illustrating different kinds of qualitative behavior of the topograph. As applications, topographs give efficient ways to compute the values of periodic and eventually periodic continued fractions, and to find all the integer solutions of Pell's equation $x^2 - dy^2 = \pm 1$.

Chapter 5 develops the classification theory for quadratic forms $ax^2 + bxy + cy^2$ in terms of the discriminant $b^2 - 4ac$. There are only a finite number of essentially distinct forms of a given discriminant, and it is shown how to compute these. Forms with symmetry play a special role, and a fairly complete picture of these is developed.

Chapter 6 turns to the fundamental representation problem, which is to find all the values a given form takes on, or in other words, to determine when an equation $ax^2 + bxy + cy^2 = n$ has integer solutions. There are two central themes here: how the factorization of n into primes plays a key role, largely reducing the problem to the case that n itself is prime; and how congruences modulo the discriminant give useful criteria for solvability, particularly in the case of primes.

Chapter 7 completes the basic theory by presenting Gauss's discovery of a way to multiply forms of a given discriminant, refining the multiplication of the values of the forms. This leads to an explanation of the seemingly mysterious fact that while there is essentially only one form of a given discriminant that represents a given prime, there can be several different forms representing nonprimes.

Finally, the rather lengthy Chapter 8 goes in a different direction to give an exposition of the alternative viewpoint toward quadratic forms by expanding the set of rational numbers to sets of numbers $a+b\sqrt{n}$ with a and b rational. Here the deeper subtleties of quadratic forms are translated into subtleties with the factorization of such numbers into “primes” and the lack of uniqueness of such factorizations. In keeping with the viewpoint of the rest of the book, we strive to make this essentially algebraic theory as geometric as possible.

At the end of the book there are several tables giving the key data for quadratic forms of small discriminant.

This book will remain available online in electronic form for free downloading after it has been published in the traditional paper form. The web address where it can be found is

http://www.math.cornell.edu/˜hatcher

Also available here will be a list of corrections as well as possible revisions and additions to the book. Readers are encouraged to send comments and corrections to the email address posted on the web page.

Note on the December 2023 revision. This version of the book contains various minor revisions to Chapters 1 through 7, including a few small additions. These changes add about ten pages to the length of the book, although some of that is just extra blank space on some pages.

0 A Preview

In this preliminary Chapter 0 we introduce by means of examples some of the main themes of Number Theory, particularly those that will be emphasized in the rest of the book.

Pythagorean Triples

Let us begin by considering right triangles whose sides all have integer lengths. The most familiar example is the $(3,4,5)$ right triangle, but there are many others as well, such as the $(5,12,13)$ right triangle. Thus we are looking for triples (a,b,c) of positive integers such that $a^2 + b^2 = c^2$. Such triples are called *Pythagorean triples* because of the connection with the Pythagorean Theorem. Our goal will be a formula that gives them all. The ancient Greeks knew such a formula, and even before the Greeks the ancient Babylonians must have known a lot about Pythagorean triples because one of their clay tablets from nearly 4000 years ago has been found which gives a list of 15 different Pythagorean triples, the largest of which is $(12709, 13500, 18541)$. (Actually, the tablet only gives the numbers a and c from each triple (a,b,c) for some unknown reason, but it is easy to compute b from a and c.)

There is an easy way to create infinitely many Pythagorean triples from a given one just by multiplying each of its three numbers by an arbitrary number n. For example, from $(3,4,5)$ we get $(6,8,10)$, $(9,12,15)$, $(12,16,20)$, and so on. This process produces right triangles that are all similar to each other, so in a sense they are not essentially different triples. In our search for Pythagorean triples there is thus no harm in restricting our attention to triples (a,b,c) whose three numbers have no common factor. Such triples are called *primitive*. The large Babylonian triple mentioned above is primitive, since the prime factorization of 13500 is $2^2 3^3 5^3$ but the other two numbers in the triple are not divisible by 2, 3, or 5.

A fact worth noting in passing is that if two of the three numbers in a Pythagorean triple (a,b,c) have a common factor n, then n is also a factor of the third number. This follows easily from the equation $a^2 + b^2 = c^2$, since for example if n divides a and b, then n^2 divides a^2 and b^2, so n^2 divides their sum c^2, hence n divides c.

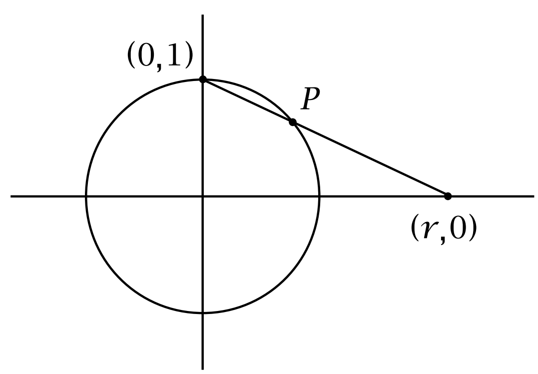
Another case is that n divides a and c. Then n^2 divides a^2 and c^2, so n^2 divides their difference $c^2 - a^2 = b^2$, hence n divides b. In the remaining case that n divides b and c the argument is similar.

A consequence of this divisibility fact is that primitive Pythagorean triples can also be characterized as the ones for which no two of the three numbers have a common factor.

If (a,b,c) is a Pythagorean triple, then we can divide the equation $a^2 + b^2 = c^2$ by c^2 to get an equivalent equation $({}^a/_c)^2 + ({}^b/_c)^2 = 1$. This equation is saying that the point $(x,y) = ({}^a/_c, {}^b/_c)$ is on the unit circle $x^2 + y^2 = 1$ in the xy-plane. The coordinates ${}^a/_c$ and ${}^b/_c$ are rational numbers, so each Pythagorean triple gives a *rational point* on the circle, a point whose coordinates are both rational. Notice that multiplying each of a, b, and c by the same nonzero integer n yields the same point (x,y) on the circle. Going in the other direction, given a rational point on the circle, we can find a common denominator for its two coordinates so that it has the form $({}^a/_c, {}^b/_c)$ and hence gives a Pythagorean triple (a,b,c). We can assume this triple is primitive by canceling any common factor of a, b, and c, and this does not change the point $({}^a/_c, {}^b/_c)$. The two fractions ${}^a/_c$ and ${}^b/_c$ must then be in lowest terms since we observed earlier that if two of a, b, c have a common factor, then all three have a common factor.

From the preceding observations we can conclude that the problem of finding all Pythagorean triples is equivalent to finding all rational points on the unit circle $x^2 + y^2 = 1$. More specifically, there is an exact one-to-one correspondence between primitive Pythagorean triples and rational points on the unit circle that lie in the interior of the first quadrant (since we want all of a,b,c,x,y to be positive).

In order to find all the rational points on the circle $x^2 + y^2 = 1$ we will use a construction that starts with one rational point and creates many more rational points from this one starting point. The four obvious rational points on the circle are the intersections of the circle with the coordinate axes, which are the points $(\pm 1, 0)$ and $(0, \pm 1)$. It does not matter which one we choose as the starting point, so let us choose $(0,1)$. Now consider a line which intersects the circle in this point $(0,1)$ and some other point P, as in the figure at the right. If the line has slope m, its equation will be $y = mx + 1$. If we denote the point where the line intersects the x-axis by $(r,0)$, then $m = -{}^1/_r$ so the equation for the line can be rewritten as $y = 1 - {}^x/_r$. Here we assume r is nonzero since $r = 0$ corresponds to the slope m being infinite and the point P being $(0,-1)$, a rational point we already know about. To find the coordinates of the point P in terms of r when $r \neq 0$ we substitute $y = 1 - {}^x/_r$ into the equation



$x^2 + y^2 = 1$ and solve for x:

$$x^2 + \left(1 - \frac{x}{r}\right)^2 = 1$$
$$x^2 + 1 - \frac{2x}{r} + \frac{x^2}{r^2} = 1$$
$$\left(1 + \frac{1}{r^2}\right)x^2 - \frac{2x}{r} = 0$$
$$\left(\frac{r^2+1}{r^2}\right)x^2 = \frac{2x}{r}$$

We are assuming $P \neq (0,-1)$ so $x \neq 0$ and we can cancel an x from both sides of the last equation above and then solve for x to get $x = {}^{2r}/_{r^2+1}$. Plugging this into the formula $y = 1 - {}^{x}/_{r}$ gives $y = 1 - {}^{2}/_{r^2+1} = {}^{r^2-1}/_{r^2+1}$. Thus the coordinates (x, y) of the point P are given by:

$$(x, y) = \left(\frac{2r}{r^2+1}, \frac{r^2-1}{r^2+1}\right)$$

Note that in these formulas we no longer have to exclude the value $r = 0$, which just gives the point $(0,-1)$. Observe also that if we let r approach $\pm\infty$ then the point P approaches $(0,1)$, as we can see either from the picture or from the formulas.

If r is a rational number, then the formula for (x, y) shows that both x and y are rational, so we have a rational point on the circle. Conversely, if both coordinates x and y of the point P on the circle are rational, then the slope m of the line must be rational, hence r must also be rational since $r = -{}^{1}/_{m}$. We could also solve the equation $y = 1 - {}^{x}/_{r}$ for r to get $r = {}^{x}/_{1-y}$, showing again that r will be rational if x and y are rational (and y is not 1). The conclusion of all this is that, starting from the initial rational point $(0,1)$ we have found formulas that give all the other rational points on the circle.

Since there are infinitely many different choices for the rational number r, there are infinitely many rational points on the circle. But we can say something much stronger than this: every arc of the circle, no matter how small, contains infinitely many rational points. This is because every arc on the circle corresponds to an interval of r-values on the x-axis, and every interval in the x-axis contains infinitely many rational numbers. Since every arc on the circle contains infinitely many rational points, we can say that the rational points are *dense* in the circle, meaning that for every point on the circle there is an infinite sequence of rational points approaching the given point.

Now we can go back and find formulas for Pythagorean triples. If we set the rational number r equal to ${}^{p}/_{q}$ with p and q integers having no common factor, then the formulas for x and y become:

$$x = \frac{2({}^{p}/_{q})}{({}^{p}/_{q})^2 + 1} = \frac{2pq}{p^2+q^2} \qquad \text{and} \qquad y = \frac{({}^{p}/_{q})^2 - 1}{({}^{p}/_{q})^2 + 1} = \frac{p^2-q^2}{p^2+q^2}$$

These formulas give the ratios $x = a/c$ and $y = b/c$ for all Pythagorean triples (a,b,c), so they determine all Pythagorean triples up to multiplication by a constant. The simplest way to realize the ratios $a/c = 2pq/p^2+q^2$ and $b/c = p^2-q^2/p^2+q^2$ is just to take:

$$(a,b,c) = (2pq, p^2 - q^2, p^2 + q^2)$$

The Pythagorean triples given by this formula may not be primitive, however. For example, if x and y are both odd then $p^2 - q^2$ and $p^2 + q^2$ are both even, as is $2pq$, so the triple could be simplified by dividing by 2. The nonprimitive triples obtained in this way are the starred entries in the table below.

(p,q)	(x,y)	(a,b,c)
$(2,1)$	$(4/5,3/5)$	$(4,3,5)$
$(3,1)^*$	$(6/10,8/10)^*$	$(6,8,10)^* \longrightarrow (3,4,5)$
$(3,2)$	$(12/13,5/13)$	$(12,5,13)$
$(4,1)$	$(8/17,15/17)$	$(8,15,17)$
$(4,3)$	$(24/25,7/25)$	$(24,7,25)$
$(5,1)^*$	$(10/26,24/26)^*$	$(10,24,26)^* \longrightarrow (5,12,13)$
$(5,2)$	$(20/29,21/29)$	$(20,21,29)$
$(5,3)^*$	$(30/34,16/34)^*$	$(30,16,34)^* \longrightarrow (15,8,17)$
$(5,4)$	$(40/41,9/41)$	$(40,9,41)$
$(6,1)$	$(12/37,35/37)$	$(12,35,37)$
$(6,5)$	$(60/61,11/61)$	$(60,11,61)$
$(7,1)^*$	$(14/50,48/50)^*$	$(14,48,50)^* \longrightarrow (7,24,25)$
$(7,2)$	$(28/53,45/53)$	$(28,45,53)$
$(7,3)^*$	$(42/58,40/58)^*$	$(42,40,58)^* \longrightarrow (21,20,29)$
$(7,4)$	$(56/65,33/65)$	$(56,33,65)$
$(7,5)^*$	$(70/74,24/74)^*$	$(70,24,74)^* \longrightarrow (35,12,37)$
$(7,6)$	$(84/85,13/85)$	$(84,13,85)$

Notice that the primitive versions of the starred triples occur higher in the table, but with a and b switched. This is a general phenomenon, as we will see in the course of proving the following basic result:

Proposition. *All primitive Pythagorean triples (a,b,c), after perhaps interchanging a and b, are obtained from the formula $(a,b,c) = (2pq, p^2 - q^2, p^2 + q^2)$ by letting p and q range over all positive integers with $p > q$, such that p and q have no common factor and are of opposite parity (one even and the other odd).*

Proof: We have seen that the formula $(a,b,c) = (2pq, p^2 - q^2, p^2 + q^2)$ yields all Pythagorean triples up to multiplication by a constant, so we just need to investigate when the formula gives a primitive triple and what to do when it gives a nonprimitive triple. As before we can assume that p and q have no common divisor, and we can assume that $p > q$ in order for the middle coordinate $b = p^2 - q^2$ to be positive.

Case 1: Suppose p and q have opposite parity. If all three of $2pq$, $p^2 - q^2$, and $p^2 + q^2$ have a common divisor $d > 1$ then d would have to be odd since $p^2 - q^2$ and

p^2+q^2 are odd when p and q have opposite parity. Furthermore, since d is a divisor of both p^2-q^2 and p^2+q^2 it must divide their sum $(p^2+q^2)+(p^2-q^2)=2p^2$ and their difference $(p^2+q^2)-(p^2-q^2)=2q^2$. However, since d is odd it would then have to divide p^2 and q^2, forcing p and q to have a common factor (since any prime factor of d would have to divide p and q). This contradicts the assumption that p and q have no common factors, so we conclude that $(2pq, p^2-q^2, p^2+q^2)$ is primitive if p and q have opposite parity.

Case 2: Suppose p and q have the same parity. Then their sum and difference are both even and we can write $p+q=2P$ and $p-q=2Q$ for some integers P and Q. Any common factor of P and Q would have to divide $P+Q=\frac{1}{2}(p+q)+\frac{1}{2}(p-q)=p$ and $P-Q=\frac{1}{2}(p+q)-\frac{1}{2}(p-q)=q$, so P and Q have no common factors. In terms of P and Q our Pythagorean triple becomes:

$$\begin{aligned}(a,b,c) &= (2pq, p^2-q^2, p^2+q^2)\\ &= (2(P+Q)(P-Q), (P+Q)^2-(P-Q)^2, (P+Q)^2+(P-Q)^2)\\ &= (2(P^2-Q^2), 4PQ, 2(P^2+Q^2))\\ &= 2(P^2-Q^2, 2PQ, P^2+Q^2)\end{aligned}$$

Canceling the factor of 2 in front of this last expression gives a new Pythagorean triple $(P^2-Q^2, 2PQ, P^2+Q^2)$ of the same type $(2pq, p^2-q^2, p^2+q^2)$ that we started with but with the first two coordinates switched. This new triple is primitive by Case 1 since P and Q cannot have the same parity, otherwise $p=P+Q$ and $q=P-Q$ would both be even, which is impossible since they have no common factor.

From Cases 1 and 2 we can conclude that if we allow ourselves to switch the first two coordinates, then we get all primitive Pythagorean triples from the formula by restricting p and q to be of opposite parity and have no common factors. □

Pythagorean Triples and Quadratic Forms

There are many questions one can ask about Pythagorean triples (a,b,c). For example, we could begin by asking which numbers actually arise as the numbers a, b, or c in some Pythagorean triple. It is sufficient to answer the question just for primitive Pythagorean triples, since the remaining ones are obtained by multiplying by arbitrary positive integers. We know all primitive Pythagorean triples arise from the formula

$$(a,b,c) = (2pq, p^2-q^2, p^2+q^2)$$

where p and q have no common factor and are of opposite parity. The latter condition just amounts to saying p and q are not both odd since they cannot both be even if they have no common factor. Determining whether a given number can be expressed in one of the forms $2pq$, p^2-q^2, or p^2+q^2 is a special case of the general question of deciding when an equation $Ap^2 + Bpq + Cq^2 = n$ has an integer solution p, q, for given integers A, B, C, and n. Expressions of the form $Ax^2 + Bxy + Cy^2$ are called *quadratic forms*. These will be the main topic studied in Chapters 4–8, where we will develop some general theory addressing the question of what values a quadratic form takes on when all the numbers involved are integers. For now, let us just look at the special cases at hand.

First let us consider which numbers occur as a or b in primitive Pythagorean triples (a, b, c). A trivial case is the equation $0^2 + 1^2 = 1^2$ which shows that 0 and 1 can be realized by the triple $(0, 1, 1)$ which is primitive, so let us focus on realizing numbers bigger than 1. If we look at the earlier table of Pythagorean triples we see that all the numbers up to 15 can be realized as a or b in primitive triples except for 2, 6, 10, and 14. This might lead us to guess that the numbers realizable as a or b in primitive Pythagorean triples are the numbers not of the form $4k + 2$. This is indeed true, and can be proved as follows. First note that since $2pq$ is even, $p^2 - q^2$ must be odd, otherwise both a and b would be even, violating primitivity. Now, every odd number is expressible in the form $p^2 - q^2$ since $2k + 1 = (k + 1)^2 - k^2$, so in fact every odd number is the difference between two consecutive squares. Taking $p = k + 1$ and $q = k$ yields a primitive triple since k and $k + 1$ always have opposite parity and no common factors. This takes care of realizing odd numbers. For even numbers, they would have to be expressible as $2pq$ with p and q of opposite parity, which forces pq to be even so $2pq$ is a multiple of 4 and hence cannot be of the form $4k + 2$. On the other hand, if we take $p = 2k$ and $q = 1$ then $2pq = 4k$ with p and q having opposite parity and no common factors.

To summarize, we have shown that all positive numbers $2k+1$ and $4k$ occur as a or b in primitive Pythagorean triples but none of the numbers $4k+2$ occur. To finish the story, note that a number $a = 4k+2$ which cannot be realized in a primitive triple can be realized by a nonprimitive triple just by taking a triple (a, b, c) with $a = 2k+1$ and doubling each of a, b, and c. Thus all numbers can be realized as a or b in Pythagorean triples (a, b, c).

Now let us ask which numbers c can occur in Pythagorean triples (a, b, c), so we are trying to find a solution of $p^2 + q^2 = c$ for a given number c. Pythagorean triples (p, q, r) give solutions when c is equal to a square r^2, but we are asking now about arbitrary numbers c. It suffices to figure out which numbers c occur in primitive triples (a, b, c), since by multiplying the numbers c in primitive triples by arbitrary numbers we get the numbers c in arbitrary triples. A look at the earlier table shows that the numbers c that can be realized by primitive triples (a, b, c) seem to be fairly rare: only $5, 13, 17, 25, 29, 37, 41, 53, 61, 65$, and 85 occur in the table. These are all odd, and in fact they are all of the form $4k + 1$. This always has to be true because p and q are of opposite parity, so one is an even number $2k$ and the other an odd number $2l + 1$. Squaring, we get $(2k)^2 = 4k^2$ and $(2l + 1)^2 = 4(l^2 + l) + 1$. Thus the square of an even number has the form $4u$ and the square of an odd number has the form $4v + 1$. Hence $p^2 + q^2$ has the form $4(u + v) + 1$, or more simply, just $4k + 1$.

The argument we just gave can be expressed more concisely using congruences modulo 4. We will assume the reader has seen something about congruences before, but to recall the terminology: two numbers a and b are said to be congruent modulo a number n if their difference $a - b$ is a multiple of n. When n is negative, congruence modulo n is equivalent to congruence modulo $|n|$, so there is no loss of generality in restricting attention just to congruence modulo positive numbers. Congruence modulo 0 is the same as equality, so there is little reason to consider this case. One writes $a \equiv b \bmod n$ to mean that a is congruent to b modulo n, with the word "modulo" abbreviated to "mod". One can tell whether two numbers are congruent mod n by dividing each of them by n and checking whether the remainders, which lie between 0 and $n - 1$, are equal. Every number is congruent mod n to one of the numbers $0, 1, 2, \cdots, n-1$, and no two of these numbers are congruent to each other, so there are exactly n congruence classes of numbers mod n, where a congruence class means all the numbers congruent to a given number. In the preceding paragraph we were in effect dealing with congruence classes mod 4 and we saw that the square of an even number is congruent to 0 mod 4 while the square of an odd number is congruent to 1 mod 4, hence $p^2 + q^2$ is congruent to $0 + 1$ or $1 + 0$ mod 4 when p and q have opposite parity, so $p^2 + q^2 \equiv 1 \bmod 4$.

Returning to the question of which numbers occur as c in primitive Pythagorean triples (a, b, c), we have seen that $c \equiv 1 \bmod 4$, but looking again at the list $5, 13, 17, 25, 29, 37, 41, 53, 61, 65, 85$ we can observe the more interesting fact that most of these numbers are primes, and the ones that are not primes are products of earlier primes in the list: $25 = 5 \cdot 5$, $65 = 5 \cdot 13$, $85 = 5 \cdot 17$. From this somewhat slim evidence one might conjecture that the numbers c occurring in primitive Pythagorean triples are exactly the numbers that are products of primes congruent to 1 mod 4. The first prime satisfying this condition that is not on the original list is 73, and this is realized as $p^2 + q^2 = 8^2 + 3^2$ in the triple $(48, 55, 73)$. The next two primes congruent to 1 mod 4 are $89 = 8^2 + 5^2$ and $97 = 9^2 + 4^2$, so the conjecture continues to look good. As further evidence for the conjecture, numbers congruent to 1 mod 4 that are not on the list such as $9 = 3 \cdot 3$, $21 = 3 \cdot 7$, $33 = 3 \cdot 11$, $45 = 3^2 \cdot 5$, $49 = 7 \cdot 7$, and $57 = 3 \cdot 19$ each have a prime factor that is not congruent to 1 mod 4.

More generally, if we ask which numbers can be expressed as $p^2 + q^2$ for integers p and q having no common divisor without requiring them to have opposite parity, then we will also get the numbers c in the starred entries of the earlier table. As we saw in the proof of the proposition about Pythagorean triples, these values of c are just the doubles of the values of c in primitive Pythagorean triples. Thus one can conjecture that the numbers expressible as $p^2 + q^2$ for positive integers p and q having no common divisor are the products of primes congruent to 1 mod 4 and the doubles of these products. This conjecture is correct, but proving it is not easy. We will do this in Chapter 6.

After this it is easy to go the last step and ask which numbers are sums p^2+q^2 for arbitrary positive integers p and q. Now we are free to multiply p and q by the same positive integer k, which multiplies p^2+q^2 by k^2. This leads to the answer that the numbers expressible as p^2+q^2, besides 0 and 1, are all the numbers n for which each prime factor congruent to 3 mod 4 occurs to an even power in the prime factorization of n. Thus the sequence of numbers that are sums of two squares begins $0,1,2,4,5,8,9,10,13,16,17,18,20,25,26,29,32,34,36,37,40,\cdots$.

Another question one can ask about Pythagorean triples is how many there are with two of the three numbers differing only by 1. In the earlier table there are several: $(3,4,5)$, $(5,12,13)$, $(7,24,25)$, $(20,21,29)$, $(9,40,41)$, $(11,60,61)$, and $(13,84,85)$. As the pairs of numbers that differ by 1 get larger, the corresponding right triangles are either approximately 45-45-90 right triangles, as with the triple $(20,21,29)$, or long thin triangles, as with $(13,84,85)$. To analyze the possibilities, note first that if two of the numbers in a triple (a,b,c) differ by 1 then the triple has to be primitive, so we can use our formula $(a,b,c)=(2pq,p^2-q^2,p^2+q^2)$. If b and c differ by 1 then we would have $(p^2+q^2)-(p^2-q^2)=2q^2=1$ which is impossible. If a and c differ by 1 then we have $p^2+q^2-2pq=(p-q)^2=1$ so $p-q=\pm1$, and in fact $p-q=+1$ since we must have $p>q$ in order for $b=p^2-q^2$ to be positive. Thus we get the infinite sequence of solutions $(p,q)=(2,1),(3,2),(4,3),\cdots$ with corresponding triples $(4,3,5),(12,5,13),(24,7,25),\cdots$. Note that these are the same triples we obtained earlier that realize all the odd values $b=3,5,7,\cdots$.

The remaining case is that a and b differ by 1. Thus we have the equation $p^2-2pq-q^2=\pm1$. The left side does not factor using integer coefficients, so it is not so easy to find integer solutions this time. In the table there are only the two triples $(4,3,5)$ and $(20,21,29)$, with $(p,q)=(2,1)$ and $(5,2)$. After some trial and error one could find the next solution $(p,q)=(12,5)$ which gives the triple $(120,119,169)$. Is there a pattern in the solutions $(2,1),(5,2),(12,5)$? One has the numbers $1,2,5,12$, and perhaps it is not too great a leap to notice that the third number is twice the second plus the first, while the fourth number is twice the third plus the second. If this pattern continued, the next number would be $29=2\cdot12+5$, giving $(p,q)=(29,12)$, and this does indeed satisfy $p^2-2pq-q^2=1$, yielding the Pythagorean triple $(696,697,985)$. These numbers are increasing rather rapidly, and the next case $(p,q)=(70,29)$ yields an even bigger Pythagorean triple $(4060,4059,5741)$. Could there be other solutions of $p^2-2pq-q^2=\pm1$ with smaller numbers that we missed? We will develop tools in Chapters 4 and 5 to find all the integer solutions, and it will turn out that the sequence we have just discovered gives them all.

Although the quadratic form $p^2 - 2pq - q^2$ does not factor using integer coefficients, it can be simplified slightly be rewriting it as $(p-q)^2 - 2q^2$. Then if we change variables by setting $(x, y) = (p - q, q)$ we obtain the quadratic form $x^2 - 2y^2$. Finding integer solutions of $x^2 - 2y^2 = n$ is equivalent to finding integer solutions of $p^2 - 2pq - q^2 = n$ since integer values of p and q give integer values of x and y, and conversely, integer values of x and y give integer values of p and q since when we solve for p and q in terms of x and y, we again get equations with integer coefficients: $(p, q) = (x + y, y)$. Thus the quadratic forms $p^2 - 2pq - q^2$ and $x^2 - 2y^2$ are completely equivalent, and finding integer solutions of $p^2 - 2pq - q^2 = \pm 1$ is equivalent to finding integer solutions of $x^2 - 2y^2 = \pm 1$.

The equation $x^2 - 2y^2 = \pm 1$ is an instance of the equation $x^2 - Dy^2 = \pm 1$ which is known as *Pell's equation* (although sometimes this term is used only when the right side of the equation is $+1$ and the other case is called the negative Pell equation). This is a very famous equation in Number Theory which has arisen in many different contexts going back hundreds of years. We will develop techniques for finding all integer solutions of Pell's equation for arbitrary values of D in Chapters 4 and 5. It is interesting that certain fairly small values of D can force the solutions to be quite large. For example, for $D = 61$ the smallest positive integer solution of $x^2 - 61y^2 = 1$ is a rather large pair:

$$(x, y) = (1766319049, 226153980)$$

As far back as the eleventh and twelfth centuries mathematicians in India knew how to find this solution. It was rediscovered in the seventeenth century by Fermat in France, who also gave the smallest solution of $x^2 - 109y^2 = 1$, an even larger pair:

$$(x, y) = (158070671986249, 15140424455100)$$

The way that the size of the smallest solution of $x^2 - Dy^2 = 1$ depends upon D is very erratic and is still not well understood today.

Pythagorean Triples and Complex Numbers

There is another way of looking at Pythagorean triples that involves complex numbers, surprisingly enough. The starting point here is the observation that $a^2 + b^2$ can be factored as $(a+bi)(a-bi)$ where $i = \sqrt{-1}$. If we rewrite the equation $a^2+b^2 = c^2$ as $(a + bi)(a - bi) = c^2$ then since the right side of the equation is a square, we might wonder whether each factor $a \pm bi$ on the left side would have to be a square too. For example, in the case of the triple $(3, 4, 5)$ we have $(3 + 4i)(3 - 4i) = 5^2$ with $3+4i = (2+i)^2$ and $3-4i = (2-i)^2$. So let us ask optimistically whether the equation $(a+bi)(a-bi) = c^2$ can be rewritten as $(p+qi)^2(p-qi)^2 = c^2$ with $a+bi = (p+qi)^2$ and $a - bi = (p - qi)^2$. We might hope also that the equation $(p + qi)^2(p - qi)^2 = c^2$

was obtained by simply squaring the equation $(p+qi)(p-qi) = c$. Let us see what happens when we multiply these various products out:

$$a + bi = (p+qi)^2 = (p^2 - q^2) + (2pq)i$$
$$\text{hence} \quad a = p^2 - q^2 \quad \text{and} \quad b = 2pq$$
$$a - bi = (p-qi)^2 = (p^2 - q^2) - (2pq)i$$
$$\text{hence again} \quad a = p^2 - q^2 \quad \text{and} \quad b = 2pq$$
$$c = (p+qi)(p-qi) = p^2 + q^2$$

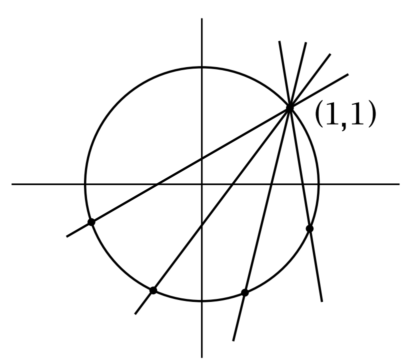
Thus we have miraculously recovered the formulas for Pythagorean triples that we obtained earlier by geometric means, with a and b switched, which does not really matter:

$$a = p^2 - q^2 \qquad b = 2pq \qquad c = p^2 + q^2$$

Our derivation of these formulas just now depended on several assumptions that we have not justified, but it does suggest that looking at complex numbers of the form $a + bi$ where a and b are integers might be a good idea. These complex numbers $a + bi$ with a and b integers are called *Gaussian integers*, after C. F. Gauss, the first mathematician to make a thorough algebraic study of them some 200 years ago. We will develop the basic properties of Gaussian integers in Chapter 8, in particular explaining why the derivation of the formulas above is valid.

Rational Points on Quadratic Curves

The same technique we used to find the rational points on the circle $x^2 + y^2 = 1$ can also be used to find all the rational points on other quadratic curves $Ax^2 + Bxy + Cy^2 + Dx + Ey = F$ with integer or rational coefficients A, B, C, D, E, F, provided that we can find a single rational point (x_0, y_0) on the curve to start the process. For example, the circle $x^2 + y^2 = 2$ contains the rational points $(\pm 1, \pm 1)$ and we can use one of these as an initial point. Taking the point $(1,1)$, we would consider lines $y - 1 = m(x-1)$ of slope m passing through this point. Solving this equation for y and plugging into the equation $x^2 + y^2 = 2$ would produce a quadratic equation $ax^2 + bx + c = 0$ whose coefficients are polynomials in the variable m, so these coefficients would be rational whenever m is rational. From the quadratic formula $x = (-b \pm \sqrt{b^2 - 4ac})/2a$ we see that the sum of the two roots is $-b/a$, a rational number if m is rational, so if one root is rational then the other root will be rational as well. The initial point $(1,1)$ on the curve $x^2 + y^2 = 2$ gives $x = 1$ as one rational root



of the equation $ax^2+bx+c=0$, so for each rational value of m the other root x will be rational. Then the equation $y-1=m(x-1)$ implies that y will also be rational, and hence we obtain a rational point (x,y) on the curve for each rational value of m. Conversely, if x and y are both rational and $x \neq 1$ then obviously $m = {}^{y-1}\!/_{x-1}$ will be rational. Thus one obtains a dense set of rational points on the circle $x^2+y^2=2$, since the slope m can be any rational number. An exercise at the end of the chapter is to work out the formulas explicitly.

Note that the point $(1,-1)$ is a rational point on the circle which does not arise from the formulas parametrizing x and y in terms of m since it corresponds to $m=\infty$. This is analogous to the earlier case of the circle $x^2+y^2=1$ where the point $(0,-1)$ corresponded to $m=\infty$ and $r=0$. For the circle $x^2+y^2=2$ we could just as well use the parameter r instead of m, with $(r,0)$ the point where the line through $(1,1)$ intersects the x-axis. There are simple formulas relating r and m, namely $r = {}^{m-1}\!/_{m}$ and $m = {}^{1}\!/_{1-r}$. From this viewpoint the exceptional slope $m=\infty$ corresponds to $r=1$ which is not exceptional for the parametrization by r, while the exceptional value $r=\infty$ corresponds to the nonexceptional value $m=0$ when the line through $(1,1)$ is parallel to the x-axis.

If we consider the circle $x^2+y^2=3$ instead of $x^2+y^2=2$ then there are no obvious rational points. And in fact this circle contains no rational points at all. For if there were a rational point, this would yield a solution of the equation $a^2+b^2=3c^2$ by integers a, b, and c with $c \neq 0$. We can assume a, b, and c have no common factor. Then a and b cannot both be even, otherwise the left side of the equation would be even, forcing c to be even, so a, b, and c would have a common factor of 2. To complete the argument we look at the equation modulo 4. As we saw earlier, the square of an even number is 0 mod 4, while the square of an odd number is 1 mod 4. Thus, modulo 4, the left side of the equation is either $0+1$, $1+0$, or $1+1$ since a and b are not both even. So the left side is either 1 or 2 mod 4. However, the right side is either $3\cdot 0$ or $3\cdot 1$ mod 4. We conclude that there can be no integer solutions of $a^2+b^2=3c^2$ with $c \neq 0$. When $c=0$ there is of course the trivial solution $(a,b,c)=(0,0,0)$ but this is not interesting so we will generally disregard it in equations of this type.

The technique we just used to show that $a^2+b^2=3c^2$ has no nontrivial integer solutions can be used in many other situations as well. The underlying reasoning is that if an equation with integer coefficients has an integer solution, then this gives a solution modulo n for all numbers n. For solutions modulo n there are only a finite number of possibilities to check, although for large n this is a large finite number. If one can find a single value of n for which there is no solution modulo n, then the original equation has no integer solutions. However, this implication is not reversible, as it is possible for an equation to have solutions modulo n for every number n and still have no actual integer solutions. A concrete example is the equation

$2x^2 + 7y^2 = 1$. This obviously has no integer solutions, yet it does have solutions modulo n for each n, although this is certainly not obvious. Note that the ellipse $2x^2 + 7y^2 = 1$ does contain rational points such as $(1/3, 1/3)$ and $(3/5, 1/5)$. These can in fact be used to show that $2x^2 + 7y^2 = 1$ has solutions modulo n for each n, as we will show in Section 2.3 of Chapter 2 when we study congruences in more detail.

In Chapter 6 we will find a complete answer to the question of when the circle $x^2 + y^2 = n$ contains rational points by showing that there are rational points on this circle only when there are integer points on it. This reduces the problem to one we considered earlier, finding the integers n that are sums of two squares.

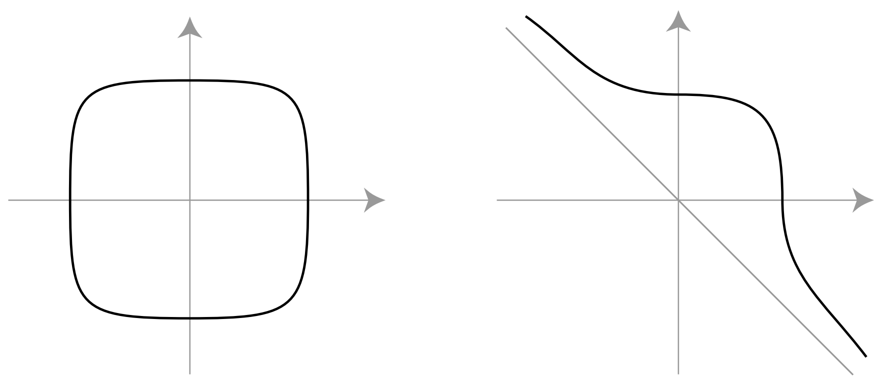
Determining when a quadratic curve contains rational points turns out to be much easier than determining when it has integer points. The general problem reduces fairly quickly to finding rational points on ellipses or hyperbolas of the special form $Ax^2 + By^2 = C$ where A, B, and C are integers that are not divisible by squares greater than 1, and such that no two of A, B, and C have a common factor. A theorem of Legendre then asserts that the curve $Ax^2 + By^2 = C$ contains rational points exactly when three congruence conditions modulo A, B, and C are satisfied, namely AC must be congruent mod B to the square of some number, and likewise BC must be a square mod A and $-AB$ must be a square mod C. (There is also the obvious condition that A and B cannot both have opposite sign from C.) For example, if $C = 1$ this reduces just to saying that each of A and B is congruent to a square modulo the other one since the congruence condition mod C holds automatically when $C = 1$. For the ellipse $2x^2 + 7y^2 = 1$ this agrees with what we saw earlier since 2 is a square mod 7, namely 3^2, and 7 is a square mod 2, namely 1^2, so Legendre's theorem guarantees that the curve has a rational point. In the case of the circle $x^2 + y^2 = 3$ the congruence conditions reduce simply to -1 being a square mod 3, which it is not since every number is congruent to 0, 1, or 2 mod 3 so the squares mod 3 are just 0 and 1 since $2^2 \equiv 1$ mod 3.

Diophantine Equations

Equations like $x^2 + y^2 = z^2$ or $x^2 - Dy^2 = 1$ that involve polynomials with integer coefficients, and where the solutions sought are required to be integers, or perhaps just rationals, are called *Diophantine equations* after the Greek mathematician Diophantus (ca. 250 A.D.) who wrote a book about these equations that was very influential when European mathematicians started to consider this topic much later in the 1600s. Usually Diophantine equations are very hard to solve because of the restriction to integer solutions. The first really interesting case is quadratic Diophantine equations. By the year 1800 there was quite a lot known about the quadratic case, and we will be focusing on this case in this book.

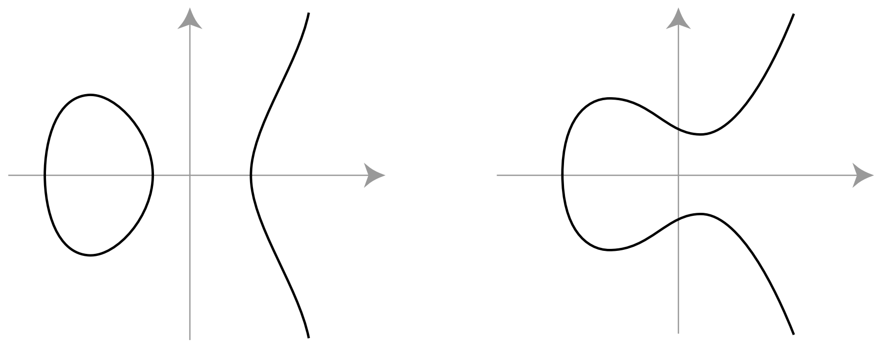
Diophantine equations of higher degree than quadratic are much more challenging to understand. Probably the most famous one is $x^n + y^n = z^n$ where n is a fixed integer greater than 2. In the 1600s when the French mathematician Fermat was reading about Pythagorean triples in his copy of Diophantus' book, he made a marginal note that, in contrast with the equation $x^2 + y^2 = z^2$, the equation $x^n + y^n = z^n$ has no solutions with positive integers x, y, z when $n > 2$. This is one of many statements that he claimed were true but never wrote proofs of for public distribution, nor have proofs been found among his manuscripts. Over the next century other mathematicians discovered proofs for all his other statements, but this one was far more difficult to verify. The issue is clouded by the fact that he only wrote this statement down the one time, whereas all his other important results were stated numerous times in his correspondence with other mathematicians of the time. So perhaps he only briefly believed he had a proof. In any case, the statement has become known as Fermat's Last Theorem. It was finally proved in the 1990s by Andrew Wiles, using some very deep mathematics developed mostly over the preceding couple decades.

We have seen that finding integer solutions of $x^2+y^2 = z^2$ is equivalent to finding rational points on the circle $x^2+y^2 = 1$, and in the same way, finding integer solutions of $x^n + y^n = z^n$ is equivalent to finding rational points on the curve $x^n + y^n = 1$. For even values of $n > 2$ this curve looks like a flattened circle or rounded square, while for odd n it has a similar shape in the first quadrant but a rather different shape elsewhere, extending out to infinity in the second and fourth quadrants, asymptotic to the line $y = -x$:



Fermat's Last Theorem is equivalent to the statement that these curves have no rational points except their intersections with the coordinate axes, where x or y is 0. These examples show that it is possible for a curve defined by an equation of degree greater than 2 to contain only a finite number of rational points (either two points or four points here, depending on whether n is odd or even) whereas quadratic curves like $x^2 + y^2 = n$ contain either no rational points or an infinite dense set of rational points.

After quadratic curves the next case that has been studied in great depth is cubic curves such as the curves defined by equations $y^2 = x^3 + ax^2 + bx + c$. These are known as elliptic curves, not because they are ellipses but because of a connection with the problem of computing the length of an arc of an ellipse. Depending on the values of the coefficients a, b, c elliptic curves can have either one or two connected pieces:

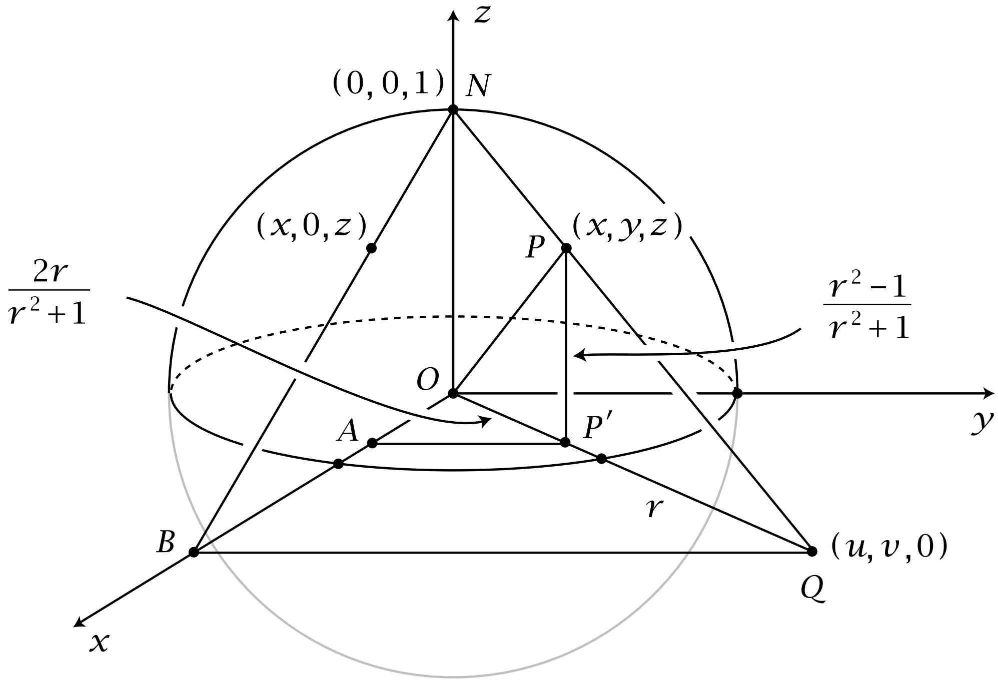


In some cases the number of rational points is finite, any number from 0 to 10 as well as 12 or 16 according to a difficult theorem of Mazur. In other cases the number of rational points is infinite and they form a dense set in the curve, or possibly just in the component that stretches to infinity when there are two components. There is no simple way known for predicting the number of rational points from the coefficients. Interestingly, elliptic curves play an important role in the proof of Fermat's Last Theorem. Their theory is much deeper than for quadratic curves, and so elliptic curves are well beyond the scope of this book.

Rational Points on a Sphere

Although we will not be discussing this later in the book, another way to generalize quadratic curves, in a different direction from considering cubic and higher degree curves, is to keep the quadratic condition but introduce more variables. After quadratic curves the next case would be quadratic surfaces, or as they are usually called, *quadric surfaces*. These are surfaces in three-dimensional space defined by an equation $Q(x,y,z) = n$ where $Q(x,y,z)$ is a quadratic function of three variables. Perhaps the simplest example is the equation $x^2 + y^2 + z^2 = 1$ which defines the sphere of radius 1 with center at the origin. Other quadric surfaces are ellipsoids, paraboloids, hyperboloids, and certain cones and cylinders.

Much of the theory of quadric surfaces parallels that for quadratic curves. To illustrate, let us consider the problem of finding all the rational points on the sphere $x^2 + y^2 + z^2 = 1$, the triples (x,y,z) of rational numbers that satisfy this equation. Some obvious rational points are the points where the sphere meets the coordinate axes such as the point $(0,0,1)$ on the z-axis. Following what we did for the circle $x^2+y^2 = 1$, consider a line from $(0,0,1)$ to a point $(u,v,0)$ in the xy-plane. This line intersects the sphere at some point (x,y,z), and we want to find formulas expressing x, y, and z in terms of u and v. To do this we use the following figure:



Suppose we look at the vertical plane containing the triangle ONQ. From our earlier analysis of rational points on a circle of radius 1 we know that if the segment OQ has length $|OQ| = r$, then $|OP'| = {}^{2r}\!/\!_{r^2+1}$ and $z = {}^{r^2-1}\!/\!_{r^2+1}$. From the right triangle OBQ we see that $u^2+v^2 = r^2$. The triangle OBQ is similar to the triangle OAP' and the scaling factor to go from OBQ to OAP' is

$$\frac{|OP'|}{|OQ|} = \frac{2r/(r^2+1)}{r} = \frac{2}{r^2+1}$$

Hence

$$x = \frac{2}{r^2+1}\cdot u = \frac{2u}{u^2+v^2+1} \qquad \text{and} \qquad y = \frac{2}{r^2+1}\cdot v = \frac{2v}{u^2+v^2+1}$$

Also we have

$$z = \frac{r^2-1}{r^2+1} = \frac{u^2+v^2-1}{u^2+v^2+1}$$

Summarizing, we have expressed x, y, and z in terms of u and v by the formulas

$$x = \frac{2u}{u^2+v^2+1} \qquad y = \frac{2v}{u^2+v^2+1} \qquad z = \frac{u^2+v^2-1}{u^2+v^2+1}$$

We can also express u and v in terms of x, y, and z. The projection of the point $P = (x,y,z)$ onto the xz-plane is the point $(x,0,z)$ which is on the line through B and N. The slope of this line is $-{}^{1}\!/\!_{u}$ so the equation for the line is $z = 1 - {}^{x}\!/\!_{u}$. Solving this for u gives $u = {}^{x}\!/\!_{1-z}$. Interchanging x and y corresponds to interchanging u and v so we also have $v = {}^{y}\!/\!_{1-z}$.

From the formulas relating (x,y,z) to (u,v) we see that x, y, and z are rational exactly when u and v are rational. Thus we have formulas for all the rational points (x,y,z) on the sphere except for the pole $(0,0,1)$ in terms of rational parameters u and v.

Here is a short table giving a few rational points on the sphere and the corresponding integer solutions of the equation $a^2 + b^2 + c^2 = d^2$:

(u, v)	(x, y, z)	(a, b, c, d)
$(1, 2)$	$(1/3, 2/3, 2/3)$	$(1, 2, 2, 3)$
$(2, 3)$	$(2/7, 3/7, 6/7)$	$(2, 3, 6, 7)$
$(1, 4)$	$(1/9, 4/9, 8/9)$	$(1, 4, 8, 9)$
$(2, 2)$	$(4/9, 4/9, 7/9)$	$(4, 4, 7, 9)$
$(1, 3)$	$(2/11, 6/11, 9/11)$	$(2, 6, 9, 11)$
$(3/2, 3/2)$	$(6/11, 6/11, 7/11)$	$(6, 6, 7, 11)$
$(3, 4)$	$(3/13, 4/13, 12/13)$	$(3, 4, 12, 13)$
$(2, 5)$	$(2/15, 5/15, 14/15)$	$(2, 5, 14, 15)$
$(1/2, 5/2)$	$(2/15, 10/15, 11/15)$	$(2, 10, 11, 15)$

These are in fact all the primitive positive solutions of $a^2 + b^2 + c^2 = d^2$ with $d \le 15$, up to permutations of a, b, and c.

As with rational points on the circle $x^2 + y^2 = 1$, rational points on the sphere $x^2 + y^2 + z^2 = 1$ are dense since rational points are dense in the xy-plane. Thus there are lots of rational points scattered all over the sphere. In linear algebra courses one is often called upon to create unit vectors (x, y, z) by taking a given vector and rescaling it to have length 1 by dividing it by its length. For example, the vector $(1, 1, 1)$ has length $\sqrt{3}$ so the corresponding unit vector is $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. It is rare that this process produces unit vectors having rational coordinates, but the formulas derived above give a way to create as many rational unit vectors as we like.

The correspondence we have described between points (x, y, z) on a sphere and points (u, v) in the plane is called *stereographic projection*. One can think of the sphere and the plane as being made of clear glass, and if one looks outward and downward from the north pole of the sphere the points of the sphere are projected onto points in the plane, and vice versa. The north pole itself does not project onto any point in the plane, but points approaching the north pole project to points approaching infinity in the plane, so one can think of the north pole as corresponding to an imaginary infinitely distant "point" in the plane. This geometric viewpoint somehow makes infinity less of a mystery, as it just corresponds to a point on the sphere, and points on a sphere are not very mysterious. (Though in the early days of polar exploration the north pole may have seemed very mysterious and infinitely distant.)

One might ask also about spheres $x^2 + y^2 + z^2 = n$, following what we did for circles $x^2 + y^2 = n$. Finding an integer point on $x^2 + y^2 + z^2 = n$ is asking whether n is a sum of three squares. One can test small values of n and one finds that most numbers are sums of three squares, so it is easier to list the ones that are not: $7, 15, 23, 28, 31, 39, 47, 55, 60, 63, 71, 79, 87, 92, 95, \cdots$. The odd numbers here are just the numbers $8k + 7$, and the even numbers seem to be 4 times the earlier numbers on the list. In fact it is easy to see that numbers congruent to 7 mod 8 cannot be expressed as sums of three squares by the following argument. The squares mod 8 are $0^2 = 0$, $(\pm 1)^2 = 1$, $(\pm 2)^2 = 4$, $(\pm 3)^2 = 9 \equiv 1$, and $4^2 = 16 \equiv 0$, so the squares of even numbers are 0 or 4 mod 8 and the squares of odd numbers are 1 mod 8. Obviously 7 cannot be realized as a sum of three terms 0, 1, or 4, so numbers congruent to 7 mod 8 cannot be sums of three squares.

To rule out numbers $4(8k+7)$ as sums of three squares, we can work mod 4 where the squares are just 0 and 1. If we have $x^2 + y^2 + z^2 = 4n$ then $x^2 + y^2 + z^2 \equiv 0$ mod 4, and the only way to get 0 as a sum of three numbers 0 or 1 is as $0 + 0 + 0$. This means each of x, y, and z must be even, so we can cancel a 4 from both sides of the equation $x^2 + y^2 + z^2 = 4n$ to get n expressed as a sum of three squares. Thus numbers $4(8k + 7)$ are never realizable as sums of three squares since $8k + 7$ is never a sum of three squares. Repeating this argument, we see that $16(8k + 7)$ is never a sum of three squares since $4(8k + 7)$ is not a sum of three squares. Similarly $4^l(8k + 7)$ is never a sum of three squares for any larger exponent l.

The converse statement that every number not of the form $4^l(8k + 7)$ is expressible as a sum of three squares is true but is much harder to prove. This was first done by Legendre.

This answers the question of when the sphere $x^2 + y^2 + z^2 = n$ contains integer points, but could it contain rational points without containing integer points? Let us show that this cannot happen. A rational point on $x^2 + y^2 + z^2 = n$ is equivalent to an integer solution of $a^2 + b^2 + c^2 = nd^2$. It will suffice to show that if n is not a sum of three squares, then neither is nd^2 for any integer d. An equivalent statement is that if n is of the form $4^l(8k+7)$ then so is nd^2. To prove this, let us write d as $2^p q$ with q odd and $p \ge 0$, hence $d^2 = 4^p q^2$ with $q^2 \equiv 1$ mod 8 since q is odd. Thus we have $nd^2 = 4^{l+p}(8k + 7)q^2$ where the product $(8k + 7)q^2$ is 7 mod 8 since $8k + 7$ is 7 mod 8 and q^2 is 1 mod 8. This shows what we wanted, that if n is of the form $4^l(8k + 7)$ then so is nd^2.

For a general quadric surface defined by a quadratic equation with integer coefficients there is a theorem due to Minkowski, analogous to Legendre's theorem for quadratic curves, that says that rational points exist exactly when certain congruence conditions are satisfied. In general, having rational points on a quadric surface is not equivalent to having integer points as it was for spheres, and the existence of integer points is a more delicate question.

Moving on to four variables, one could ask about integer or rational points on the spheres $x^2 + y^2 + z^2 + w^2 = n$ in four-dimensional space. Integers that could not be expressed as the sum of three squares can be realized as sums of four squares, for example $7 = 2^2 + 1^2 + 1^2 + 1^2$ and $15 = 3^2 + 2^2 + 1^2 + 1^2$, and it is a theorem of Lagrange that every positive number can be expressed as the sum of four squares. Thus the spheres $x^2 + y^2 + z^2 + w^2 = n$ always contain integer points.

Minkowski's theorem remains true for quadratic equations with integer coefficients in any number of variables, as does the fact that the existence of a single rational solution implies that rational solutions are dense.

Exercises

1. (a) Make a list of the 16 primitive Pythagorean triples (a, b, c) with $c \le 100$, regarding (a, b, c) and (b, a, c) as the same triple.
(b) How many more would there be if we allowed nonprimitive triples?
(c) How many triples (primitive or not) are there with $c = 65$?

2. (a) Find all the positive integer solutions of $x^2 - y^2 = 512$ by factoring $x^2 - y^2$ as $(x + y)(x - y)$ and considering the possible factorizations of 512.
(b) Show that the equation $x^2 - y^2 = n$ has only a finite number of integer solutions for each value of $n > 0$.
(c) Find a value of $n > 0$ for which the equation $x^2 - y^2 = n$ has at least 100 different positive integer solutions.

3. (a) Show that there are only a finite number of Pythagorean triples (a, b, c) with a equal to a given number n.
(b) Show that there are only a finite number of Pythagorean triples (a, b, c) with c equal to a given number n.

4. Find an infinite sequence of primitive Pythagorean triples where two of the numbers in each triple differ by 2.

5. Find a right triangle whose sides have integer lengths and whose acute angles are close to 30 and 60 degrees by first finding the irrational value of r that corresponds to a right triangle with acute angles exactly 30 and 60 degrees, then choosing a rational number close to this irrational value of r.

6. Find a right triangle whose sides have integer lengths and where one of the two shorter sides is approximately twice as long as the other, using a method like the one in the preceding problem. (One possible answer might be the $(8, 15, 17)$ triangle, or a triangle similar to this, but you should do better than this.)

7. Find a rational point on the sphere $x^2 + y^2 + z^2 = 1$ whose three coordinates are nearly equal.

8. (a) Derive formulas that give all the rational points on the circle $x^2 + y^2 = 2$ in terms of a rational parameter m, the slope of the line through the point $(1, 1)$ on the circle. (The value $m = \infty$ should be allowed as well, yielding the point $(1, -1)$.) The calculations may be a little messy, but they eventually simplify to give formulas that are not too complicated:

$$x = \frac{m^2 - 2m - 1}{m^2 + 1} \qquad y = \frac{-m^2 - 2m + 1}{m^2 + 1}$$

(b) Using these formulas, find five different rational points on the circle in the first quadrant, and hence five solutions of $a^2 + b^2 = 2c^2$ with positive integers a, b, c.

(c) The equation $a^2 + b^2 = 2c^2$ can be rewritten as $c^2 = \frac{1}{2}(a^2 + b^2)$, which says that c^2 is the average of a^2 and b^2, or in other words, the squares a^2, c^2, b^2 form an arithmetic progression. One can assume $a < b$ by switching a and b if necessary. Find four such arithmetic progressions of three increasing squares where in each case the three numbers have no common divisors.

9. (a) Find formulas that give all the rational points on the upper branch of the hyperbola $y^2 - x^2 = 1$.

(b) Can you find any relationship between these rational points and Pythagorean triples?

10. (a) Show that the equation $x^2 - 2y^2 = \pm 3$ has no integer solutions by considering this equation modulo 8.

(b) Show that there are no primitive Pythagorean triples (a, b, c) with a and b differing by 3.

11. Show there are no rational points on the circle $x^2 + y^2 = 3$ using congruences modulo 3 instead of modulo 4.

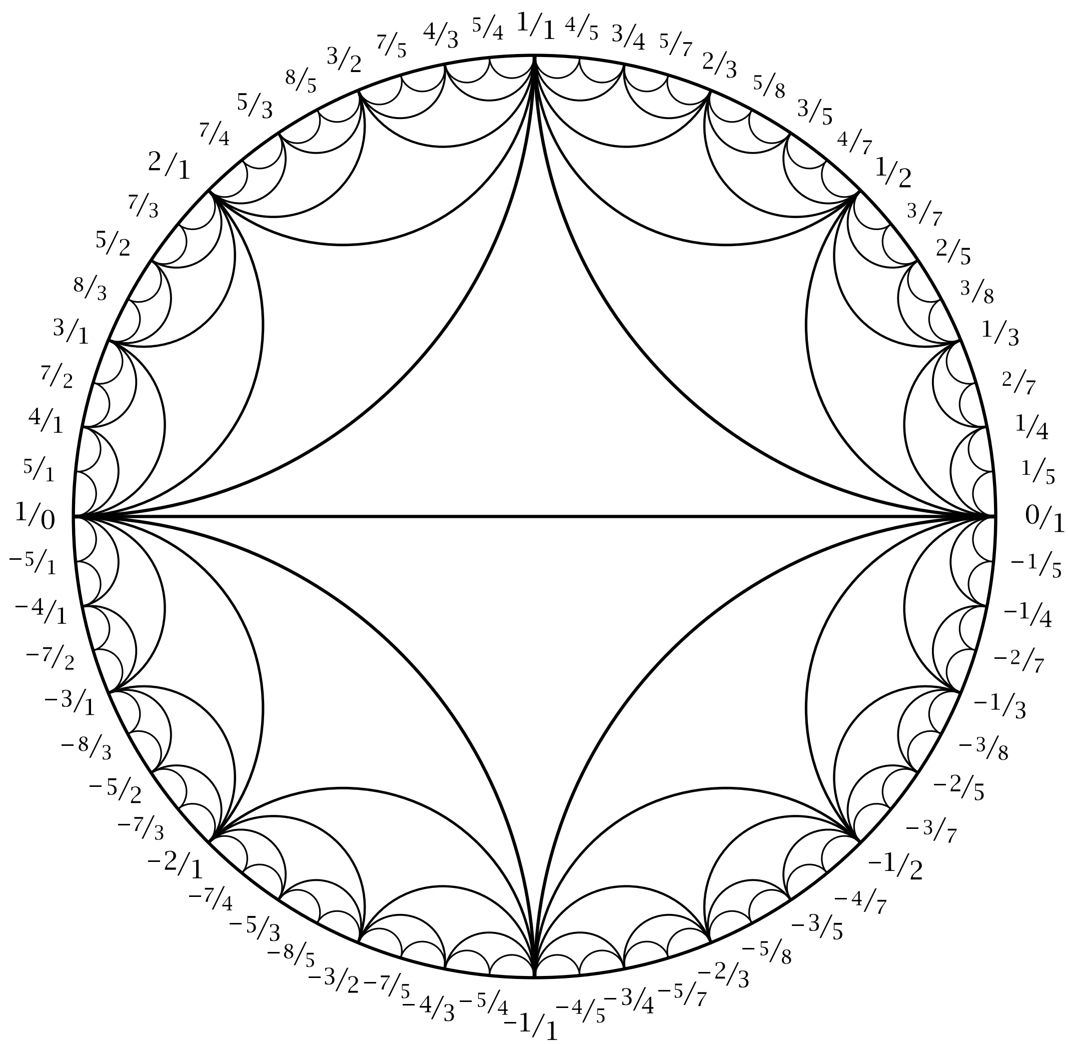
12. Show that for every Pythagorean triple (a, b, c) the product abc must be divisible by 60. (It suffices to show that abc is divisible by 3, 4, and 5.)

13. Use congruences modulo 8 to show that primitive solutions of $a^2 + b^2 + c^2 = d^2$ must have d odd and must have two of a, b, c even and the other odd.

14. Show that if the curve $x^n + y^n = 1$ has a rational point with x and y nonzero, then it has a rational point with x and y positive. *Hint*: Consider the equation $a^n + b^n = c^n$.

1 The Farey Diagram

Our goal is to use geometry to study numbers. Of the various kinds of numbers, the simplest are integers, along with their ratios, the rational numbers. Usually one thinks of rational numbers geometrically as points along a line, interspersed with irrational numbers as well. In this chapter we introduce a two-dimensional pictorial representation of rational numbers that displays certain interesting relations between them that we will be exploring. This diagram, along with several variants of it that will be introduced later, is known as the *Farey diagram*. The origin of the name will be explained when we get to one of these variants. Here is the diagram:

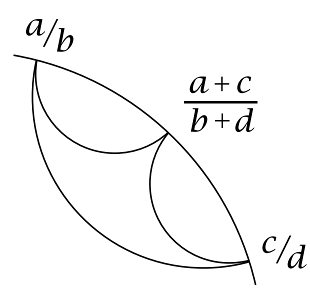


What is shown here is not the whole diagram but only a finite part of it. The actual diagram has infinitely many curvilinear triangles, getting smaller and smaller out near the boundary circle. The diagram can be constructed by first inscribing the two big triangles in the circle, then adding the four triangles that share an edge with the two big triangles, then the eight triangles sharing an edge with these four, then sixteen more triangles, and so on forever. With a little practice one can draw the diagram without lifting one's pencil from the paper: First draw the outer circle starting at the left or right side, then the diameter, then make the two large triangles, then the four next-largest triangles, and so on.

Our first task will be to explain how the vertices of all the triangles are labeled with rational numbers. Perhaps the reader can guess what the rules are before we spell them out in detail.

1.1 The Mediant Rule

The vertices of the triangles in the Farey diagram are labeled with fractions a/b, including the fraction $1/0$ for ∞, according to the following scheme. In the upper half of the diagram, first label the vertices of the big triangle $1/0$, $0/1$, and $1/1$. Then add labels for successively smaller triangles by the rule that, if the labels at the two ends of the long edge of a triangle are a/b and c/d, then the label on the third vertex of the triangle is $a+c/b+d$, so the numerators and denominators are added separately, contrary to the usual way of adding fractions. The fraction $a+c/b+d$ is called the *mediant* of a/b and c/d.

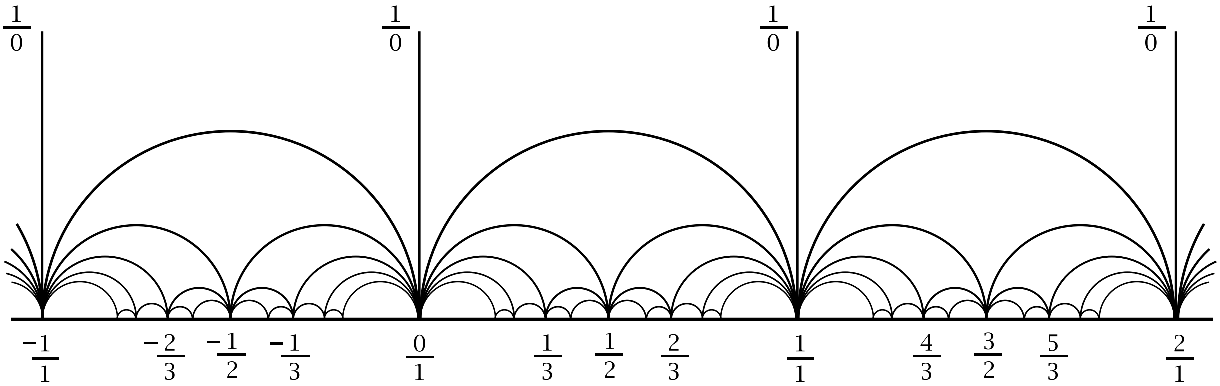


The labels in the lower half of the diagram follow the same scheme, starting with the labels $-1/0$, $0/1$, and $-1/1$ on the large triangle. Using $-1/0$ instead of $1/0$ as the label of the vertex at the far left means that we are regarding $+\infty$ and $-\infty$ as the same. The labels in the lower half of the diagram are the negatives of those in the upper half, and the labels in the left half are the reciprocals of those in the right half.

For fractions with a nonzero denominator our usual rule will be to write them with a positive denominator, so the sign of the fraction is the sign of the numerator.

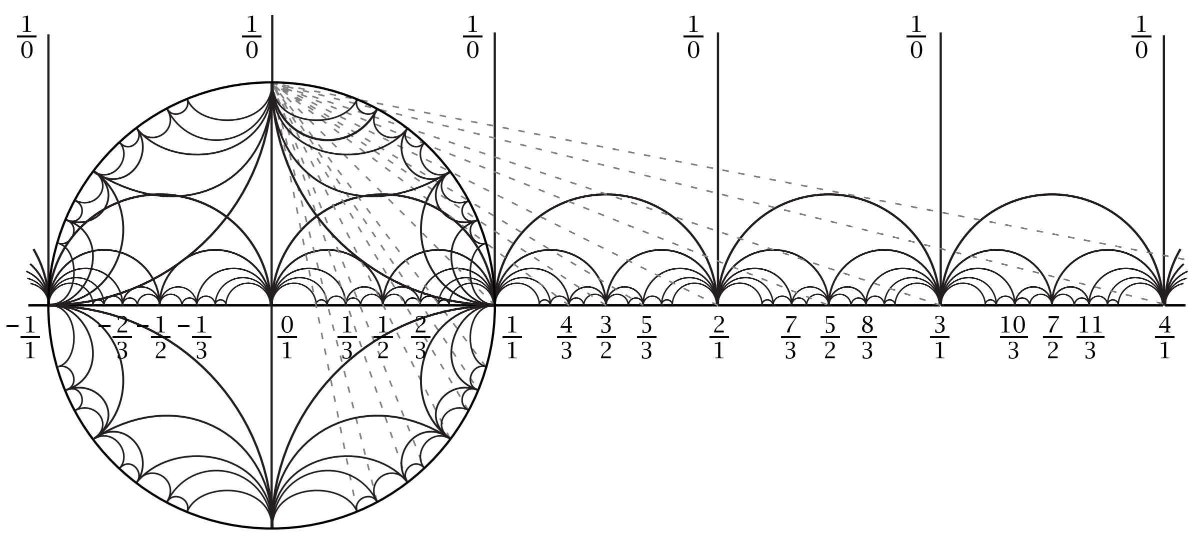
The labels generated by the mediant rule occur in their proper order around the circle, increasing from $-\infty$ to $+\infty$ as one goes around the circle in the counterclockwise direction. This is obviously true for the integer labels, and to verify it for the others it suffices to show that the mediant $a+c/b+d$ of a/b and c/d is always a number between a/b and c/d (hence the term "mediant"). Thus we want to show that if $a/b < c/d$ then $a/b < a+c/b+d < c/d$. These fractions all have positive denominators, so the inequality $a/b < c/d$ is equivalent to $ad < bc$ and $a/b < a+c/b+d$ is equivalent to $ab + ad < ab + bc$. Obviously $ad < bc$ implies $ab + ad < ab + bc$, so $a/b < c/d$ implies $a/b < a+c/b+d$. Similarly $a+c/b+d < c/d$ is equivalent to $ad + cd < bc + cd$ which also follows from $ad < bc$, so $a/b < c/d$ implies $a+c/b+d < c/d$.

There is another version of the Farey diagram with the boundary circle straightened out to a line:



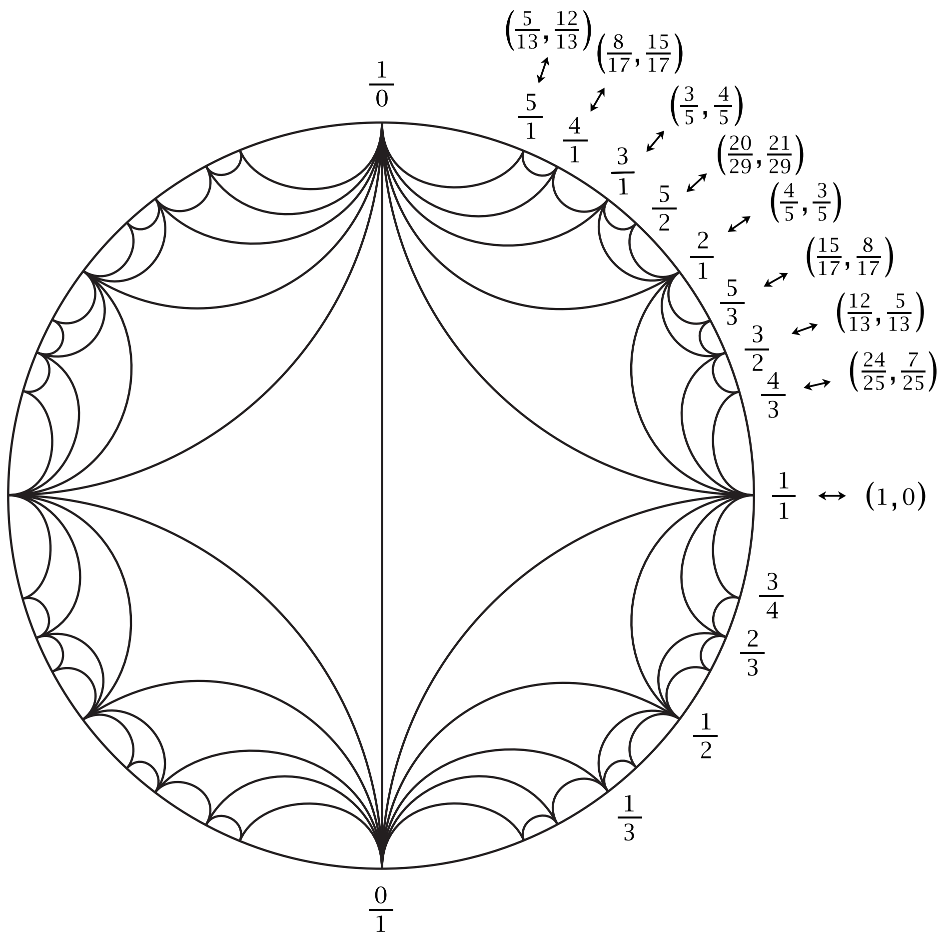
Here the diagram fills up the upper half of the xy-plane, with the vertex $1/0$ of the original Farey diagram positioned "at infinity" so it is not actually shown in the new version. The edges of the diagram with one endpoint at $1/0$ are drawn as vertical lines with lower endpoints at the integer points on the x-axis. All the other edges of the diagram are semicircles with endpoints on the x-axis, and we can position these so that the vertex labeled a/b is actually the number a/b on the x-axis. This is possible since when we construct the diagram by adding more and more curvilinear triangles, we can place the new vertex of each triangle at any point between its outer two vertices, so we just choose this new vertex to be at the mediant of the outer two vertices. With this rule the part of the diagram between each pair of consecutive integers n and $n+1$ looks the same since the mediant of $n+a/b$ and $n+c/d$ is $n+a+c/b+d$ as one can easily check by a simple calculation.

In the previous chapter we described how rational points (x,y) on the unit circle $x^2+y^2=1$ correspond to rational points p/q on the x-axis by means of lines through the point $(0,1)$ on the circle. Using this correspondence, we can label the rational points on the circle by the corresponding rational points on the x-axis and then construct a new Farey diagram in the circle by filling in triangles by the mediant rule just as before.



This gives a version of the circular Farey diagram that is rotated by 90 degrees to put $1/0$ at the top of the circle, and there are also some perturbations of the positions of the other vertices and the shapes of the triangles. For our purposes these perturbations will not make much of a difference since it will usually be just the combinatorial pattern of the triangles that is important. We drew the circular Farey diagram the way we did at the beginning of the chapter because it looks more symmetric and is easier to draw since one does not have to figure out the exact positions of the vertices.

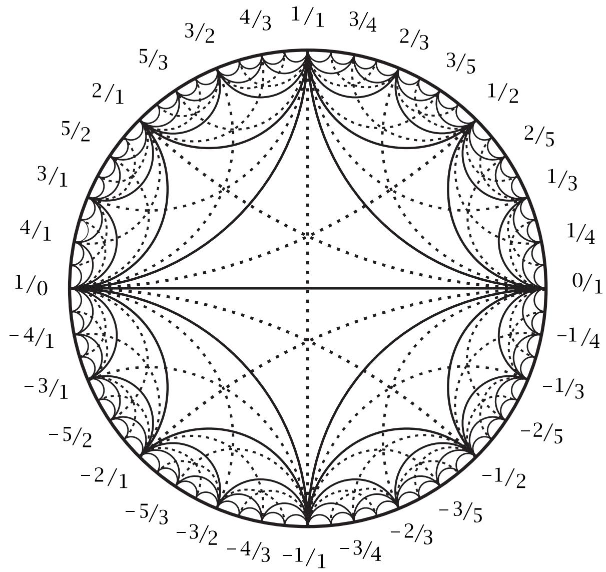
The next figure shows the relationship between the new circular Farey diagram and Pythagorean triples (a,b,c) using the formulas $(a,b,c) = (2pq, p^2-q^2, p^2+q^2)$ that we found in the previous chapter. The vertex with label p/q thus has coordinates $(x,y) = (a/c, b/c) = \left(2pq/p^2+q^2, p^2-q^2/p^2+q^2\right)$.



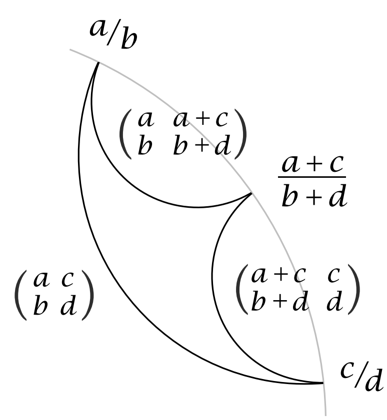
The construction we have described for the Farey diagram involves an inductive process where more and more edges and vertex labels are added in succession. With a construction like this it is not easy to tell by a simple calculation whether or not two given rational numbers a/b and c/d are joined by an edge in the diagram. Fortunately there is such a criterion:

Proposition 1.1. *For each pair of fractions a/b and c/d, including $\pm 1/0$, there exists an edge in the Farey diagram with endpoints labeled a/b and c/d if and only if the determinant $ad - bc$ of the matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ is equal to ± 1.*

What this means is that if one starts with the rational numbers together with $\pm 1/0$ arranged in order around a circle and one inserts circular arcs inside this circle meeting it perpendicularly and joining each pair of fractions a/b and c/d such that $ad - bc = \pm 1$, with the circular arc replaced by a diameter in case a/b and c/d are diametrically opposite on the circle, then no two of these arcs will cross, and they will divide the interior of the circle into nonoverlapping curvilinear triangles. This is really quite remarkable when you think about it, and it does not happen for other values of the determinant besides ± 1. For example, for determinant ± 2 the edges would be the dotted arcs in the figure at the right. Here there are three arcs crossing in each triangle of the original Farey diagram, and these arcs divide each triangle of the Farey diagram into six smaller triangles.



Proof: First we show by an inductive argument that for an edge in the diagram joining two fractions a/b and c/d the associated matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ has determinant ± 1. The induction starts with the edge joining $\pm 1/0$ to $0/1$ where the determinant condition obviously holds. All the other edges are added in stages, first the four edges creating the two biggest triangles, then the eight edges creating the next four triangles, and so on. Consider a triangle created at some stage by adding a new vertex labeled $a+c/b+d$ as the mediant of vertices a/b and c/d from an earlier stage, as in the figure at the right. We may assume by induction that $ad - bc = \pm 1$ for the long edge of the triangle which was added at an earlier stage. The determinant condition then holds also for the two shorter edges of the triangle since $a(b+d) - b(a+c) = ad - bc$ and $(a+c)d - (b+d)c = ad - bc$. Thus the determinant condition continues to hold after each stage of the construction of the diagram, so it holds for all edges.



Now we prove the converse, the statement that if $ad - bc = \pm 1$ then there is an edge in the diagram joining a/b and c/d. We may assume $b \geq 0$ and $d \geq 0$ by multiplying both numerator and denominator of either fraction by -1 if necessary, which multiplies the determinant $ad - bc$ by -1. The order of the two fractions a/b and c/d does not matter since interchanging the two columns of the matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ also multiplies the determinant by -1. If b or d is 0, say $b = 0$, then the determinant condition becomes $ad = \pm 1$ so $d = 1$ and $a = \pm 1$. In this case the fractions a/b

and c/d are $\pm 1/0$ and $c/1$ so they lie at the ends of an edge of the diagram, one of the vertical edges to $1/0$ in the upper halfplane version of the diagram. Thus for the rest of the proof we may assume $b > 0$ and $d > 0$.

The previous figure shows that adding a new triangle to the diagram creates two new edges corresponding to matrices obtained from $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ by replacing one of the columns by the sum of the two columns. To finish the proof we will show that for each matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ of determinant ± 1 with $b > 0$ and $d > 0$ it is possible to perform a finite sequence of the inverse operations of subtracting one column from the other and end up with a matrix that we already know corresponds to an edge in the diagram. We will do this by always subtracting the column with smaller second entry from the column with larger second entry, so that these two entries remain positive. We stop the process when the two entries in the second row become equal. For example, here is how the process works for the matrix $\begin{pmatrix} 3 & 7 \\ 8 & 19 \end{pmatrix}$:

$$\begin{pmatrix} 3 & 7 \\ 8 & 19 \end{pmatrix} \to \begin{pmatrix} 3 & 4 \\ 8 & 11 \end{pmatrix} \to \begin{pmatrix} 3 & 1 \\ 8 & 3 \end{pmatrix} \to \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Here the last matrix corresponds to the edge joining $1/1$ and $0/1$. Reversing the steps reducing $\begin{pmatrix} 3 & 7 \\ 8 & 19 \end{pmatrix}$ to $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, we are adding one column to the other at each stage so each new matrix produced in this way corresponds to an edge of the diagram. In particular this shows that the original matrix $\begin{pmatrix} 3 & 7 \\ 8 & 19 \end{pmatrix}$ corresponds to an edge of the diagram.

For the general argument we start with a matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ of determinant ± 1 with $b > 0$ and $d > 0$. If $b \neq d$ then we subtract the column with smaller second entry from the column with larger second entry, and repeat this operation until the two entries in the second row are equal. We cannot get a 0 in the second row since this would mean that the previous matrix already had equal entries in the second row. Once we get a matrix with equal entries in the second row, these entries will divide the determinant which is ± 1 so these entries must be 1. Thus the matrix is of the form $\begin{pmatrix} a & c \\ 1 & 1 \end{pmatrix}$, with determinant $a - c = \pm 1$ so a and c differ by 1. The corresponding fractions are then $n/1$ and $n+1/1$ for some integer n, and there is an edge of the diagram joining these two fractions, one of the large semicircles in the upper halfplane diagram. Hence when we reverse the sequence of column subtractions by performing a sequence of column additions, each successive matrix will correspond to an edge of the diagram and in particular $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ will correspond to an edge of the diagram. $\square$

The sign of the determinant $ad - bc$ has a simple interpretation for fractions a/b and c/d with positive denominators since in this case the inequality $ad - bc > 0$ is equivalent to $a/b > c/d$ and $ad - bc < 0$ is equivalent to $a/b < c/d$. Thus the sign of the determinant tells which of a/b or c/d is larger.

Here is an interesting consequence of the preceding proposition:

Corollary 1.2. *The mediant rule for labeling the vertices in the Farey diagram always produces labels a/b that are fractions in lowest terms.*

This would follow automatically if it was always true that the mediant of two fractions in lowest terms is again in lowest terms, but this is not always the case. For example, the mediant of $1/3$ and $2/3$ is $3/6$, and the mediant of $2/7$ and $3/8$ is $5/15$. Somehow cases like this do not occur in the Farey diagram.

Before deducing the corollary let us introduce a bit of standard terminology. For a fraction a/b to be in lowest terms means that a and b have no common factor greater than 1. This is equivalent to saying that the prime factorizations of a and b have no prime factor in common. When this is the case we say that a and b are *coprime*. An alternative terminology is to say that a and b are *relatively prime*.

Proof: From the way the Farey diagram is constructed, each labeled vertex a/b is joined to some other labeled vertex c/d by an edge of the diagram. By the easier half of Proposition 1.1 we have $ad - bc = \pm 1$. This implies that a and b are coprime since any common divisor of a and b must divide the products ad and bc, hence also the difference $ad - bc = \pm 1$, but the only divisors of ± 1 are ± 1. ☐

Proposition 1.1 can also be used to prove another basic fact about the Farey diagram:

Proposition 1.3. *Every fraction p/q in lowest terms occurs as the label on some vertex in the Farey diagram.*

Proof: We may assume p and q are nonzero since $0/1$ and $1/0$ certainly occur as labels in the diagram. Since the negative labels in the diagram are just the negatives of the positive labels, we can assume p and q are in fact positive. It will suffice to show that if p and q are coprime, then there is an edge in the diagram whose endpoints are labeled p/q and r/s for some integers r and s. By Proposition 1.1 this is equivalent to the existence of integers r and s such that $ps - qr = \pm 1$.

Consider a matrix $\begin{pmatrix} x & y \\ p & q \end{pmatrix}$ where the integers x and y are yet to be determined. In the proof of Proposition 1.1 there was a procedure for repeatedly subtracting the column with smaller second entry from the column with larger second entry until a matrix with equal second entries is obtained. Subtracting one column from the other does not affect coprimeness of the two second entries, so when the procedure is applied to a matrix $\begin{pmatrix} x & y \\ p & q \end{pmatrix}$ with p and q coprime, the result is a matrix whose second entries are equal and coprime, so these entries must be 1. Now let us choose a matrix of determinant ± 1 whose lower two entries are 1, say the matrix $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. If we start with this matrix and apply the reverse of the sequence of operations performed on $\begin{pmatrix} x & y \\ p & q \end{pmatrix}$ to get 1's in the second row, the resulting sequence of operations of adding one column to the other converts $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ into a matrix $\begin{pmatrix} r & s \\ p & q \end{pmatrix}$ of the same determinant ± 1. This means that we have found integers r and s such that $rq - ps = \pm 1$, or equivalently $ps - qr = \pm 1$. ☐

Implicit in this proof is a method for solving Diophantine equations of the form $px - qy = \pm 1$ for any two given coprime positive integers p and q. In Section 2.3 we will make this procedure explicit and streamline it to be more efficient.

Exercises

1. There is another version of the Farey diagram in which the vertex labeled p/q is placed at the point (q, p) in the plane, so p/q is the slope of the line through the origin and (q, p). The edges of this new Farey diagram are straight line segments connecting the pairs of vertices that are connected in the original Farey diagram. For example there is a triangle with vertices $(1,0)$, $(0,1)$, and $(1,1)$ corresponding to the big triangle in the upper half of the circular Farey diagram. With this model of the Farey diagram the operation of forming the mediant of two fractions just corresponds to standard vector addition $(a,b) + (c,d) = (a+c, b+d)$.

What you are asked to do in this problem is just to draw the portion of the new Farey diagram consisting of all the triangles whose vertices (q, p) satisfy $0 \le q \le 5$ and $0 \le p \le 5$. Note that since fractions p/q labeling vertices are always in lowest terms, the points (q, p) such that q and p have a common divisor greater than 1 are not vertices of the diagram.

2. Consider a vertex of the Farey diagram labeled a/b with $b > 1$. Show that of all the labels on vertices connected to the a/b vertex by an edge of the diagram, exactly two have denominator smaller than b.

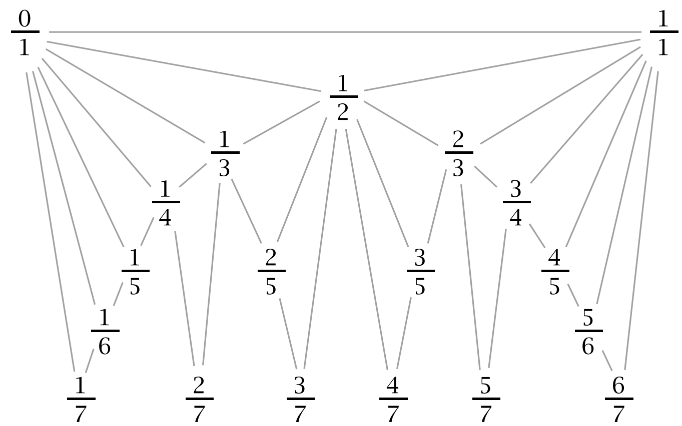
3. If a/b, c/d, and e/f are fractions in lowest terms such that e/f is the mediant of a/b and c/d, is it necessarily true that there is a triangle in the Farey diagram with vertices a/b, c/d, and e/f? Give either a proof or a counterexample.

4. (a) Reduce each of the matrices $\begin{pmatrix} 7 & 3 \\ 16 & 7 \end{pmatrix}$ and $\begin{pmatrix} 67 & 14 \\ 24 & 5 \end{pmatrix}$ to either $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ by repeatedly subtracting one column from the other as in the proof of Proposition 1.1.
(b) Use Proposition 1.1 to show that this can be done for any matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ with nonnegative entries and determinant ± 1.

5. Determine whether the following statement is always true: If $a/b < a'/b'$ and $c/d < c'/d'$ then the mediant of a/b and c/d is less than the mediant of a'/b' and c'/d'.

1.2 Farey Series

We can build the set of rational numbers by starting with the integers and then inserting in succession the halves, thirds, fourths, fifths, sixths, and so on. Let us look at what happens if we restrict to rational numbers between 0 and 1. Starting with 0 and 1 we first insert $1/2$, then $1/3$ and $2/3$, then $1/4$ and $3/4$, skipping $2/4$ which we already have, then inserting $1/5$, $2/5$, $3/5$, and $4/5$, then $1/6$ and $5/6$, etc. A natural way to depict this way of listing rational numbers between 0 and 1 is to place the terms with equal denominators in successive rows, with line segments connecting each new term to its two nearest neighbors among the terms in the previous rows as shown in the figure for denominators up to 7. Inspecting the figure, it appears that each new term is the mediant of its two neighbors, and we will show that in fact this always happens. This means that we are just constructing a straight-line version of the part of the Farey diagram between $0/1$ and $1/1$.



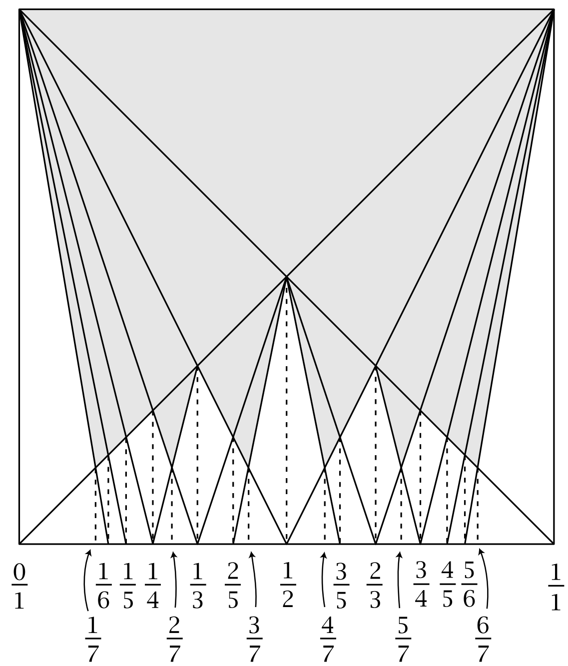
The discovery of this curious mediant property in the early 1800s was initially attributed to a geologist and amateur mathematician named Farey, although it turned out that he was not the first person to have noticed it. In spite of this confusion, the sequence of fractions a/b between 0 and 1 with denominator less than or equal to a given number n is called the nth *Farey series* F_n. For example, here is F_7:

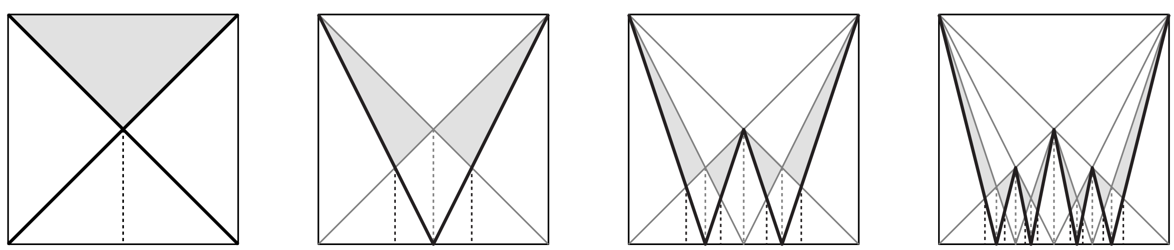
$$\frac{0}{1}\quad\frac{1}{7}\quad\frac{1}{6}\quad\frac{1}{5}\quad\frac{1}{4}\quad\frac{2}{7}\quad\frac{1}{3}\quad\frac{2}{5}\quad\frac{3}{7}\quad\frac{1}{2}\quad\frac{4}{7}\quad\frac{3}{5}\quad\frac{2}{3}\quad\frac{5}{7}\quad\frac{3}{4}\quad\frac{4}{5}\quad\frac{5}{6}\quad\frac{6}{7}\quad\frac{1}{1}$$

These numbers trace out the up-and-down path across the bottom of the preceding figure. For the next Farey series F_8 we would insert $1/8$ between $0/1$ and $1/7$, $3/8$ between $1/3$ and $2/5$, $5/8$ between $3/5$ and $2/3$, and finally $7/8$ between $6/7$ and $1/1$.

The mediant property of the Farey series F_n holds not just when each new term is added as described above, but in fact for every three consecutive terms of the series. For example in F_7, the mediant of $1/5$ and $2/7$ is $3/12 = 1/4$, so the mediant fraction must be reduced to lowest terms when the middle of the three denominators is not greater than the other two. This extended mediant property of Farey series will be deduced from a more general fact about mediants later in this section.

A more compact version of the preceding diagram that puts the part of the Farey diagram between 0 and 1 into a square is shown in the figure at the right. This can be constructed in stages as indicated in the sequence of figures below. Starting with a square, one first adds its diagonals and a vertical line from their intersection point down to the bottom edge of the square. The vertical line divides the region below the shaded triangle into two quadrilaterals. Each quadrilateral has one of its diagonals already present, and for the second stage of the construction we add the other diagonal and drop a vertical line from the intersection point of the two diagonals down to the bottom edge of the square. The process is then repeated for each subsequent step, adding a second diagonal in each unshaded quadrilateral and then a vertical line from the intersection point of the two diagonals down to the bottom edge of the square.



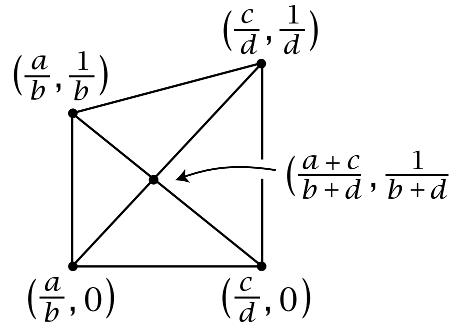


Let us choose the square to lie in in the upper halfplane with sides of length 1, with the bottom edge of the square along the x-axis and the lower left corner of the square at the origin. We then use the mediant rule to label the vertices of the shaded triangles as we proceed downward in the square, starting with the labels $^0\!/_1$ and $^1\!/_1$ at the upper left and right corners of the square. The positions of the vertices within the square can be described very simply:

- The vertex labeled $^a\!/_b$ is located at the point $(^a\!/_b, {}^1\!/_b)$.

This is obviously true for the vertices labeled $^0\!/_1$ and $^1\!/_1$ at the upper corners of the square, and also for the vertex labeled $^1\!/_2$ at the centerpoint $(^1\!/_2, {}^1\!/_2)$ of the square. For the remaining vertices we proceed by induction downward in the diagram. Each step of the induction is a special case of the following geometric characterization of mediants:

- For any two fractions a/b and c/d consider a quadrilateral in the xy-plane with vertices at the points shown in the figure at the right. Then the diagonals of the quadrilateral intersect at $(a+c/b+d, 1/b+d)$. Thus the mediant of a/b and c/d is the x-coordinate of the intersection point of the diagonals.



To verify this let us first show that $(a+c/b+d, 1/b+d)$ is on the diagonal from $(a/b, 0)$ to $(c/d, 1/d)$. To do this it suffices to show that the line segments from $(a/b, 0)$ to $(a+c/b+d, 1/b+d)$ and from $(a+c/b+d, 1/b+d)$ to $(c/d, 1/d)$ have the same slope. These slopes are

$$\frac{1/b+d}{a+c/b+d - a/b} = \frac{b}{b(a+c) - a(b+d)} = \frac{b}{bc - ad}$$

and

$$\frac{1/d - 1/b+d}{c/d - a+c/b+d} = \frac{b+d-d}{c(b+d) - d(a+c)} = \frac{b}{bc - ad}$$

so they are equal. The same argument works for the other diagonal by interchanging a/b and c/d. Thus the diagonals intersect at the point $(a+c/b+d, 1/b+d)$.

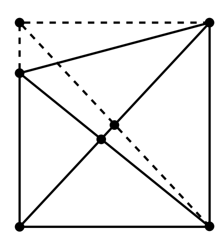
Note that the denominator $bc - ad$ in the slope formulas above is ± 1 when a/b and c/d are the endpoints of an edge of the Farey diagram. Thus each diagonal line in the square Farey diagram has integer slope, and this integer is, up to sign, the denominator of the rational number where the line meets the x-axis.

The fact that the y-coordinate of the vertex labeled a/b in the square diagram is $1/b$ implies that the successive Farey series can be obtained by taking the vertices that lie above the line $y = 1/2$, then the vertices above $y = 1/3$, then above $y = 1/4$, and so on. This explains why each new term inserted when F_n is enlarged to F_{n+1} is the mediant of its two neighbors in F_n. We see also that at most one new term of F_{n+1} is inserted between any two adjacent terms of F_n since there cannot be two triangles in the diagram with the same upper edge but different lower vertices.

From the geometric interpretation of mediants given above we can deduce a general fact about mediants:

- The mediant of two fractions a/b and c/d is closer to the fraction with larger denominator, unless the two denominators are equal in which case the mediant is halfway between the two fractions.

This can be seen by comparing the diagonals of the quadrilateral for a/b and c/d with the diagonals of the rectangle obtained by moving one of the upper two vertices of the quadrilateral vertically to the same height as the other upper vertex.



The following general fact justifies the earlier assertion that each term in the Farey series F_n is the mediant of the two adjacent terms.

- For rational numbers $a/b < c/d < e/f$, if there are edges in the Farey diagram joining a/b to c/d and c/d to e/f, then c/d is the mediant of a/b and e/f reduced to lowest terms.

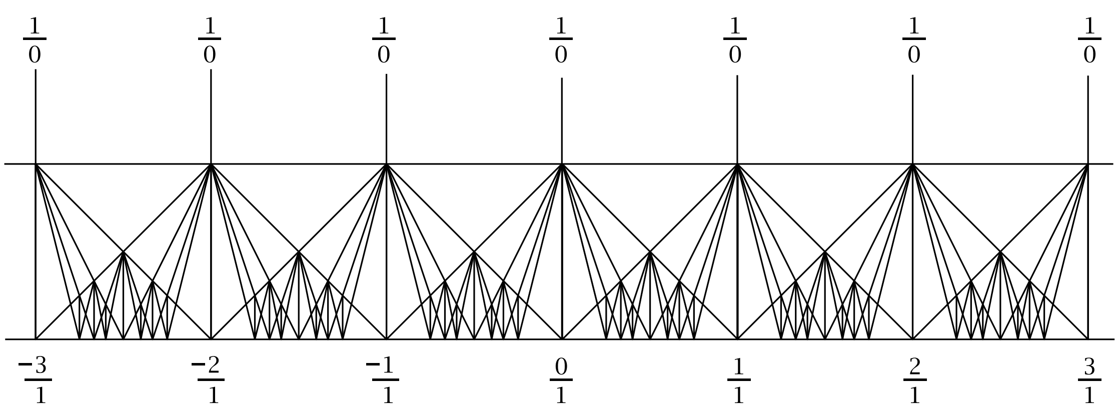
To see this we compute the mediant $a+e/b+f$. The assumption that there is an edge joining a/b and c/d means that $ad - bc = \pm 1$, so if $a/b < c/d$ we have $ad < bc$ so $ad - bc = -1$ and hence $bc - ad = 1$. Similarly, if there is an edge joining c/d to e/f we have $de - cf = 1$. From these equations we have:

$$a = ade - acf \qquad b = bde - bcf$$
$$e = bce - ade \qquad f = bcf - adf$$
$$\text{hence} \quad \frac{a+e}{b+f} = \frac{c(be-af)}{d(be-af)} = \frac{c}{d}$$

Note that in the last step, the fraction c/d is in lowest terms since we assumed c/d is a vertex of the Farey diagram, and the factor $be - af$ that we canceled to obtain c/d is, up to sign, the determinant of the matrix $\left(\begin{smallmatrix} a & e \\ b & f \end{smallmatrix}\right)$ associated to the pair of fractions a/b and e/f, and this determinant is ± 1 exactly when a/b and e/f are joined by an edge in the diagram.

As an example, all the fractions $m-1/m$ are connected to $1/1$ in the Farey diagram as are the fractions $n+1/n$ on the other side of $1/1$, and the mediant of $m-1/m$ and $n+1/n$ is $m+n/m+n = 1/1$. Here the number $be - af$ that is being canceled to get a fraction in lowest terms is $m(n+1) - (m-1)n = m + n$ which is the number of triangles in the Farey diagram between the edges from $1/1$ to $m-1/m$ and $n+1/n$.

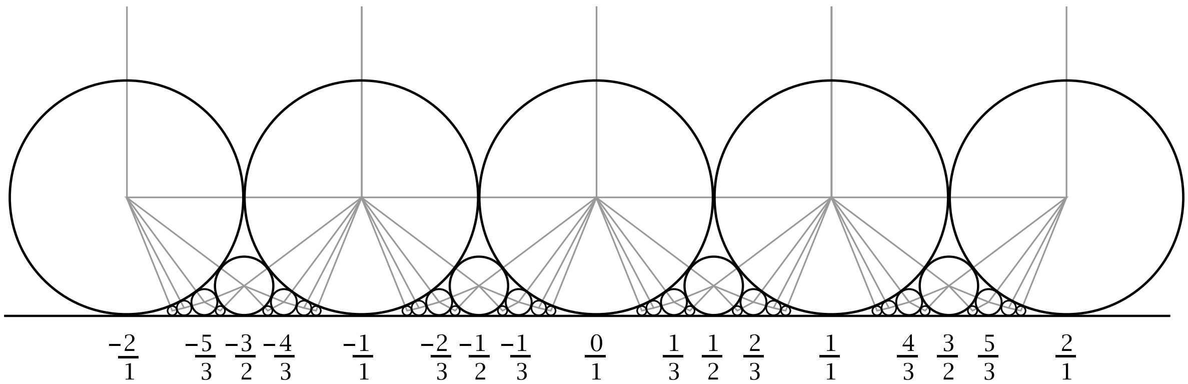
We can form a linear version of the full Farey diagram by placing copies of the square diagram we have been considering side by side along the x-axis:



Here the vertical segments in the horizontal strip of squares are not part of the resulting Farey diagram, which consists just of the triangles with nonvertical edges, along with the infinite “triangles” above the strip with a vertex at $1/0$. The original halfplane Farey diagram can be obtained from this linear Farey diagram by shrinking each vertical segment in the horizontal strip down to its lower endpoint while bending each straight edge of a triangle into a semicircle with endpoints on the x-axis.

Ford Circles

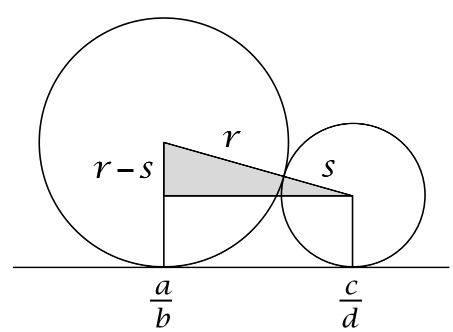
Another version of the Farey diagram can be constructed from an array of circles in the upper halfplane tangent to the x-axis and to each other as in the following figure:



This arrangement of tangent circles can be built in stages, starting with circles of diameter 1 tangent to the x-axis at the integer points. At the next stage a smaller circle is inserted in each gap between adjacent pairs of circles from the first stage. This creates new gaps, and one then puts a still smaller circle in each of these gaps. The process can then be repeated indefinitely all along the x-axis.

If we connect the centers of each pair of tangent circles by a line segment passing through the point of tangency, we obtain a pattern of triangles that is combinatorially equivalent to the pattern of triangles in the linear Farey diagram, but compressed closer to the x-axis. The vertices of these triangles are the centers of the various tangent circles, and we can label these centers by rational numbers, starting with an integer label $n/1$ at the center of the large circle tangent to the x-axis at the point n, and then labeling all the other centers by applying the mediant rule repeatedly.

The surprising thing about this construction is that the circle whose center is labeled a/b is tangent to the x-axis at exactly the point a/b on the x-axis. This can be verified as follows. For an edge of the Farey diagram with endpoints labeled a/b and c/d let us draw two circles tangent to each other and tangent to the x-axis at the points a/b and c/d. Let the radii of these two circles be r and s respectively. Note that r and s are not uniquely determined by a/b and c/d. In fact we can choose r arbitrarily and then this determines s, with s becoming small as r becomes large, and vice versa. We can find a formula for how r and s are related by applying the Pythagorean theorem to the right triangle



shown in the figure. The horizontal side of this triangle has length $|c/d - a/b|$ and the vertical side has length $|r - s|$. The condition for the two circles to be tangent is that the hypotenuse of the triangle has length $r + s$. Thus we have:

$$(r-s)^2 + \left(\frac{c}{d} - \frac{a}{b}\right)^2 = (r+s)^2$$

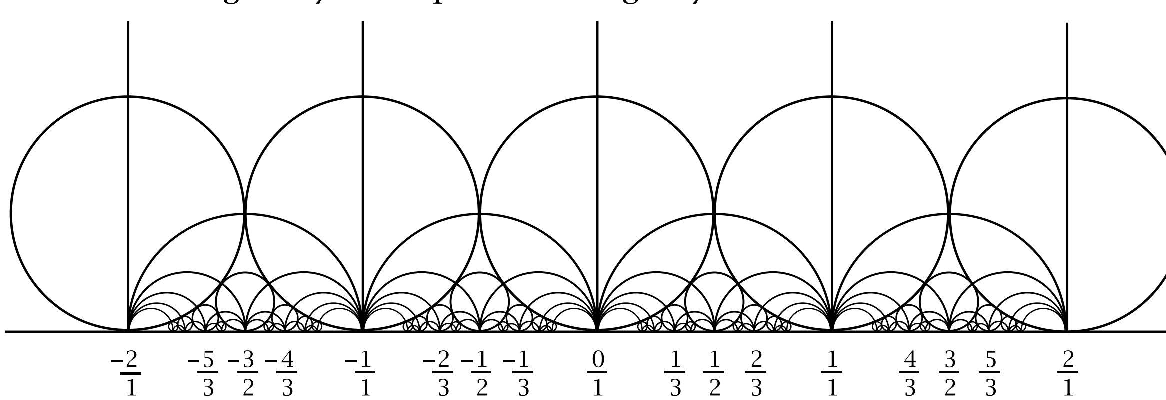
This simplifies to:

$$\left(\frac{bc - ad}{bd}\right)^2 = 4rs$$

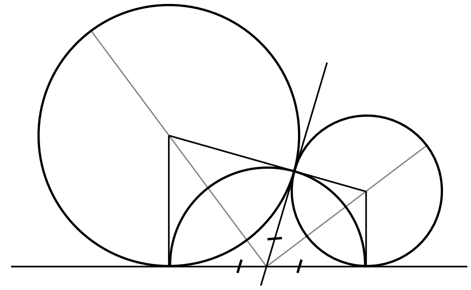
Since we assumed the fractions a/b and c/d were the endpoints of an edge in the Farey diagram, we have $ad - bc = \pm 1$ so the preceding equation simplifies further to $(\frac{1}{bd})^2 = 4rs$. The easiest way to assure that this holds is to let $r = 1/2b^2$ and $s = 1/2d^2$, so that r depends only on a/b and s depends only on c/d. Thus we are choosing the diameter of each circle to be the reciprocal of the square of the denominator of the fraction where the circle is tangent to the x-axis. This is consistent with how we chose the initial large circles tangent to the x-axis at integer points. Then when we build the Farey diagram inductively by adding more and more vertices labeled according to the mediant rule, each new vertex labeled $a+c/b+d$ between vertices labeled a/b and c/d is the center of a circle of diameter $1/(b+d)^2$ tangent to the x-axis at $a+c/b+d$ and tangent to each of the two circles labeled a/b and c/d of diameters $1/b^2$ and $1/d^2$ that are tangent to the x-axis at a/b and c/d.

The circles tangent to the x-axis constructed in this way are called *Ford circles* after their discoverer L. R. Ford. From the formula for their diameters we see that the Ford circles whose diameter is greater than a fixed number are just the ones associated to the fractions in a Farey series, if we restrict attention to the circles tangent to the x-axis at points between 0 and 1.

Another very nice feature of Ford circles is that when we superimpose them on the upper halfplane Farey diagram, the semicircles of the Farey diagram intersect the Ford circles orthogonally at the points of tangency of the Ford circles:



The fact that the circles and semicircles intersect orthogonally at the tangency points of the circles can be verified by considering the tangent lines to the circles at the points where two circles are tangent. The key fact is that for any two nonparallel tangent lines to a circle, the distances from the points of tangency to the intersection point of the two tangent lines are equal. This is because reflecting across the radial line through the intersection point takes one tangent line to the other.



Exercises

1. Compute the Farey series F_{10}.

2. Draw a figure showing how Ford circles are positioned in a circular Farey diagram by the following procedure. Start with a circle C of radius 1 which will be the outer boundary of the Farey diagram. Next, draw two tangent circles of radius $1/2$ inside C and tangent to C at two opposite points of C. Label these two tangency points $1/0$ and $0/1$. Now continue drawing smaller circles inside C with the same tangency patterns as the Ford circles in the upper halfplane Farey diagram, and label the tangency points of these circles with C according to the mediant rule. After a number of these circles have been drawn, superimpose the semicircles of the Farey diagram itself.

3. Suppose two Ford circles tangent to the x-axis at points a/b and c/d are tangent to each other. Show that the point of tangency between the two circles is the point

$$\left(\frac{ab+cd}{b^2+d^2}, \frac{1}{b^2+d^2}\right)$$

so in particular the coordinates of this point are rational. *Hint*: What proportion of the way along the line segment joining the two centers is the point of tangency? This same proportion will apply to x-coordinates and y-coordinates separately.

2 Continued Fractions

Continued fractions are expressions of the following sort:

$$\frac{7}{16} = \cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{2}}} \qquad \frac{67}{24} = 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{4}}}}$$

The numerators in these two examples are all 1, and we will only be considering continued fractions of this type, although there are situations outside the scope of this book where other numerators are allowed.

To compute the value of a continued fraction one starts in the lower right corner and works one's way upward. For example, in the continued fraction for $7/16$ one starts with $3 + 1/2 = 7/2$, then taking 1 over this gives $2/7$, and adding 2 to this gives $16/7$, and finally 1 over this gives $7/16$. In the case of the continued fraction for $67/24$ the fractions arising by this process are $5/4$, $4/5$, $19/5$, $5/19$, $24/19$, $19/24$, and finally $67/24$. As we will see, there is a fairly simple way to express every rational number as a continued fraction.

The main theme of this chapter will be the close relationship between continued fractions and the Farey diagram. For example, the fact that all rational numbers occur as labels on vertices in the Farey diagram is a reflection of the fact that every rational number has an expression as a continued fraction. In fact the continued fraction for a rational number p/q will tell how to locate the vertex labeled p/q in the diagram, and conversely, from the location of the vertex p/q one can read off the continued fraction for p/q.

We will also consider continued fractions with infinitely many terms extending downward to the right. These will give expressions for irrational numbers, somewhat like expressing irrational numbers as infinite decimals. Continued fractions have the advantage that rational numbers are expressible as finite continued fractions whereas the decimal representations for rational numbers are not generally finite but are instead just eventually periodic. Infinite continued fractions that are eventually periodic correspond to a special class of irrational numbers, those that are roots of quadratic equations with integer coefficients, like $\sqrt{2}$. Thus continued fractions are better than decimals in some ways, but on the other hand simple operations like addition and multiplication of rational numbers do not have nice descriptions in terms of continued fractions. In spite of these limitations continued fractions are quite useful in Number Theory. Among other things, they can be used to solve certain Diophantine equations including linear Diophantine equations, as we will see in Section 2.3.

2.1 Finite Continued Fractions

$$\frac{p}{q} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}}$$

The continued fractions we will be considering have the form shown at the right. The numbers a_i are assumed to be positive integers except for a_0 which can be any integer, possibly negative or zero. When a_0 is zero it can be omitted from the formula. To write a continued fraction in more compact form on a single line, we will often write it as ${}^p\!/_q = a_0 + {}^1\!\nearrow_{a_1} + {}^1\!\nearrow_{a_2} + \cdots + {}^1\!\nearrow_{a_n}$ with diagonal arrows to indicate the extended horizontal bars in the previous notation, for example ${}^7\!/_{16} = {}^1\!\nearrow_{2} + {}^1\!\nearrow_{3} + {}^1\!\nearrow_{2}$ and ${}^{67}\!/_{24} = 2 + {}^1\!\nearrow_{1} + {}^1\!\nearrow_{3} + {}^1\!\nearrow_{1} + {}^1\!\nearrow_{4}$. An even more concise notation that is sometimes used is $[a_0; a_1, a_2, \cdots, a_n]$, or just $[a_1, a_2, \cdots, a_n]$ when there is no a_0 term. However, we will use the more suggestive arrow notation in this book.

To compute the continued fraction for a given rational number, one starts in the upper left corner and works one's way downward, as the following example shows:

$$\frac{67}{24} = 2 + \frac{19}{24} = 2 + \frac{1}{24/19} = 2 + \frac{1}{1 + 5/19} = 2 + \cfrac{1}{1 + \cfrac{1}{19/5}}$$

$$= 2 + \cfrac{1}{1 + \cfrac{1}{3 + 4/5}} = 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{5/4}}} = 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{4}}}}$$

The key steps are the equations ${}^{67}\!/_{24} = 2 + {}^{19}\!/_{24}$, ${}^{24}\!/_{19} = 1 + {}^{5}\!/_{19}$, ${}^{19}\!/_{5} = 3 + {}^{4}\!/_{5}$, and ${}^{5}\!/_{4} = 1 + {}^{1}\!/_{4}$. If we clear fractions in each of these equations we obtain the first four of the five equations at the right which show a sequence of repeated divisions starting with a given pair of positive integers, 67 and 24 in this case. One first divides the smaller number into the larger to obtain a quotient and a remainder which is smaller than the divisor. Then at each successive step one divides the previous remainder into the previous divisor. The process stops when one obtains a remainder of zero. This process is known as the ***Euclidean algorithm***. The numbers in the shaded box are the quotients of the successive divisions and are sometimes called the *partial quotients*. These are the numbers a_i in the continued fraction for ${}^{67}\!/_{24}$.

$$\begin{array}{rcl}
67 &=& \boxed{2}\cdot 24 + 19 \\
24 &=& \boxed{1}\cdot 19 + 5 \\
19 &=& \boxed{3}\cdot 5 + 4 \\
5 &=& \boxed{1}\cdot 4 + 1 \\
4 &=& \boxed{4}\cdot 1 + 0
\end{array}$$

One of the classical uses for the Euclidean algorithm is to find the greatest common divisor of two given numbers. If one applies the algorithm to two numbers p and q, dividing the smaller into the larger, then the remainder into the first divisor, and so on, then the greatest common divisor of p and q turns out to be the last nonzero remainder. For example, starting with $p = 72$ and $q = 201$ the calculation is shown at the right, and the last nonzero remainder is 3, which is the greatest common divisor of 72 and 201. (In fact the fraction $^{201}\!/_{72}$ equals $^{67}\!/_{24}$, which explains why the successive quotients for this example are the same as in the preceding example.)

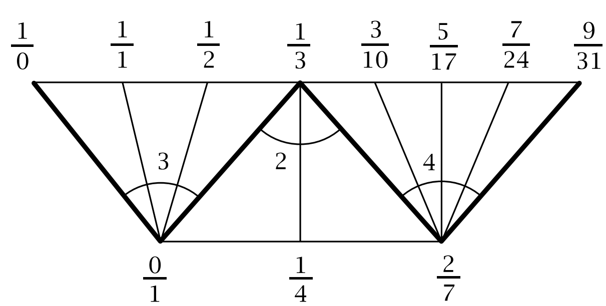
$$\begin{aligned}
201 &= 2 \cdot 72 + 57 \\
72 &= 1 \cdot 57 + 15 \\
57 &= 3 \cdot 15 + 12 \\
15 &= 1 \cdot 12 + 3 \\
12 &= 4 \cdot 3 + 0
\end{aligned}$$

It is easy to see from the displayed equations why 3 has to be the greatest common divisor of 72 and 201, since from the first equation it follows that any divisor of 72 and 201 must also divide 57, then the second equation shows it must divide 15, the third equation then shows it must divide 12, and the fourth equation shows it must divide 3, the last nonzero remainder. Conversely, if a number divides the last nonzero remainder 3, then the last equation shows it must also divide 12, and the next-to-last equation then shows it must divide 15, and so on until we conclude that it divides all the numbers not in the shaded rectangle, including the original two numbers 72 and 201. The same reasoning applies in general.

A more obvious way to try to compute the greatest common divisor of two numbers would be to factor each of them into a product of primes, then look to see which primes occurred as factors of both, and to what power. But to factor a large number into its prime factors is a very laborious and time-consuming process. For example, even a large computer would have a hard time factoring a number with a hundred or more digits into primes, so it would not be feasible to find the greatest common divisor of a pair of numbers of this size in this way. However, the computer would have no trouble applying the Euclidean algorithm to find their greatest common divisor.

Having seen what continued fractions are, let us now see what they have to do with the Farey diagram. Some examples will illustrate this best, so let us first look at the continued fraction for $^{9}\!/_{31}$ which is $1/\!\!\nearrow 3 + 1/\!\!\nearrow 2 + 1/\!\!\nearrow 4$. This has $3, 2, 4$ as its sequence of partial quotients, and we use these three numbers to build a strip of $3 + 2 + 4$ triangles grouped into “fans” of 3, 2, and 4 triangles:

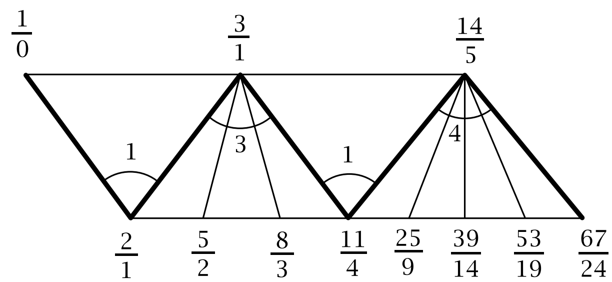
$$\frac{9}{31} = \cfrac{1}{3 + \cfrac{1}{2 + \cfrac{1}{4}}}$$



Now we begin labeling the vertices of this strip. On the left edge we start with the labels $^{1}\!/_{0}$ and $^{0}\!/_{1}$. Then we use the mediant rule for computing the third label of each triangle in succession as we move from left to right in the strip. Thus we insert, in order, the labels $1/1$, $1/2$, $1/3$, $1/4$, $2/7$, $3/10$, $5/17$, $7/24$, and finally $9/31$.

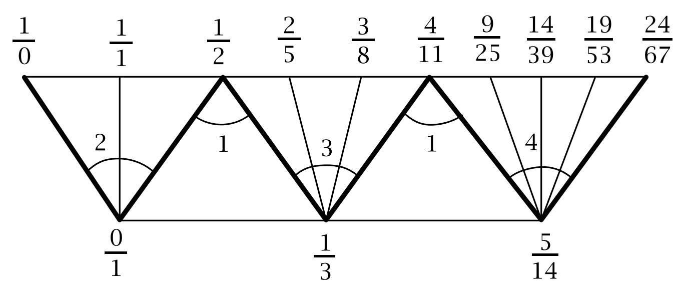
It may seem like just an accident that the final label is the fraction $9/31$ that we started with since the continued fraction for $9/31$ is computed from the equations $31/9 = 3 + 4/9$ and $9/4 = 2 + 1/4$ and these numbers have nothing to do with the fractions labeling the vertices along the strip before the final label $9/31$ miraculously appears. Nevertheless, we will see in Theorem 2.1 that what happened in this example always happens, at least for fractions p/q between 0 and 1. For fractions outside this interval the procedure works if we modify it by replacing the numerator 0 of the label $0/1$ with a_0, the initial integer in the continued fraction $p/q = a_0 + 1/a_1 + \cdots + 1/a_n$. Thus $0/1$ is replaced by $a_0/1$. This is illustrated by the $67/24$ example:

$$\frac{67}{24} = 2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{4}}}}$$

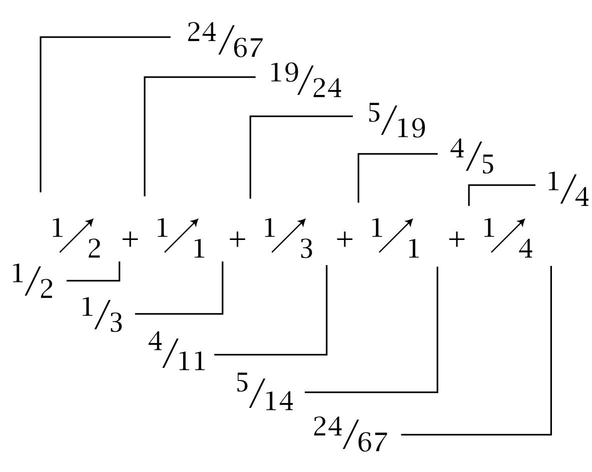


For comparison, here is the corresponding strip for the reciprocal, $24/67$:

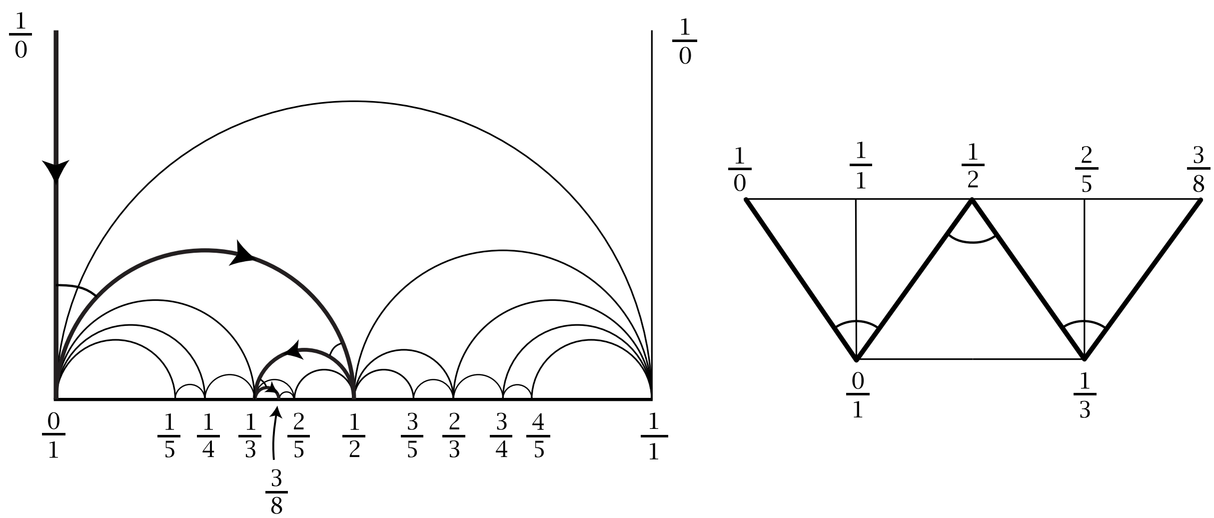
$$\frac{24}{67} = \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{4}}}}}$$



In the strip of triangles for a fraction p/q there is a zigzag path from $1/0$ to p/q that we have indicated by the heavily shaded edges. The labels on the vertices that this zigzag path passes through are the fractions that occur as the values of successively longer initial segments of the continued fraction, the continued fractions formed by the terms to the left of each plus sign in $a_0 + 1/a_1 + 1/a_2 + \cdots + 1/a_n$. This is illustrated at the right for the example of $24/67$. These fractions are called the ***convergents*** for the given fraction. Thus the convergents for $24/67$ are $1/2$, $1/3$, $4/11$, $5/14$, and $24/67$ itself. The figure also shows the values of the terminal segments, the terms to the right of each plus sign. These are the fractions one computes in order to find the value of the continued fraction.



It is interesting to see what the zigzag paths corresponding to continued fractions look like in the upper halfplane Farey diagram. The next figure shows the simple example of the continued fraction for $3/8$. We can see here that the five triangles of the strip correspond to the four curvilinear triangles lying directly above $3/8$ in the Farey diagram, plus the fifth "triangle" extending upward to infinity, bounded on the left and right by the vertical lines above $0/1$ and $1/1$, and bounded below by the semicircle from $0/1$ to $1/1$.



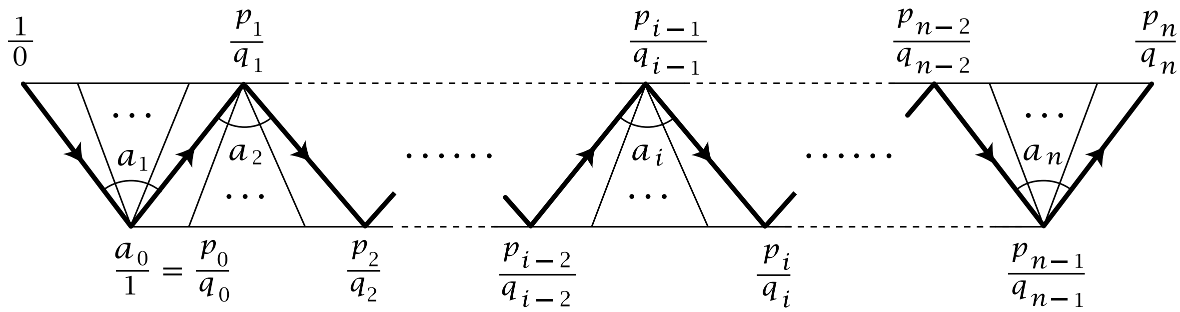
This example is typical of the general case, where the zigzag path for a continued fraction $p/q = a_0 + 1/a_1 + \cdots + 1/a_n$ becomes a "pinball path" in the Farey diagram, starting down the vertical line from $1/0$ to $a_0/1$, then turning left across a_1 triangles, then right across a_2 triangles, then left across a_3 triangles, continuing to alternate left and right turns until reaching the final vertex p/q. Two consequences of this are:

- The convergents are alternately smaller than and greater than p/q. The convergents to the left of p/q are getting successively closer to p/q from the left and the convergents to the right of p/q are getting successively closer to p/q from the right. We will see later in this section that in fact each convergent is closer to p/q than the previous one on the opposite side of p/q.
- The triangles that form the strip of triangles for p/q are exactly the triangles in the Farey diagram that lie directly above the point p/q on the x-axis. In other words, the strip of triangles for p/q consists of the triangles that the vertical line through the vertex p/q crosses.

Here is a general statement describing the relationship between continued fractions and the Farey diagram that we have observed in the preceding examples:

Theorem 2.1. *The convergents for a continued fraction $p/q = a_0 + 1/a_1 + \cdots + 1/a_n$ are the vertices along a zigzag path consisting of a finite sequence of edges in the Farey diagram, starting at $1/0$ and ending at p/q. The path starts along the edge from $1/0$ to $a_0/1$, then turns left across a fan of a_1 triangles, then right across a fan of a_2 triangles, etc., alternating left and right turns and finally ending at p/q.*

Proof: The continued fraction $p/q = a_0 + 1/a_1 + \cdots + 1/a_n$ determines a strip of triangles:



We will show that the label p_n/q_n on the final vertex in this strip is equal to p/q, the value of the continued fraction. Replacing n by i, we conclude that this holds also for each initial seqment $a_0 + 1/a_1 + \cdots + 1/a_i$ of the continued fraction. This is just saying that the vertices p_i/q_i along the strip are the convergents to p/q, which is what the theorem claims.

Each successive vertex label p_i/q_i along the zigzag path for the continued fraction $p/q = a_0 + 1/a_1 + \cdots + 1/a_n$ is computed in terms of the two preceding vertex labels according to the following formula:

$$\frac{p_i}{q_i} = \frac{a_i p_{i-1} + p_{i-2}}{a_i q_{i-1} + q_{i-2}}$$

This is because, as one can see in the figure above, the mediant rule is being applied a_i times, "adding" p_{i-1}/q_{i-1} to the previously obtained fraction each time until the next label p_i/q_i is obtained.

To prove that $p_n/q_n = p/q$ we will use 2×2 matrices. Consider the following product:

$$P = \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix}$$

We can multiply this product out starting either from the left or from the right. Suppose first that we multiply starting at the left. The two columns of the first matrix give the two fractions $1/0$ and $a_0/1$ labeling the left edge of the strip of triangles. Multiplying the first matrix by the second matrix gives:

$$\begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} = \begin{pmatrix} a_0 & 1 + a_0 a_1 \\ 1 & a_1 \end{pmatrix} = \begin{pmatrix} p_0 & p_1 \\ q_0 & q_1 \end{pmatrix}$$

The two columns here give the fractions at the ends of the second edge of the zigzag path. The same thing happens for subsequent matrix multiplications, as multiplying by the next matrix in the product takes the matrix corresponding to one edge of the zigzag path to the matrix corresponding to the next edge:

$$\begin{pmatrix} p_{i-2} & p_{i-1} \\ q_{i-2} & q_{i-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_i \end{pmatrix} = \begin{pmatrix} p_{i-1} & p_{i-2} + a_i p_{i-1} \\ q_{i-1} & q_{i-2} + a_i q_{i-1} \end{pmatrix} = \begin{pmatrix} p_{i-1} & p_i \\ q_{i-1} & q_i \end{pmatrix}$$

In the end, when all the matrices have been multiplied, we obtain the matrix corresponding to the last edge in the strip from p_{n-1}/q_{n-1} to p_n/q_n. Thus the second column of the product P is $\binom{p_n}{q_n}$, and what remains to show is that this equals $\binom{p}{q}$ where p/q is the value of the continued fraction $a_0 + 1/a_1 + \cdots + 1/a_n$.

The value of the continued fraction $a_0 + 1/a_1 + \cdots + 1/a_n$ is computed by working from right to left. If we let r_i/s_i be the value of the tail $1/a_i + 1/a_{i+1} + \cdots + 1/a_n$ of the continued fraction, then we have:

$$\frac{r_n}{s_n} = \frac{1}{a_n}, \qquad \frac{r_i}{s_i} = \frac{1}{a_i + \dfrac{r_{i+1}}{s_{i+1}}} = \frac{s_{i+1}}{a_i s_{i+1} + r_{i+1}}, \qquad \text{and} \qquad \frac{p}{q} = a_0 + \frac{r_1}{s_1} = \frac{a_0 s_1 + r_1}{s_1}$$

Expressed in terms of matrices these equations become:

$$\begin{pmatrix} r_n \\ s_n \end{pmatrix} = \begin{pmatrix} 1 \\ a_n \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & a_i \end{pmatrix}\begin{pmatrix} r_{i+1} \\ s_{i+1} \end{pmatrix} = \begin{pmatrix} s_{i+1} \\ r_{i+1} + a_i s_{i+1} \end{pmatrix} = \begin{pmatrix} r_i \\ s_i \end{pmatrix}$$

$$\text{and} \quad \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r_1 \\ s_1 \end{pmatrix} = \begin{pmatrix} r_1 + a_0 s_1 \\ s_1 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$$

This means that when we multiply out the product P starting from the right, the second columns will be successively $\binom{r_n}{s_n}$, $\binom{r_{n-1}}{s_{n-1}}$, $\cdots$, $\binom{r_1}{s_1}$, and finally $\binom{p}{q}$. We have already shown that the second column of P is $\binom{p_n}{q_n}$, so $p/q = p_n/q_n$ and the proof is complete. □

An interesting fact that can be deduced from the preceding proof is that for a continued fraction $1/a_1 + \cdots + 1/a_n$ with no initial integer a_0, if we reverse the order of the numbers a_i, this leaves the denominator unchanged. For example:

$$1/2 + 1/3 + 1/4 = \frac{13}{30} \qquad \text{and} \qquad 1/4 + 1/3 + 1/2 = \frac{7}{30}$$

To see why this must always be true we use the operation of transposing a matrix to interchange its rows and columns. For a 2×2 matrix this just amounts to interchanging the upper-right and lower-left entries, so the transpose of a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Transposing a product of matrices reverses the order of the factors, so one has $(AB)^T = B^T A^T$ as the reader can check by direct calculation. In the product

$$\begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & a_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix}$$

the individual matrices on the left side of the equation are symmetric with respect to transposition, so the transpose of the product is obtained by just reversing the order of the factors:

$$\begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & a_{n-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} = \begin{pmatrix} p_{n-1} & q_{n-1} \\ p_n & q_n \end{pmatrix}$$

Thus we see that reversing the order of the terms $a_1, \cdots, a_n$ leaves the denominator q_n unchanged, as claimed.

There is also a fairly simple relationship between the numerators. In the example of $13/30$ and $7/30$ we see that the product of the numerators, 91, is congruent to 1 modulo the denominator. In the general case the product of the numerators is

$p_n q_{n-1}$ and this is congruent to $(-1)^{n+1}$ modulo the denominator q_n. To verify this, we note that the determinant of each factor $\left(\begin{smallmatrix} 0 & 1 \\ 1 & a_i \end{smallmatrix}\right)$ is -1 so since the determinant of a product is the product of the determinants, we have $p_{n-1}q_n - p_n q_{n-1} = (-1)^n$, which implies that $p_n q_{n-1}$ is congruent to $(-1)^{n+1}$ modulo q_n.

Exercises

1. (a) Compute the values of the continued fractions ${}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{3} + {}^{1}\!\nearrow_{5} + {}^{1}\!\nearrow_{7}$ and ${}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{2}$.

(b) Compute the continued fraction expansions of ${}^{19}/_{44}$ and ${}^{101}/_{1020}$.

(c) Draw the strips of triangles corresponding to the continued fractions in parts (a) and (b).

2. (a) Compute the continued fraction for ${}^{38}/_{83}$ and display the steps of the Euclidean algorithm for 38 and 83 as a sequence of equations involving only integers.

(b) For the same number ${}^{38}/_{83}$ compute the associated strip of triangles grouped into fans, including the labeling of the vertices of all the triangles.

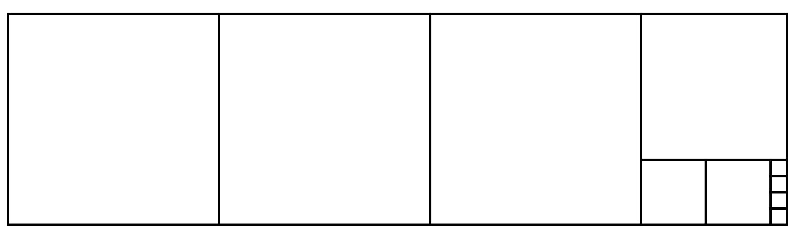
(c) Take the continued fraction ${}^{1}\!\nearrow_{a_1} + {}^{1}\!\nearrow_{a_2} + \cdots + {}^{1}\!\nearrow_{a_n}$ you got in part (a) and reverse the order of the numbers a_i to get a continued fraction ${}^{1}\!\nearrow_{a_n} + {}^{1}\!\nearrow_{a_{n-1}} + \cdots + {}^{1}\!\nearrow_{a_1}$. Compute the value ${}^{p}/_{q}$ of this continued fraction, and also compute the strip of triangles for this fraction ${}^{p}/_{q}$. What is the relationship between ${}^{p}/_{q}$ and ${}^{38}/_{83}$?

3. Let ${}^{p_n}/_{q_n}$ be the value of the continued fraction ${}^{1}\!\nearrow_{a_1} + {}^{1}\!\nearrow_{a_2} + \cdots + {}^{1}\!\nearrow_{a_n}$ where each of the n terms a_i is equal to 2. Thus ${}^{p_1}/_{q_1} = {}^{1}/_{2}$, ${}^{p_2}/_{q_2} = {}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{2} = {}^{2}/_{5}$, etc.

(a) Find equations expressing p_n and q_n in terms of p_{n-1} and q_{n-1}, and use these to write down the values of ${}^{p_n}/_{q_n}$ for $n = 1, 2, 3, 4, 5, 6, 7$.

(b) Compute the strip of triangles for ${}^{p_7}/_{q_7}$.

4. (a) A rectangle with sides of length 13 and 48 can be partitioned into squares in the way shown in the figure at the right. Determine the lengths of the sides of all the squares, and relate the numbers of squares of each size to the continued fraction for ${}^{13}/_{48}$.



(b) Draw the analogous figure decomposing a rectangle of sides 19 and 42 into squares, and relate this to the continued fraction for ${}^{19}/_{42}$.

5. This exercise is intended to illustrate the proof of Theorem 2.1 in the concrete case of the continued fraction ${}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{3} + {}^{1}\!\nearrow_{4} + {}^{1}\!\nearrow_{5}$.

(a) Write down the product $A_1A_2A_3A_4 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & a_1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & a_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & a_3 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & a_4 \end{smallmatrix}\right)$ associated to ${}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{3} + {}^{1}\!\nearrow_{4} + {}^{1}\!\nearrow_{5}$.

(b) Compute the four matrices A_1, A_1A_2, $A_1A_2A_3$, $A_1A_2A_3A_4$ and relate these to the edges of the zigzag path in the strip of triangles for ${}^1\!\nearrow_2 + {}^1\!\nearrow_3 + {}^1\!\nearrow_4 + {}^1\!\nearrow_5$.

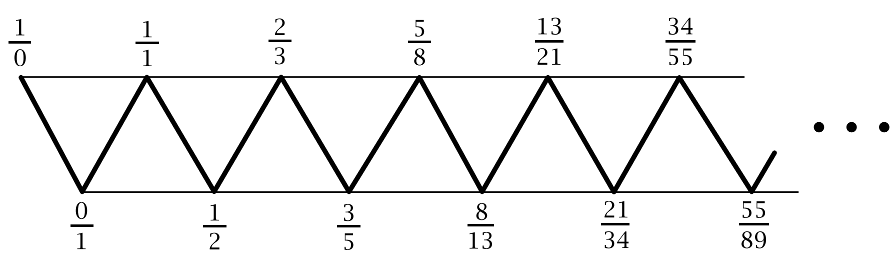
(c) Compute the four matrices A_4, A_3A_4, $A_2A_3A_4$, $A_1A_2A_3A_4$ and relate these to the successive fractions that one gets when one computes the value of ${}^1\!\nearrow_2 + {}^1\!\nearrow_3 + {}^1\!\nearrow_4 + {}^1\!\nearrow_5$, namely ${}^1\!\nearrow_5$, ${}^1\!\nearrow_4 + {}^1\!\nearrow_5$, ${}^1\!\nearrow_3 + {}^1\!\nearrow_4 + {}^1\!\nearrow_5$, and ${}^1\!\nearrow_2 + {}^1\!\nearrow_3 + {}^1\!\nearrow_4 + {}^1\!\nearrow_5$.

6. Compute the strip of triangles corresponding to the continued fraction for $7/19$ and compare this with the sequence of matrices reducing $\begin{pmatrix} 3 & 7 \\ 8 & 19 \end{pmatrix}$ to $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ by a sequence of operations subtracting one column from the other. (See the proof of Proposition 1.1.)

7. Show that the continued fraction for a rational number is unique except for replacing a final term ${}^1\!\nearrow_{a_n}$ by ${}^1\!\nearrow_{a_n-1} + {}^1\!\nearrow_1$ when $a_n > 1$. For example ${}^1\!\nearrow_3 + {}^1\!\nearrow_5 = {}^1\!\nearrow_3 + {}^1\!\nearrow_4 + {}^1\!\nearrow_1$.

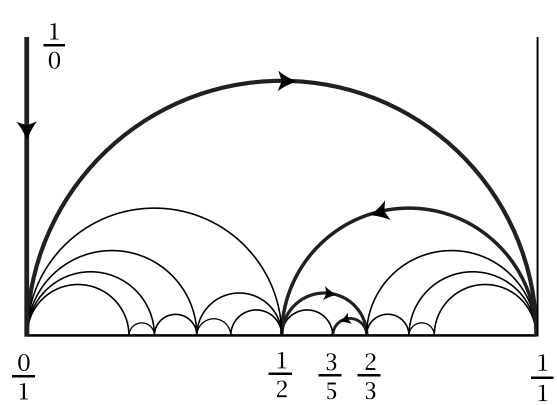
2.2 Infinite Continued Fractions

We have seen that all rational numbers can be expressed as continued fractions $a_0 + {}^1\!\nearrow_{a_1} + {}^1\!\nearrow_{a_2} + \cdots + {}^1\!\nearrow_{a_n}$. To complete the picture we will see that irrational numbers can be represented as continued fractions with an infinite number of terms, of the form $a_0 + {}^1\!\nearrow_{a_1} + {}^1\!\nearrow_{a_2} + {}^1\!\nearrow_{a_3} + \cdots$. A simple example is ${}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + \cdots$. This corresponds to an infinite strip of triangles in the Farey diagram:



Here the vertex labels along the zigzag path after the initial $1/0$ are the ratios of successive terms of the famous Fibonacci sequence $0, 1, 1, 2, 3, 5, 8, 13, 21, \cdots$ where each number after the initial 0 and 1 is the sum of its two predecessors.

The way the zigzag path looks in the upper halfplane Farey diagram is shown in the figure at the right. After the initial vertical edge from $1/0$ to $0/1$ this path consists of an infinite sequence of semicircles, each one shorter than the preceding one and sharing a common endpoint. The left endpoints of the semicircles form an increasing sequence of numbers which have to be approaching a



certain limiting value x. We know x has to be finite since it is certainly less than each of the right-hand endpoints of the semicircles, the convergents $1/1, 2/3, 5/8, \cdots$. Similarly, the right endpoints of the semicircles form a decreasing sequence of numbers approaching a limiting value y greater than each of the left-hand endpoints $0/1, 1/2, 3/5, \cdots$. Obviously $x \le y$. Is it possible that x is not equal to y? If this happened, the infinite sequence of semicircles would be approaching the semicircle from x to y. Above this semicircle there would then be an infinite number of semicircles, all the semicircles in the infinite sequence. Between x and y there would have to be a rational number p/q since there is always a rational number between any two real numbers, so above p/q there would be an infinite number of semicircles, hence an infinite number of triangles in the Farey diagram. But we know that there are only finitely many triangles above any rational number p/q, namely the triangles that appear in the strip for the continued fraction for p/q. This contradiction shows that x has to be equal to y. Thus the sequence of convergents along the edges of the infinite strip of triangles converges to a unique real number x.

This argument works for arbitrary infinite continued fractions, so we have shown the following general result:

Proposition 2.2. *For every infinite continued fraction* $a_0 + 1/\!\!\nearrow a_1 + 1/\!\!\nearrow a_2 + 1/\!\!\nearrow a_3 + \cdots$ *the convergents converge to a unique limit.*

This limit is by definition the value of the infinite continued fraction. This is similar to the situation for infinite decimals, where the value of an infinite decimal is the limit of the values of its finite initial segments.

As a complement to the preceding proposition we have:

Proposition 2.3. *Every irrational number has an expression as an infinite continued fraction, and this continued fraction is unique.*

Proof: In the upper halfplane Farey diagram consider the vertical line L going upward from a given irrational number x on the x-axis. The lower endpoint of L is not a vertex of the Farey diagram since x is irrational. Thus as we move downward along L we cross a sequence of triangles, entering each triangle by crossing its upper edge and leaving the triangle by crossing one of its two lower edges at a point between the two endpoints of this edge. When we exit one triangle, we are entering another triangle so the sequence of triangles and edges we cross must be infinite. The left and right endpoints of the edges in the sequence must be approaching the single point x by the argument we gave earlier, so the edges themselves are approaching x. It cannot happen that an infinite number of successive edges in the sequence have a common vertex since these edges would then be approaching this vertex, which would mean that x was rational. Thus the triangles crossed by the line L form an infinite strip consisting of an infinite sequence of fans with their pivot vertices on alternate sides of the strip. The zigzag path along this strip then gives a continued fraction for x.

For the uniqueness, we have seen that an infinite continued fraction for x corresponds to a zigzag path in the infinite strip of triangles lying above x. This set of triangles is unique so the strip is unique, and there is only one path in this strip that starts at $1/0$ and then does left and right turns alternately, starting with a left turn. The initial turn must be to the left because the first two convergents are a_0 and $a_0 + 1/a_1$, with $a_0 + 1/a_1 > a_0$ since $a_1 > 0$. After the path traverses the initial edge from $1/0$ to $a_0/1$ no subsequent edge of the path can be in the border of the strip since this would entail two successive left turns or two successive right turns. □

From the preceding arguments we can see fairly explicitly why the triangles in the upper halfplane Farey diagram completely cover the upper halfplane, so every point (x, y) with $y > 0$ lies either in the interior of some triangle or on the common edge between two triangles. To see this, consider the vertical line L in the upper halfplane through the given point (x, y). If x is an integer then (x, y) is on one of the vertical edges of the diagram, so we can assume x is not an integer and hence L is not one of the vertical edges of the diagram. The line L will then be contained in the strip of triangles corresponding to the continued fraction for x. This is a finite strip if x is rational and an infinite strip if x is irrational. In either case the point (x, y), being in L, will be in one of the triangles of the strip or on an edge separating two triangles in the strip.

Periodic and Eventually Periodic Continued Fractions

Now that we have an exact correspondence between infinite continued fractions and irrational numbers, there are two natural questions that come to mind: Given an infinite continued fraction, how can one compute its value, and conversely, how can one find the infinite continued fraction for a given irrational number? These questions have very nice answers for a special class of irrational numbers, the numbers whose continued fractions have a pattern that repeats periodically from some point onward, as for example:

$$1/\!\!\nearrow 2 + 1/\!\!\nearrow 4 + 1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7 + 1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7 + 1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7 + \cdots$$

This includes the case that the whole continued fraction is periodic, for example:

$$1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7 + 1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7 + 1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7 + \cdots$$

A more concise notation is to write a bar over a block of terms in a continued fraction that repeats infinitely often. Thus the two continued fractions above can be written as:

$$1/\!\!\nearrow 2 + 1/\!\!\nearrow 4 + \overline{1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7} \qquad \text{and} \qquad \overline{1/\!\!\nearrow 3 + 1/\!\!\nearrow 5 + 1/\!\!\nearrow 7}$$

The value of a periodic or eventually periodic continued fraction can be computed by simple algebraic manipulations, as we illustrate now by finding the value of the continued fraction $\overline{1/1} = 1/1 + 1/1 + 1/1 + \cdots$ involving Fibonacci numbers that we looked at earlier. Suppose we set $x = 1/1 + 1/1 + 1/1 + \cdots$. Then if we take the reciprocals of both sides of this equation we get $1/x = 1 + 1/1 + 1/1 + 1/1 + \cdots$. The right side of this equation is just $1 + x$, so we can easily solve for x:

$$\begin{aligned} &\frac{1}{x} = 1 + x \\ &x^2 + x - 1 = 0 \\ &x = (-1 \pm \sqrt{5})/2 \end{aligned}$$

We know x is positive, so this rules out the negative root and we are left with the final value $x = (-1 + \sqrt{5})/2$. The reciprocal $1/x = 1 + x = (1 + \sqrt{5})/2 \approx 1.618$ is known as the golden ratio because of its many interesting and beautiful properties.

As another example let us find the value of $1/3 + \overline{1/1 + 1/2}$. To do this we first find the value of the periodic part, so we set:

$$x = \overline{1/1 + 1/2} = 1/1 + 1/2 + 1/1 + 1/2 + 1/1 + 1/2 + \cdots$$

Taking reciprocals, we get:

$$\frac{1}{x} = 1 + 1/2 + 1/1 + 1/2 + 1/1 + 1/2 + \cdots$$

Subtracting 1 from both sides gives:

$$\frac{1}{x} - 1 = 1/2 + 1/1 + 1/2 + 1/1 + 1/2 + \cdots$$

The next step will be to take reciprocals of both sides, so before doing this we rewrite the left side as $1-x/x$. Then taking reciprocals gives:

$$\begin{aligned} \frac{x}{1-x} &= 2 + 1/1 + 1/2 + 1/1 + 1/2 + \cdots \\ &= 2 + x \end{aligned}$$

Thus we have $x/1-x = 2+x$ which simplifies to the quadratic equation $x^2+2x-2 = 0$ with roots $x = -1 \pm \sqrt{3}$. Again the negative root is discarded and we get $x = -1+\sqrt{3}$. From this we can determine the value of the original continued fraction $1/3 + \overline{1/1 + 1/2}$ which is $1/(3 + x) = 1/(2 + \sqrt{3}) = 2 - \sqrt{3}$.

Let us consider now the complementary question of how the continued fraction for a given irrational number can be computed. Recall first how the continued fraction $a_0 + 1/a_1 + 1/a_2 + \cdots + 1/a_n$ for a rational number is computed, as in the example of $67/24 = 2 + 1/1 + 1/3 + 1/1 + 1/4$ earlier in the chapter. We first write $67/24 = 2 + 19/24$ which gives $a_0 = 2$, then we write $24/19 = 1 + 5/19$ so $a_1 = 1$, then $19/5 = 3 + 4/5$ so $a_2 = 3$, then $5/4 = 1 + 1/4$ so $a_3 = 1$ and finally $4/1 = 4 + 0$ so $a_4 = 4$. This finishes the process and we have $67/24 = a_0 + 1/a_1 + 1/a_2 + 1/a_3 + 1/a_4 = 2 + 1/1 + 1/3 + 1/1 + 1/4$.

In summary, the steps are:

(1) Write the given number x as $x = a_0 + r_1$ where a_0 is an integer and $0 \le r_1 < 1$.
(2) Write $1/r_1$ as $1/r_1 = a_1 + r_2$ where a_1 is an integer and $0 \le r_2 < 1$.
(3) Write $1/r_2$ as $1/r_2 = a_2 + r_3$ where a_2 is an integer and $0 \le r_3 < 1$.

And so on, repeatedly.

If x is a rational number, the "remainders" r_i are rational numbers with decreasing denominators until we reach a remainder r_n which is zero and the process stops after finitely many steps. We can apply the same procedure if x is irrational, but in this case the equations defining the remainders r_i show that each successive r_i must be irrational and in particular nonzero. Thus the process goes on forever, yielding an infinite continued fraction.

One can see this is the continued fraction for x by the following argument. Suppose the continued fraction for x is $a_0 + 1/a_1 + 1/a_2 + \cdots$. We can write this continued fraction as $a_0 + r_1$ for $r_1 = 1/a_1 + 1/a_2 + \cdots$. This r_1 is a number strictly between 0 and 1 since the convergents for r_1 all lie between 0 and 1 and r_1 lies between any two of its successive convergents. Thus we have $x = a_0 + r_1$ with $0 < r_1 < 1$ so a_0 is the largest integer less than x. Inverting $r_1 = 1/a_1 + 1/a_2 + \cdots$ gives $1/r_1 = a_1 + 1/a_2 + 1/a_3 + \cdots$. The preceding argument can now be repeated with $1/r_1$ in place of x to get $1/r_1 = a_1 + r_2$ with $r_2 = 1/a_2 + 1/a_3 + \cdots$ and $0 < r_2 < 1$. Then one repeats with $1/r_2$ in place of $1/r_1$, and so on.

However, there are a couple subtle points in this argument that are somewhat hidden by the notation. (These subtle points were also lurking in the background in the earlier calculations of the values of the continued fractions $\overline{1/1}$ and $1/3 + \overline{1/1 + 1/2}$.) First, we defined x and r_1 to be the infinite continued fractions $a_0 + 1/a_1 + 1/a_2 + \cdots$ and $1/a_1 + 1/a_2 + \cdots$ and then said that $x = a_0 + r_1$. For finite continued fractions this is true because they are evaluated from right to left, so the last step in evaluating $a_0 + 1/a_1 + 1/a_2 + \cdots + 1/a_n$ is to add a_0 to $1/a_1 + 1/a_2 + \cdots + 1/a_n$. Infinite continued fractions cannot be evaluated from right to left since there is no right end to start the evaluation. Instead they are evaluated from left to right as the limit of the sequence of convergents. The convergents are the values of finite continued fractions, and for these the desired result holds so the convergents for $a_0 + 1/a_1 + 1/a_2 + \cdots$ are obtained by adding a_0 to the convergents for $1/a_1 + 1/a_2 + \cdots$. Adding a fixed number a_0 to each term of a convergent sequence of numbers adds a_0 to the limit of the sequence, so the result holds for infinite continued fractions as well as finite continued fractions.

A similar issue arises when we said that the continued fraction for the reciprocal $1/r_1$ of $r_1 = 1/a_1 + 1/a_2 + \cdots$ is $a_1 + 1/a_2 + \cdots$. Again this is correct for finite continued fractions since they are evaluated from right to left, so if one stops the evaluation of $1/a_1 + 1/a_2 + \cdots + 1/a_n$ before the last step of inverting $a_1 + 1/a_2 + \cdots + 1/a_n$ one has the reciprocal of $1/a_1 + 1/a_2 + \cdots + 1/a_n$. Thus the convergents for the infinite continued fraction $1/a_1 + 1/a_2 + \cdots$ are the reciprocals of the convergents for $a_1 + 1/a_2 + \cdots$ so the limits of the convergents for the two infinite continued fractions will also be reciprocals of each other.

Here is how the procedure works for computing the continued fraction for $\sqrt{2}$:

(1) $\sqrt{2} = 1 + (\sqrt{2} - 1)$ where $a_0 = 1$ since $\sqrt{2}$ is between 1 and 2. Thus $r_1 = \sqrt{2} - 1$.

(2) $1/r_1 = 1/\sqrt{2}-1 = 1/\sqrt{2}-1 \cdot \sqrt{2}+1/\sqrt{2}+1 = \sqrt{2}+1$ which is between 2 and 3 so we have $1/r_1 = 2 + (\sqrt{2} - 1)$ with $a_1 = 2$ and $r_2 = \sqrt{2} - 1$.

Notice that something unexpected has happened: The remainder $r_2 = \sqrt{2}-1$ is exactly the same as the previous remainder r_1. There is then no need to do the calculation of $1/r_2$ since we know it will have to be $\sqrt{2} + 1$. This means that when we continue with step (3), this will be exactly the same as step (2), and the same will be true for all subsequent steps. Thus we can immediately write down the continued fraction for $\sqrt{2}$:

$$\sqrt{2} = 1 + 1/2 + 1/2 + 1/2 + \cdots$$

We can check this calculation by finding the value of the continued fraction in the same way that we did earlier for $1/1 + 1/1 + 1/1 + \cdots$. It suffices to compute the value of $1/2 + 1/2 + 1/2 + \cdots$ and then add 1. We set $x = 1/2 + 1/2 + 1/2 + \cdots$ and then take reciprocals to get $1/x = 2 + 1/2 + 1/2 + 1/2 + \cdots = 2 + x$. From $1/x = 2 + x$ we get the quadratic equation $x^2 + 2x - 1 = 0$ with roots $x = -1 \pm \sqrt{2}$. Since x is positive we can discard the negative root. Thus we have $-1 + \sqrt{2} = 1/2 + 1/2 + 1/2 + \cdots$. Adding 1 to both sides of this equation gives the continued fraction for $\sqrt{2}$.

We can compute the continued fraction for $\sqrt{3}$ by the same method as for $\sqrt{2}$, but something slightly different happens:

(1) $\sqrt{3} = 1 + (\sqrt{3} - 1)$ with $a_0 = 1$ since $\sqrt{3}$ is between 1 and 2. Thus $r_1 = \sqrt{3} - 1$.

(2) $1/r_1 = 1/\sqrt{3}-1 = 1/\sqrt{3}-1 \cdot \sqrt{3}+1/\sqrt{3}+1 = \sqrt{3}+1/2$. This is between 1 and 2 since its numerator $\sqrt{3} + 1$ is between 2 and 3. Thus $a_1 = 1$ and $\sqrt{3}+1/2 = 1 + (\sqrt{3}-1/2)$ with $r_2 = \sqrt{3}-1/2$.

(3) $1/r_2 = 2/\sqrt{3}-1 = 2/\sqrt{3}-1 \cdot \sqrt{3}+1/\sqrt{3}+1 = \sqrt{3} + 1 = 2 + (\sqrt{3} - 1)$ with $a_2 = 2$ and $r_3 = \sqrt{3} - 1$.

Now the remainder $r_3 = \sqrt{3} - 1$ is the same as r_1, so instead of the same step being repeated infinitely often as happened for $\sqrt{2}$, the same two steps will repeat infinitely often. Thus we have computed the continued fraction for $\sqrt{3}$:

$$\sqrt{3} = 1 + 1/1 + 1/2 + 1/1 + 1/2 + 1/1 + 1/2 + \cdots$$

This agrees with our earlier calculation of the value of $\overline{1/1 + 1/2}$ to be $-1 + \sqrt{3}$.

It is true in general that for every positive integer n that is not a square, the continued fraction for $\sqrt{n}$ has the form $a_0 + \overline{1/a_1 + 1/a_2 + \cdots + 1/a_k}$. The length of the period (the repeating block) can be large even for fairly small values of n, for example:

$$\sqrt{46} = 6 + \overline{{}^1\!/_1 + {}^1\!/_3 + {}^1\!/_1 + {}^1\!/_1 + {}^1\!/_2 + {}^1\!/_6 + {}^1\!/_2 + {}^1\!/_1 + {}^1\!/_1 + {}^1\!/_3 + {}^1\!/_1 + {}^1\!/_{12}}$$

This example illustrates two other curious facts about the continued fraction for an irrational number $\sqrt{n}$:

- The last term of the period (12 in the example above) is always twice the first term a_0 (the initial 6).
- If the last term of the period is omitted, the preceding terms in the period form a palindrome, reading the same backwards as forwards.

We will see in Section 4.2 how these two properties follow from certain symmetry properties of the infinite strip of triangles in the Farey diagram associated to the continued fraction for $\sqrt{n}$.

It is natural to ask exactly which irrational numbers have continued fractions that are periodic or eventually periodic. The answer is given by:

Lagrange's Theorem. *The irrational numbers whose continued fractions are eventually periodic are exactly the numbers of the form $a + b\sqrt{n}$ where a and b are rational numbers, $b \neq 0$, and n is a positive integer that is not a square.*

These numbers $a + b\sqrt{n}$ are called *quadratic irrationals* because they are roots of quadratic equations with integer coefficients. The easier half of the theorem is the statement that the value of an eventually periodic infinite continued fraction is always a quadratic irrational. This can be proved by showing that the method we used for finding a quadratic equation satisfied by an eventually periodic continued fraction works in general. Rather than following this purely algebraic approach, however, we will develop a more geometric version of the procedure in the next chapter, so we will wait until then to give the argument that proves this half of Lagrange's Theorem, in Proposition 3.4. The more difficult half of the theorem is the assertion that the continued fraction expansion of every quadratic irrational is eventually periodic. It is not at all apparent from the examples of $\sqrt{2}$ and $\sqrt{3}$ why this should be true in general, but in Chapters 4 and 5 we will develop some theory that will make it clear, with the actual proof being given in Proposition 4.1 and Theorem 5.2. Along the way we will also develop more efficient methods for computing the continued fraction for a quadratic irrational and for computing the value of an eventually periodic infinite continued fraction.

What can be said about the continued fraction expansions of irrational numbers that are not quadratic, such as $\sqrt[3]{2}$, π, or e, the base for natural logarithms? It happens that e has a continued fraction whose terms have a very nice pattern, even though they are not periodic or eventually periodic:

$$e = 2 + \underbrace{{}^1\!/_1 + {}^1\!/_2 + {}^1\!/_1} + \underbrace{{}^1\!/_1 + {}^1\!/_4 + {}^1\!/_1} + \underbrace{{}^1\!/_1 + {}^1\!/_6 + {}^1\!/_1} + \cdots$$

Thus the terms are grouped by threes with successive even numbers as middle denominators. Even simpler are the continued fractions for certain numbers built from e that have arithmetic progressions for their denominators:

$$\frac{e-1}{e+1} = {}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{6} + {}^{1}\!\nearrow_{10} + {}^{1}\!\nearrow_{14} + \cdots$$

$$\frac{e^2-1}{e^2+1} = {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{3} + {}^{1}\!\nearrow_{5} + {}^{1}\!\nearrow_{7} + \cdots$$

The continued fractions for e and $(e-1)/(e+1)$ were discovered by Euler in 1737 while the formula for $(e^2-1)/(e^2+1)$ was found by Lambert in 1766 as a special case of a slightly more complicated formula for $(e^x-1)/(e^x+1)$.

For $\sqrt[3]{2}$ and π, however, the continued fractions have no known pattern. For π the continued fraction begins:

$$\pi = 3 + {}^{1}\!\nearrow_{7} + {}^{1}\!\nearrow_{15} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{292} + \cdots$$

Here the first four convergents are 3, ${}^{22}\!/_{7}$, ${}^{333}\!/_{106}$, and ${}^{355}\!/_{113}$. We recognize ${}^{22}\!/_{7}$ as the familiar approximation $3\,{}^{1}\!/_{7}$ to π. The convergent ${}^{355}\!/_{113}$ is a particularly good approximation to π since its decimal expansion begins 3.14159282 whereas $\pi = 3.14159265\cdots$. It is no accident that the convergent ${}^{355}\!/_{113}$ obtained by truncating the continued fraction just before the 292 term gives a good approximation to π since it is a general fact that a convergent immediately preceding a large term in the continued fraction always gives an especially good approximation. This is because the next edge in the zigzag path will be rather small when viewed in the upper halfplane Farey diagram since it is the lower edge of a fan with a large number of triangles, and the value of the continued fraction lies somewhere between the two ends of this small edge.

It is easy to calculate an initial string of terms in the continued fraction for π using a reasonably capable calculator if one knows the decimal expansion of π with enough accuracy. One just repeats the two steps of subtracting the integer part and inverting, preferably using the calculator's $1/x$ button. For example, starting with 3.1415926535 one can get the initial segment $3 + {}^{1}\!\nearrow_{7} + {}^{1}\!\nearrow_{15} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{292}$ this way. People who like computational challenges have used computers to find large numbers of terms of the continued fraction for π, more than a billion terms in fact.

There are nice continued fractions for π if one allows numerators larger than 1, as in the following formula discovered by Euler:

$$\pi = 3 + {}^{1^2}\!\nearrow_{6} + {}^{3^2}\!\nearrow_{6} + {}^{5^2}\!\nearrow_{6} + {}^{7^2}\!\nearrow_{6} + \cdots$$

However, it is the continued fractions with numerator 1 that have the best properties, so we will not consider the more general sort in this book.

Convergents as Rational Approximations

Let us explore in a little more detail how the convergents in the continued fraction for an irrational number x give good rational approximations to x.

As an example, consider the case $x = \sqrt{2} = 1 + \overline{{}^1\!\!\nearrow_2}$. It is a little easier to compute the convergents for $2 + \overline{{}^1\!\!\nearrow_2} = 1 + \sqrt{2}$ and then subtract 1 from each of these. For $2 + {}^1\!\!\nearrow_2 + {}^1\!\!\nearrow_2 + {}^1\!\!\nearrow_2 + \cdots$ the convergents are:

$$\frac{2}{1} \quad \frac{5}{2} \quad \frac{12}{5} \quad \frac{29}{12} \quad \frac{70}{29} \quad \frac{169}{70} \quad \frac{408}{169} \quad \frac{985}{408} \quad \cdots$$

The sequence of numbers $1, 2, 5, 12, 29, 70, 169, \cdots$ generating these fractions can be constructed in a way somewhat analogous to the Fibonacci sequence, except that each number is *twice* the preceding number plus the number before that. This is because each convergent is obtained from the previous one by inverting the fraction and adding 2, and therefore the next convergent after ${}^a/_b$ is $2 + {}^b/_a = {}^{2a+b}/_a$.

Subtracting 1 from each of the fractions in the list displayed above, we obtain the convergents for $\sqrt{2}$ as shown at the right. Notice that once an initial string of digits in the decimal expansions of the convergents occurs twice in succession, then this string is unchanged from then on. This is because for any two successive convergents, all subsequent convergents lie between these two since the convergents occur along a zigzag path in the Farey diagram. This is true generally for all infinite continued fractions.

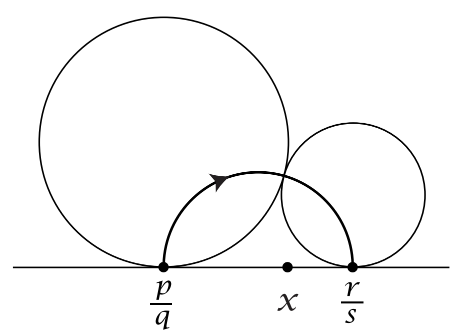
$$\begin{aligned}
\sqrt{2} &= 1.41421356\cdots \\
{}^1/_1 &= 1.00000000\cdots \\
{}^3/_2 &= 1.50000000\cdots \\
{}^7/_5 &= 1.40000000\cdots \\
{}^{17}/_{12} &= 1.41666666\cdots \\
{}^{41}/_{29} &= 1.41379310\cdots \\
{}^{99}/_{70} &= 1.41428571\cdots \\
{}^{239}/_{169} &= 1.41420118\cdots \\
{}^{577}/_{408} &= 1.41421568\cdots
\end{aligned}$$

Information about how well an irrational number is approximated by the convergents in its continued fraction can be deduced from the geometry of Ford circles, which were introduced at the end of Chapter 1. Here is one general statement that can be made:

- Each convergent ${}^p/_q$ in the continued fraction for an irrational number x is within ${}^1/_{q^2}$ of x, so $|x - {}^p/_q| < {}^1/_{q^2}$.

For example, if a convergent has a denominator of 100 or greater then the convergent approximates x to within ${}^1/_{10000}$. Thus the approximation ${}^{239}/_{169}$ to $\sqrt{2}$ must be accurate to four decimal places.

To justify the ${}^1/_{q^2}$ estimate, suppose the convergent ${}^p/_q$ is connected to the next convergent ${}^r/_s$ by an edge of the zigzag path. We then have $s \ge q$ so the Ford circle $C_{p/q}$ at ${}^p/_q$, which has diameter ${}^1/_{q^2}$, is at least as large as the Ford circle $C_{r/s}$. The number x lies between ${}^p/_q$ and ${}^r/_s$, so its distance to ${}^p/_q$ is less than twice the radius ${}^1/_{2q^2}$ of $C_{p/q}$. In other words this distance is less than ${}^1/_{q^2}$, as claimed.



For many convergents the estimate $1/q^2$ can be improved by a factor of 2:

- At least one convergent p/q out of every two successive convergents to x is within $1/2q^2$ of x.

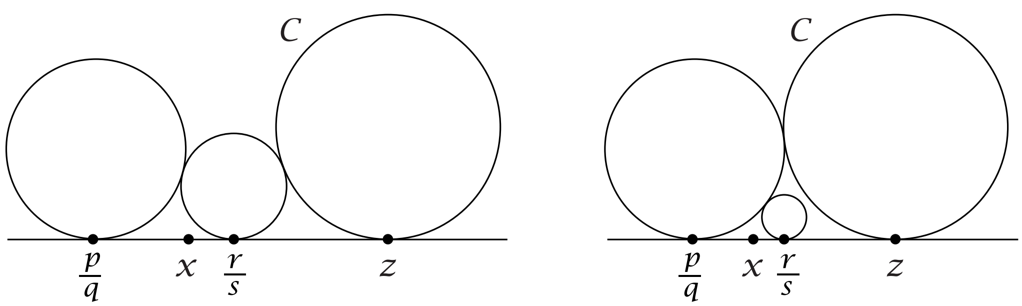
We can see this from the previous figure. For x to be within $1/2q^2$ of p/q means that x is a point in the projection of the interior of $C_{p/q}$ to the x-axis, and similarly for r/s and $C_{r/s}$. Since x lies between p/q and r/s, it must be in at least one of these two projections, except possibly in the case that $q = s$ (which can only happen when q and s are 1) when the midpoint of the interval between p/q and r/s is not in the projection of the interior of either $C_{p/q}$ or $C_{r/s}$. But this midpoint is a rational number so it cannot be x.

Next we have a very strong optimality statement about the convergents to an irrational number x:

- If p/q is a convergent in the continued fraction for x then no rational number with denominator less than or equal to q is closer to x than p/q is.

To see why this is true consider two consecutive convergents p/q and r/s as before, and let t/u be any rational number with $u \le q$, so $C_{t/u}$ is at least as large as $C_{p/q}$. The circle $C_{t/u}$ is either disjoint from or tangent to $C_{p/q}$ and $C_{r/s}$. Clearly t/u cannot be between p/q and r/s since there is no room to fit the large circle $C_{t/u}$ there. If t/u is on the opposite side of p/q from r/s then t/u would be farther from x than p/q is, so t/u would not be a closer approximation to x than p/q is.

The remaining possibility is that t/u is on the opposite side of r/s from p/q. Let C be any circle in the upper halfplane with the same geometric properties as $C_{t/u}$, namely, C is tangent to the x-axis at a point z on the opposite side of r/s from p/q, C is either disjoint from or tangent to $C_{p/q}$ and $C_{r/s}$, and C is at least as large as $C_{p/q}$.

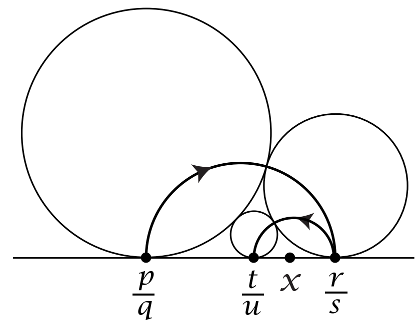


We wish to show that z is farther from x than p/q is. Sliding C along the x-axis farther from r/s moves z farther from x so we may assume C touches either $C_{p/q}$ or $C_{r/s}$. Then increasing the size of C while keeping it tangent to $C_{p/q}$ or $C_{r/s}$ also moves z farther from x so we may assume C has the same size as $C_{p/q}$. It is then evident that z is farther from x than p/q is, finishing the argument.

The last fact we will deduce from the diagram of Ford circles is that the convergents converge monotonically. We know that the convergents to the left of x are getting steadily closer to x, and the same is true for the convergents to the right. And in fact:

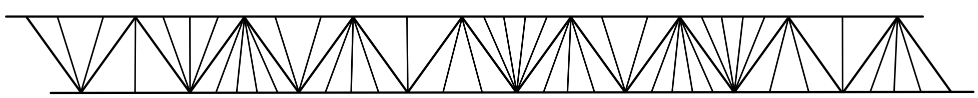
- Each convergent in the continued fraction for x is closer to x than the previous convergent.

To verify this, suppose that p/q and r/s are two consecutive convergents to x, so we wish to show that x is closer to r/s than to p/q. Consider the next convergent t/u after r/s. The circle $C_{r/s}$ is tangent to both $C_{p/q}$ and $C_{t/u}$, while $C_{t/u}$ is either tangent to $C_{p/q}$ or $C_{t/u}$ is one of the other Ford circles tangent to $C_{r/s}$ farther from $C_{p/q}$. The point x lies between r/s and t/u so to show that x is closer to r/s than to p/q it will suffice to consider just the case that $C_{t/u}$ is tangent to $C_{p/q}$. Then t/u is the mediant of p/q and r/s so as we saw in Section 1.2, t/u is closer to r/s than to p/q since $s \ge q$, or possibly t/u is equidistant from r/s and p/q if $s = q$. In either case, since x lies between t/u and r/s, it must then be closer to r/s than to p/q, which is what we wanted to show.



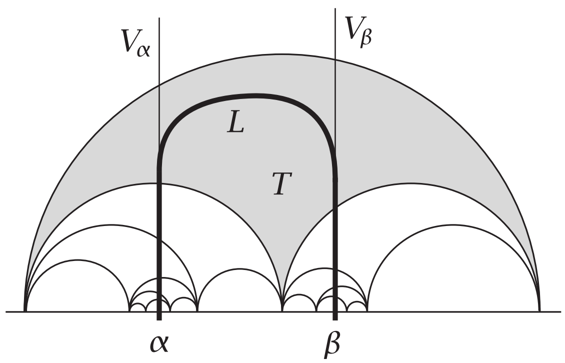
Doubly Infinite Strips

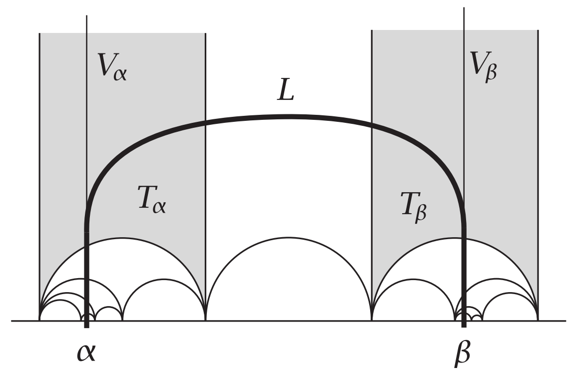
We have been considering strips of triangles in the Farey diagram consisting of fans, each fan having a finite number of triangles and each fan intersecting the next along an edge of a zigzag path in the strip. For finite continued fractions the strip has finitely many fans, while for infinite continued fractions the strip has an infinite sequence of fans at one end. In later chapters we will often be considering strips that extend infinitely far at both ends. We can think of these strips as being “doubly infinite” since they are infinite in both directions.



To see how such a doubly infinite strip lies in the upper halfplane model of the Farey diagram, let L be a line running down the middle of the strip from end to end. Viewing L as a path in the upper halfplane model of the Farey diagram, L cannot cross only vertical edges, the edges with one end at $1/0$, otherwise the strip would consist of a single infinite fan, which is not allowed as an infinite strip. Thus L must cross some semicircular edges. As we move along L crossing such a semicircular edge in the downward direction into the adjacent triangle, the next edge that L crosses will be one of the other two shorter semicircular edges of this triangle, moving downward again. All subsequent crossings will then be downward as well. The semicircles crossed are becoming smaller and smaller with diameters approaching zero, as we saw in our initial discussion of infinite continued fractions, and there is a unique limiting point α on the x-axis for this end of the strip of triangles. This is the unique point that lies between the two endpoints of each semicircular edge crossed by L on its downward path.

Consider the vertical line V_α going upward from α. Near its lower end V_α will pass through triangles of the strip. If the whole line V_α does not stay entirely within the strip as we move upward, it will eventually leave the strip by crossing the upper semicircular edge of a triangle T of the strip as in the figure on the left below.





In this case the line L, which passes through the same upward sequence of triangles as V_α until reaching T, must exit T by turning and crossing the other smaller semicircular edge of T in the downward direction. After crossing this edge, L will then continue downward forever, passing through all the triangles of the other end of the strip and limiting on an irrational number β. The vertical line V_β going upward from β will pass through the same set of triangles until reaching the triangle T where it will also exit the strip by crossing the upper edge of T. We can then deform L so that it consists of the parts of V_α and V_β below T joined by a bending arc within T. Notice that the vertex $1/0$ is not a vertex of the strip in this case.

The other possibility is that V_α stays in the strip forever as we move upward, so eventually it lies in a triangle T_α of the strip having $1/0$ as a vertex as in the figure on the right above. One end of the line L runs parallel to V_α until it reaches T_α, then it turns right or left to cross a finite number of other triangles having $1/0$ as a vertex before turning downward to cross the lower edge of one of these triangles T_β. After this it will travel monotonically downward, limiting on an irrational number β in the x-axis. We can deform L to consist of parts of V_α and the vertical line V_β through β, joined by an arc crossing from T_α to T_β.

One conclusion we can draw from this analysis of the infinite strip is that its endpoints α and β cannot be the same number. This can be seen from the two figures above where in the first figure α and β lie below the two different lower edges of the triangle T, and in the second figure α and β lie below the two different triangles T_α and T_β with a vertex at $1/0$.

Another consequence is that the labels x/y on the vertices along the infinite strip must have denominators y approaching infinity at the ends of the strip and numerators x approaching either $+\infty$ or $-\infty$ depending on the sign of the endpoint α or β being approached. This is because the labels are given by repeated applications of the mediant rule as we move vertically down either end of L toward α or β so $|x|$ and y always increase as each new triangle is added to the strip. (Near the ends of the strip the labels x/y are approaching α or β so neither x nor y is 0.)

We can also deduce that for each pair of distinct irrationals α and β there is a unique infinite strip in the Farey diagram whose ends converge to α and β. This is because α and β determine the vertical lines V_α and V_β in the figures, and these determine the triangles T or T_α and T_β since in the case that α and β lie in the same interval in the x-axis between consecutive integers, T is the smallest triangle of the Farey diagram whose projection to the x-axis contains both α and β, while in the case that α and β lie in different intervals between consecutive integers, the triangles T_α and T_β are the triangles with vertex $1/0$ that project to these two intervals.

A nice way to construct an infinite strip joining any two irrationals α and β is to take all the triangles in the Farey diagram that meet the semicircle in the upper halfplane with endpoints α and β. This semicircle can cross an edge of the Farey diagram only once since if two semicircles in the upper halfplane with endpoints on the x-axis intersect in more than one point, they must coincide. Nor can two semicircles with endpoints on the x-axis be tangent unless the point of tangency is one of the endpoints, but this does not happen here since α and β are irrational while the endpoints of edges of the Farey diagram are rational. From these observations we see that if the semicircle from α to β intersects a triangle of the Farey diagram, then it crosses this triangle from one edge to another edge. The semicircle cannot cross an infinite number of triangles having a common vertex, otherwise the semicircle would contain points arbitrarily close to the common vertex, which is impossible since the common vertex cannot be either of the irrational numbers α and β. Thus the union of all the triangles crossed by the semicircle is an infinite strip.

We have seen that an infinite strip is uniquely determined by its endpoints, so this implies that the semicircle from α to β crosses exactly the same triangles as the line we constructed earlier consisting of two vertical segments joined at the top by a 180 degree bend. This may seem odd at first glance, but what it means is that the height of the vertical segments cannot be too large compared to the distance between them.

The construction of a strip connecting two irrational numbers α and β via the semicircle with endpoints α and β works equally well when α or β is rational, but in this case the strip has only a finite number of triangles at a rational end. A very special case is when α and β are the endpoints of an edge of the Farey diagram, when the strip degenerates to just this edge.

The doubly infinite strips we will be most interested in are the ones that are periodic along their whole length. As we will see, the irrational numbers α and β

at the ends of such a strip will be the two roots of a quadratic equation with integer coefficients.

Exercises

1. Compute the values of the following infinite continued fractions:

(a) $\overline{{}^{1}\!\nearrow_{4}}$

(b) $\overline{{}^{1}\!\nearrow_{n}}$ for an arbitrary positive integer n

(c) $\overline{{}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{3}}$ and ${}^{1}\!\nearrow_{1} + \overline{{}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{3}}$

(d) $\overline{{}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{6}}$ and ${}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{4} + \overline{{}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{1} + {}^{1}\!\nearrow_{6}}$

(e) $\overline{{}^{1}\!\nearrow_{2} + {}^{1}\!\nearrow_{3} + {}^{1}\!\nearrow_{5}}$

2. (a) Compute the continued fractions for $\sqrt{5}$ and $\sqrt{23}$.

(b) Using the continued fraction for $\sqrt{5}$, find the first convergent which gives a rational approximation to $\sqrt{5}$ accurate to four decimal places.

3. Compute the continued fractions for $\sqrt{n^2+1}$ and $\sqrt{n^2+n}$ where n is an arbitrary positive integer.

2.3 Linear Diophantine Equations

As an application of continued fractions let us see how they can be used to solve linear Diophantine equations $ax + by = n$, where a, b, and n are integers and the solutions are required to be integers as well. We can assume a, b, and n are nonzero, otherwise the equation is rather trivial. Changing the signs of x or y if necessary, we can rewrite the equation in the form $ax - by = n$ where a and b are both positive. Solving this equation means finding multiples of a and b that differ by n.

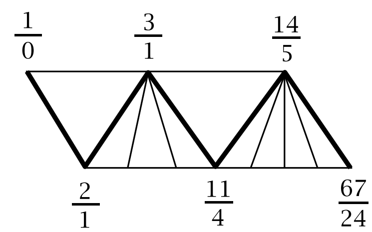
If a and b have greatest common divisor $d > 1$, then since d divides a and b it must divide $ax - by$, so d must divide n if the equation $ax - by = n$ is to have any solutions at all. If d does divide n we can divide both sides of the equation by d to get a new equation having the same solutions but with the new coefficients a and b coprime. For example, the equation $6x - 15y = 21$ reduces in this way to the equation $2x - 5y = 7$. Thus we can assume from now on that a and b are coprime. We will show that solutions always exist in this case, in fact infinitely many solutions, and we will see how to compute them.

To find a solution of $ax - by = n$ it suffices to do the case $n = 1$ since if we have a solution of $ax - by = 1$, we can multiply x and y by n to get a solution of $ax - by = n$. For example, for the equation $2x - 5y = 1$ the smallest multiple of 2 that is one greater than a multiple of 5 is $2\cdot 3 > 5\cdot 1$, so a solution of $2x - 5y = 1$ is $(x, y) = (3, 1)$. A solution of $2x - 5y = 7$ is then $(x, y) = (21, 7)$.

The idea for solving $ax - by = 1$ when a and b are coprime is to utilize the criterion from Proposition 1.1 that the Farey diagram contains an edge joining ${}^a/_b$ to ${}^c/_d$ exactly when $ad - bc = \pm 1$. In the case that $ad - bc = +1$ a solution of $ax - by = 1$ is then $(x, y) = (d, c)$, and when $ad - bc = -1$ a solution of $ax - by = 1$ is $(x, y) = (-d, -c)$.

For a given coprime pair of positive integers a and b we can compute the continued fraction for ${}^a/_b$ and the corresponding strip of triangles in the Farey diagram from ${}^1/_0$ to ${}^a/_b$. The last edge in the zigzag path in this strip connects a fraction ${}^c/_d$ to ${}^a/_b$, so we have $ad - bc = \pm 1$. Since ${}^c/_d$ is the next to last vertex along the zigzag path, the continued fraction for ${}^c/_d$ is obtained from the continued fraction for ${}^a/_b$ by omitting the last term. From this truncated continued fraction we can then compute ${}^c/_d$ and hence a solution of $ax - by = 1$.

As an example, let us solve $67x - 24y = 1$. The continued fraction for ${}^{67}/_{24}$ is $2 + {}^1\!\nearrow_1 + {}^1\!\nearrow_3 + {}^1\!\nearrow_1 + {}^1\!\nearrow_4$. Omitting the last term gives $2 + {}^1\!\nearrow_1 + {}^1\!\nearrow_3 + {}^1\!\nearrow_1$ which equals ${}^{14}/_5$. Thus we have $67\cdot 5 - 24\cdot 14 = \pm 1$. The sign can be determined by observing that ${}^{67}/_{24}$ lies to the right of ${}^{14}/_5$ in the circular Farey diagram so ${}^{67}/_{24} < {}^{14}/_5$, hence $67\cdot 5 < 24\cdot 14$ and therefore $67\cdot 5 - 24\cdot 14 = -1$. Thus we obtain the solution $(x, y) = (-5, -14)$.



The fact that ${}^{67}/_{24}$ lies to the right of ${}^{14}/_5$ in the Farey diagram is a consequence of the strip of triangles having an even number of fans. With an odd number of fans the situation would be reversed. The number of fans is the number of terms in the continued fraction after the initial integer, so we see that it is not really necessary to draw the strip of triangles to figure out the correct sign.

Another way to determine the sign without using the diagram is by computing $67\cdot 5 - 24\cdot 14 \bmod 10$ to see whether we get $+1$ or $-1 \bmod 10$. Computing mod 10 means ignoring all but the last digit, so we get $7\cdot 5 - 4\cdot 4 = 19 \equiv -1 \bmod 10$ and hence the sign is negative.

We can get other solutions to $67x - 24y = 1$ by using other edges of the Farey diagram with endpoint ${}^{67}/_{24}$ instead of the edge from ${}^{14}/_5$. For example we could use the edge to ${}^{67}/_{24}$ in the lower border of the strip of triangles. By the mediant rule this edge joins ${}^{53}/_{19}$ to ${}^{67}/_{24}$, so we have $67\cdot 19 - 24\cdot 53 = \pm 1$ and this time the plus sign is correct, giving the solution $(x, y) = (19, 53)$. All the other edges connected to ${}^{67}/_{24}$ are obtained by repeatedly "adding" ${}^{67}/_{24}$ either to ${}^{14}/_5$ for edges above ${}^{67}/_{24}$, or to ${}^{53}/_{19}$ for edges below ${}^{67}/_{24}$. In the former case these are the edges leading to the fractions ${}^{14+67k}/_{5+24k}$ for positive integers k, and in the latter case they are the edges to ${}^{53+67k}/_{19+24k}$ for positive integers k. Notice that if we let k be negative in one of these formulas, we get the fractions given by the other formula. For example in ${}^{53+67k}\!/_{19+24k}$ the values $k = -1, -2, \cdots$ give the fractions ${}^{-14}\!/_{-5} = {}^{14}\!/_{5}$, ${}^{-81}\!/_{-29} = {}^{81}\!/_{29}, \cdots$ which are the values of ${}^{14+67k}\!/_{5+24k}$ for $k = 0, 1, \cdots$. This means that the general solution of $67x - 24y = 1$ is $(x,y) = (19+24k, 53+67k)$ for arbitrary integers k. Alternatively, we could write the general solution as $(x,y) = (-5-24k, -14-67k)$ or as $(x,y) = (-5+24k, -14+67k)$ since k can be replaced by $-k$.

This example illustrates a general fact:

Proposition 2.4. *For coprime integers a and b, if one solution of $ax - by = n$ is $(x,y) = (p,q)$ then the general solution is $(x,y) = (p+bk, q+ak)$ for k an arbitrary integer.*

Here we do not need to assume a and b are positive, so by changing the sign of b we can write the equation as $ax + by = n$ with general solution $(p - bk, q + ak)$, or alternatively as $(p + bk, q - ak)$.

Proof: One solution $(x,y) = (p,q)$ of $ax - by = n$ is given. For an arbitrary solution (x,y) we look at the difference $(x-p, y-q)$ which we denote as (x_0, y_0). This satisfies $ax_0 - by_0 = 0$, or in other words, $ax_0 = by_0$. Since a and b are coprime, the equation $ax_0 = by_0$ must have the form $a(bk) = b(ak)$ for some integer k, with $x_0 = bk$ and $y_0 = ak$. Hence every solution of $ax - by = n$ has the form $(x,y) = (p + x_0, q + y_0) = (p + bk, q + ak)$. It is easy to check that these formulas for x and y give solutions to $ax - by = n$ for all values of k. □

The Diophantine equation $ax - by = n$ can be interpreted as a congruence condition by rewriting it as $ax - n = by$ which implies that $ax \equiv n \bmod b$. Conversely, if $ax \equiv n \bmod b$ then this means that $ax - n = by$ for some integer y, so $ax - by = n$. Thus a solution (x,y) of $ax - by = n$ gives a solution x of $ax \equiv n \bmod b$, and a solution x of $ax \equiv n \bmod b$ gives a solution (x,y) of $ax - by = n$ since this equation allows y to be computed from a, b, n, and x if b is nonzero.

The special case $ax - by = 1$ is equivalent to $ax \equiv 1 \bmod b$ which says that x is a multiplicative inverse to a mod b. We know that $ax - by = 1$ has a solution exactly when a and b are coprime, so this means that a has a multiplicative inverse mod b exactly when a is coprime to b. For example the congruence classes mod 15 that are coprime to 15 are $1, 2, 4, 7, 8, 11, 13, 14$ and we can find multiplicative inverses for each of these by observing that the products $1 \cdot 1$, $2 \cdot 8$, $4 \cdot 4$, $7 \cdot 13$, $11 \cdot 11$, and $14 \cdot 14$ are each congruent to 1 mod 15. Thus the numbers $1, 4, 11$, and 14 are their own inverses mod 15 while the other inverses occur in pairs, the pair $2, 8$ and the pair $7, 13$. We could shorten these calculations by noting that if $ax \equiv 1 \bmod b$ then $(-a)(-x) \equiv 1 \bmod b$, so for example $2 \cdot 8 \equiv 1 \bmod 15$ implies $(-2)(-8) \equiv 1 \bmod 15$ or in other words $13 \cdot 7 \equiv 1 \bmod 15$. Similarly $4 \cdot 4 \equiv 1 \bmod 15$ implies $11 \cdot 11 \equiv 1 \bmod 15$.

The function which assigns to each positive integer n the number of congruence classes mod n of numbers coprime to n is called the ***Euler phi function*** $\varphi(n)$. Thus in the preceding example of multiplicative inverses mod 15 we have $\varphi(15) = 8$ from the eight numbers $1, 2, 4, 7, 8, 11, 13, 14$. Later in this section we will obtain a formula for $\varphi(n)$.

Linear Diophantine equations with more than two variables can be solved by reduction to the case of two variables. Consider for example a three-variable equation $ax + by + cz = n$. Any number that divides all three coefficients a, b, c must also divide n if a solution is to exist, and in this case we can simplify the equation by dividing it by the greatest common divisor of a, b, and c, so we may as well assume that the greatest common divisor of a, b, and c is 1.

As an example that is typical of the general case for three variables, consider the equation $6x + 10y + 15z = 7$. Here the greatest common divisor of 6, 10, and 15 is 1, although when taken two at a time they have larger common divisors: 2 for 6 and 10, 3 for 6 and 15, and 5 for 10 and 15.

The idea for solving $6x+10y+15z = 7$ is to write it first as $2(3x+5y)+15z = 7$ and then to rewrite this as the two equations $3x + 5y = w$ and $2w + 15z = 7$. The first equation $3x + 5y = w$ has solutions for every w since 3 and 5 are coprime, and we can find the solutions by first solving $3x + 5y = 1$ and then multiplying these solutions by w. Since the coefficients 3 and 5 are so small, we can find a solution of $3x + 5y = 1$ by inspection rather than computing continued fractions, and we see that $(x, y) = (2, -1)$ is a solution. Then $(x, y) = (2w, -w)$ is a solution of $3x + 5y = w$. Applying Proposition 2.4, the general solution of $3x + 5y = w$ can therefore be written as $(x, y) = (2w + 5s, -w - 3s)$ for s an arbitrary integer.

Next we solve $2w + 15z = 7$. A solution of $2w + 15z = 1$ is $(w, z) = (8, -1)$ so a solution of $2w + 15z = 7$ is $(w, z) = (56, -7)$. The general solution of $2w + 15z = 7$ is then $(w, z) = (56 + 15t, -7 - 2t)$ for arbitrary integers t. Alternatively, we could notice that $2w + 15z = 7$ has the simpler solution $(w, z) = (-4, 1)$, obtained either by inspection or by letting $t = -4$ in the pair $(56 + 15t, -7 - 2t)$. Hence the general solution of $2w + 15z = 7$ can also be written as $(w, z) = (-4 + 15t, 1 - 2t)$.

Using $(w, z) = (-4 + 15t, 1 - 2t)$ we now substitute $w = -4 + 15t$ into the earlier formula $(x, y) = (2w + 5s, -w - 3s)$ to obtain the final answer in terms of the arbitrary intgers s and t:

$$\begin{aligned}(x, y, z) &= (2(-4 + 15t) + 5s, -(-4 + 15t) - 3s, 1 - 2t)\\ &= (-8 + 5s + 30t, 4 - 3s - 15t, 1 - 2t)\end{aligned}$$

In the spirit of Proposition 2.4 we can say that a particular solution of $6x+10y+15z = 7$ is $(-8, 4, 1)$, obtained by setting $s = t = 0$, and the general solution is obtained by adding this particular solution to $(5s + 30t, -3s - 15t, -2t)$ which is the general solution of the associated equation $6x + 10y + 15z = 0$ with right side zero.

The situation for equations with more variables is similar to what happened in this example, with an equation in n variables breaking up into $n-1$ equations in two variables. Each of these has solutions depending on an integer parameter, so the solutions of the n-variable equation depend on $n-1$ independent parameters.

The Chinese Remainder Theorem

We can apply what we have learned about linear Diophantine equations to derive a general fact about congruences often referred to as the Chinese Remainder Theorem since it was used in ancient Chinese manuscripts to solve mathematical puzzles of a certain type. Here is the statement:

Proposition 2.5. *A collection of congruence conditions*

$$\begin{aligned} x &\equiv a_1 \mod m_1 \\ x &\equiv a_2 \mod m_2 \\ &\cdots \\ x &\equiv a_k \mod m_k \end{aligned}$$

always has a simultaneous solution provided that no two of the moduli m_i have a common divisor greater than 1, and in this case the collection of all solutions forms a single congruence class modulo the product $m_1 \cdots m_k$.

Without the hypothesis that the various moduli m_i are coprime there may not be a common solution. For example the two congruences $x \equiv 5 \bmod 6$ and $x \equiv 7 \bmod 15$ have no common solution since the first congruence implies $x \equiv 2 \bmod 3$ while the second congruence implies $x \equiv 1 \bmod 3$. Here we are using the following general fact about congruences that will be used often:

- If a congruence $a \equiv b$ holds mod n then it holds mod d for each divisor d of n.

This is true because if n divides $a-b$ then so does d for each divisor d of n.

Proof of Proposition 2.5: Let us first prove the existence of a common solution x when there are just two congruences $x \equiv a_1 \bmod m_1$ and $x \equiv a_2 \bmod m_2$. In this case the desired number x will have the form $x = a_1 + x_1 m_1 = a_2 + x_2 m_2$ for some pair of yet-to-be-determined numbers x_1 and x_2. We can rewrite the equation $a_1 + x_1 m_1 = a_2 + x_2 m_2$ as $m_1 x_1 - m_2 x_2 = a_2 - a_1$. We know that this equation has a solution (x_1, x_2) with integers x_1 and x_2 whenever m_1 and m_2 are coprime. This is obtained by first finding integers n_1 and n_2 such that $m_1 n_1 + m_2 n_2 = 1$ and then multiplying this equation by $a_2 - a_1$ to get $(a_2 - a_1) m_1 n_1 + (a_2 - a_1) m_2 n_2 = a_2 - a_1$. Then in the equation $m_1 x_1 - m_2 x_2 = a_2 - a_1$ we may choose $x_1 = (a_2 - a_1) n_1$ and

$x_2 = (a_2 - a_1)(-n_2)$. Thus we have:

$$\begin{aligned} x &= a_1 + x_1 m_1 \\ &= a_1 + m_1(a_2 - a_1)n_1 \\ &= a_1(1 - m_1 n_1) + a_2 m_1 n_1 \\ &= a_1 m_2 n_2 + a_2 m_1 n_1 \quad \text{since} \quad 1 - m_1 n_1 = m_2 n_2 \end{aligned}$$

Summarizing, we have the solution $x = a_1 m_2 n_2 + a_2 m_1 n_1$ where n_1 and n_2 satisfy $m_1 n_1 + m_2 n_2 = 1$.

For a system of more than two congruences we may suppose by induction on the number of congruences that we have a number $x = a$ satisfying all but the last congruence $x \equiv a_k \bmod m_k$. From the preceding paragraph we know that a number x exists satisfying the two congruences $x \equiv a \bmod m_1 \cdots m_{k-1}$ and $x \equiv a_k \bmod m_k$ since $m_1 \cdots m_{k-1}$ and m_k are coprime. This gives the desired solution to all k congruences $x \equiv a_i \bmod m_i$ since $x \equiv a \bmod m_1 \cdots m_{k-1}$ implies $x \equiv a \bmod m_i$ for each $i < k$, and $a \equiv a_i \bmod m_i$ for each $i < k$ by the inductive hypothesis.

Now we show that all the different solutions of the given set of congruences form a single congruence class mod $m_1 \cdots m_k$. If x and y are two solutions then the difference $x - y$ is congruent to 0 mod each of the numbers $m_1, \cdots, m_k$, which means that it is divisible by each m_i and hence by their product since they have no common factors. Thus $x \equiv y \bmod m_1 \cdots m_k$, which shows that all the solutions lie in a single congruence class mod $m_1 \cdots m_k$. Moreover every number in this congruence class is a solution since if x is one solution and $y \equiv x \bmod m_1 \cdots m_k$ then $y \equiv x \bmod m_i$ for each i, so $x \equiv a_i \bmod m_i$ implies $y \equiv a_i \bmod m_i$. □

As an illustration of the method in this proof let us find all numbers that are congruent to 7 mod 9 and to 8 mod 11. First we find a solution of $9n_1 + 11n_2 = 1$ by the earlier methods. One such solution is $(n_1, n_2) = (5, -4)$. The formula $x = a_1 m_2 n_2 + a_2 m_1 n_1$ then gives $x = -7 \cdot 11 \cdot 4 + 8 \cdot 9 \cdot 5 = -308 + 360 = 52$. We are free to change this by adding any multiple of $9 \cdot 11$, so the general solution is $52 + 99t$ for arbitrary integers t. If we were to modify the problem by adding a third congruence condition such as $x \equiv 4 \bmod 7$ then we would just be solving the two congruences $x \equiv 52 \bmod 99$ and $x \equiv 4 \bmod 7$ by the same method.

There is a geometric picture that gives a way of visualizing what the Chinese Remainder Theorem is saying. Consider the case of two simultaneous congruences $x \equiv a \bmod m$ and $x \equiv b \bmod n$ where m and n are coprime. We can then label the mn unit squares in an $m \times n$ rectangle by the numbers $1, 2, 3, \cdots$ starting in the lower left corner and continuing upward to the right at a 45 degree angle as shown in the following figure for the case of a 9×4 rectangle:

28	20	12	4	32	24	16	8	36
19	11	3	31	23	15	7	35	27
10	2	30	22	14	6	34	26	18
1	29	21	13	5	33	25	17	9

Whenever we run over the top edge, we jump back to the bottom in order to continue, and when we reach the right edge, we jump back to the left edge. This amounts to taking congruence classes mod m horizontally and mod n vertically. What the Chinese Remainder Theorem says is that when m and n are coprime, each unit square in the $m \times n$ rectangle is labeled exactly once by a number from 1 to mn. (Without the coprimeness some squares would have no labels while others would have multiple labels.) The figure thus illustrates that specifying a congruence class mod mn is equivalent to specifying a pair of congruence classes mod m and mod n via the projections onto the two axes.

For the case of three simultaneous congruences there is an analogous picture with a three-dimensional rectangular box partitioned into unit cubes. More generally, for k congruences one would be dealing with a k-dimensional box.

A common situation for applying the Chinese Remainder Theorem is to start with a number n factored as $n = p_1^{r_1} \cdots p_k^{r_k}$ for distinct primes $p_1, \cdots, p_k$, so that a congruence $x \equiv a$ mod n is equivalent to a set of k congruences $x \equiv a_i$ mod $p_i^{r_i}$. If we add the condition that each a_i is not divisible by the corresponding prime p_i then a simultaneous solution $x = a$ for all k congruences must be coprime to n since $a \equiv a_i$ mod $p_i^{r_i}$ implies $a \equiv a_i$ mod p_i and we assume a_i is nonzero mod p_i so a is also nonzero mod p_i. Conversely, if a is coprime to n and satisfies a set of congruences $a \equiv a_i$ mod $p_i^{r_i}$ and hence $a \equiv a_i$ mod p_i, then a_i must be nonzero mod p_i since a is. Thus congruence classes mod n of numbers a coprime to n are equivalent to congruence classes mod $p_i^{r_i}$ of numbers a_i coprime to p_i, one for each i.

In the geometric picture for the case $k = 2$ with a rectangular array of unit squares, if we require a_1 to be coprime to p_1 then we are omitting the numbers in certain vertical columns of squares, the columns whose horizontal coordinate is a multiple of p_1. Similarly, when we require a_2 to be coprime to p_2 we omit the numbers in the horizontal rows whose vertical coordinate is a multiple of p_2. The numbers in the boxes that are not omitted are then the numbers coprime to $n = p_1^{r_1} p_2^{r_2}$. Here is the picture for the case $n = 3^2 \cdot 2^2$:

28	20	12	4	32	24	16	8	36
19	11	3	31	23	15	7	35	27
10	2	30	22	14	6	34	26	18
1	29	21	13	5	33	25	17	9

Here the 12 unshaded squares are what is left after columns 3, 6, and 9 are excluded along with rows 2 and 4. In other words we delete multiples of 2 and 3, leaving the numbers $1, 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35$ as the numbers less than 36 that are coprime to 36.

In the corresponding three-dimensional picture for $k = 3$ we would be omitting the cubes in certain slices parallel to the three coordinate planes, and similarly for $k > 3$.

The Euler Phi Function

We can now obtain a formula for the Euler phi function $\varphi(n)$ which counts the number of congruence classes mod n of integers coprime to n. The arguments above show that $\varphi(n) = \varphi(p_1^{r_1}) \cdots \varphi(p_k^{r_k})$ when $n = p_1^{r_1} \cdots p_k^{r_k}$ for distinct primes p_i. For a prime p we have $\varphi(p^r) = p^r - p^{r-1} = p^{r-1}(p-1)$ since we are counting how many numbers remain from $1, 2, 3, \cdots, p^r$ after we delete $p, 2p, 3p, \cdots, (p^{r-1})p = p^r$. Thus we have a formula for $\varphi(n)$:

$$\begin{aligned}\varphi(n) &= p_1^{r_1-1}(p_1-1)p_2^{r_2-1}(p_2-1)\cdots p_k^{r_k-1}(p_k-1)\\ &= n\left(\frac{p_1-1}{p_1}\right)\left(\frac{p_2-1}{p_2}\right)\cdots\left(\frac{p_k-1}{p_k}\right)\end{aligned}$$

If we omit the factor n from this last product, the remaining product of the terms ${}^{(p_i-1)}/_{p_i}$ tells what proportion of the numbers less than n are coprime to n. Notice that this does not depend on the exponents r_i. For example $\varphi(36) = \varphi(4)\varphi(9) = 2 \cdot 6 = 12$, which is $1/2 \cdot 2/3 = 1/3$ times 36, in agreement with the preceding figure.

The way that $\varphi(n)$ varies with n is rather erratic since the prime factorizations of adjacent numbers are not related. For example we have $\varphi(1000) = \varphi(2^3 5^3) = 2^2(2-1)5^2(5-1) = 400$, in agreement with the fact that the numbers coprime to 2 and 5 are the numbers with last digit 1, 3, 7, or 9, which means four out of every ten numbers or 400 out of the first 1000 numbers. For the adjacent numbers 999 and 1001 we have $\varphi(999) = \varphi(3^3 \cdot 37) = 18 \cdot 36 = 648$ and $\varphi(1001) = \varphi(7 \cdot 11 \cdot 13) = 6 \cdot 10 \cdot 12 = 720$.

An Example with a Quadratic Diophantine Equation

The Chinese Remainder Theorem can be applied to give an example of a Diophantine equation that has a solution mod n for each positive integer n but does not have an actual integer solution. The example is the equation $2x^2 + 7y^2 = 1$. This obviously has no integer solutions, although it does have rational solutions such as $(x, y) = (1/3, 1/3)$ and $(3/5, 1/5)$. We can use either of these rational solutions to get a solution mod n for certain values of n in the following way. Let us take the solution $(3/5, 1/5)$ for example. This rational solution will give an integer solution mod n provided that 5 has a multiplicative inverse "$1/5$" mod n. For example for $n = 14$ a multiplicative inverse for 5 is 3 since $5 \cdot 3 \equiv 1$ mod 14. If we multiply the equation $2(3/5)^2 + 7(1/5)^2 = 1$ by 5^2 to get $2 \cdot 3^2 + 7 \cdot 1^2 = 5^2$ and then multiply by 3^2, the inverse of 5^2 mod 14, we get $2 \cdot 9^2 + 7 \cdot 3^2 \equiv 1$ mod 14.

This argument gives a solution of $2x^2 + 7y^2 \equiv 1$ mod n whenever 5 has a multiplicative inverse mod n. As we saw earlier in this section, this happens whenever 5 is coprime to n, which means that 5 does not divide n. Similarly, using the other rational solution $(1/3, 1/3)$ we can solve $2x^2 + 7y^2 = 1$ mod n whenever 3 does not divide n by finding a multiplicative inverse for 3 mod n.

There remains the possibility that n is divisible by both 3 and 5, and this is where the Chinese Remainder Theorem will be used. Consider for example the case $n = 30$. We can factor this as $5 \cdot 6$ where one factor is not divisible by 3 and the other is not divisible by 5. By the method above we can obtain a solution of $2x^2 + 7y^2 \equiv 1$ mod 5 from $(1/3, 1/3)$ using $3 \cdot 2 \equiv 1$ mod 5 so $(1/3, 1/3)$ becomes $(2, 2)$. For $2x^2 + 7y^2 \equiv 1$ mod 6 we use $(3/5, 1/5)$ and the fact that $5 \cdot 5 \equiv 1$ mod 6 so $(3/5, 1/5)$ becomes $(3 \cdot 5, 5) \equiv (3, 5)$ mod 6. Thus we want to find (x, y) with $(x, y) \equiv (2, 2)$ mod 5 and $(x, y) \equiv (3, 5)$ mod 6. This we do by two applications of the Chinese Remainder Theorem, once for x and once for y. We use the earlier formula $a_1m_2n_2 + a_2m_1n_1$ where $5n_1 + 6n_2 = 1$ so $n_1 = -1$ and $n_2 = 1$. This yields $x = 2 \cdot 6 \cdot 1 - 3 \cdot 5 \cdot 1 = -3$ and $y = 2 \cdot 6 \cdot 1 - 5 \cdot 5 \cdot 1 = -13$. Thus $2(-3)^2 + 7(-13)^2 \equiv 1$ mod 5 and mod 6. This implies the congruence also holds mod 30 since the difference $2(-3)^2 + 7(-13)^2 - 1$ is divisible by 5 and by 6, hence by 30 since 5 and 6 are coprime. This method for the case $n = 30$ works for any n divisible by 3 and 5 since any such n can be factored as $n = kl$ where k is not divisible by 3 and l is not divisible by 5.

One might ask how rational solutions of $2x^2 + 7y^2 = 1$ such as $(1/3, 1/3)$ and $(3/5, 1/5)$ can be found. Rational solutions of $2x^2 + 7y^2 = 1$ are equivalent to integer solutions of $2x^2 + 7y^2 = z^2$, so we are looking for integers x and y such that $2x^2+7y^2$ is a square. This is a special case of the general problem of solving quadratic Diophantine equations $ax^2 + bxy + cy^2 = n$ which will be a central theme of the book starting in Chapter 4.

A Digression on Rational Points on Quadratic Curves

A key point in the preceding example was the existence of rational solutions of $2x^2 + 7y^2 = 1$, which correspond to rational points on the curve $2x^2 + 7y^2 = 1$, so let us consider now the general problem of determining when a quadratic curve $ax^2 + bxy + cy^2 = d$ contains rational points. Here a, b, c, and d are rational numbers but there is no loss of generality in assuming they are integers since we can multiply the equation by a common denominator for a, b, c, and d if they are not all integers.

The first step is to reduce to the case that $b = 0$. If $a \neq 0$ we can write:

$$ax^2 + bxy + cy^2 = a\Big(x + \frac{b}{2a}y\Big)^2 + \Big(c - \frac{b^2}{4a}\Big)y^2$$

Then if we change variables to $X = x + {}^{b}\!/_{2a}\,y$ and $Y = y$ this converts the equation $ax^2 + bxy + cy^2 = d$ into the equation $aX^2 + c'Y^2 = d$ for $c' = c - {}^{b^2}\!/_{4a}$. Rational values of x and y give rational values for X and Y, and conversely rational values for X and Y give rational values for x and y since the change of variables is reversible, with $x = X - {}^{b}\!/_{2a}\,Y$ and $y = Y$. If $a = 0$ and $c \neq 0$ we can change variables as above but with a and c reversed. If both a and c are 0 the equation is $bxy = d$ which always has rational solutions if $b \neq 0$.

Thus it suffices to determine whether curves $ax^2 + by^2 = c$ have rational points. Again we can multiply through by a common denominator to make a, b, and c integers. We assume a, b, and c are nonzero to avoid trivial cases. To have solutions we obviously need to assume that a and b do not have one sign and c the opposite sign.

If rational numbers x and y satisfy $ax^2 + by^2 = c$ we can put them over a common denominator and write them as quotients X/Z and Y/Z for integers X, Y, Z, and then the equation becomes $aX^2 + bY^2 = cZ^2$ for which we are seeking integer solutions (X, Y, Z). With three variables instead of two it may appear that we have made the problem more complicated, but an advantage of the new equation is that it is *homogeneous* in the sense that all three terms have the same degree, namely 2. This means that if (X, Y, Z) is a solution, then so is (kX, kY, kZ) for any constant k. In particular, rational solutions can always be converted to integer solutions. The homogeneous equation has the trivial solution $(0, 0, 0)$ but this is not very interesting so we will always exclude this trivial solution. In fact we will need solutions with $Z \neq 0$ to get actual points $(x, y) = (X/Z, Y/Z)$ on the curve $ax^2 + by^2 = c$.

Thus we are asking when an equation $ax^2 + by^2 = cz^2$ has an integer or rational solution $(x, y, z) \neq (0, 0, 0)$. There are a few preliminary simplifications in the coefficients a, b, c that can be made. Suppose first that a factors as $a'd^2$ for some integers a' and $d > 1$. The equation can then be written as $a'(dx)^2 + by^2 = cz^2$, and finding rational solutions of $ax^2 + by^2 = cz^2$ is equivalent to finding rational solutions of $a'x^2 + by^2 = cz^2$. Square factors of b and c can be absorbed into y^2

and z^2 in the same way. Thus there is no loss of generality in assuming that each of the coefficients a, b, c in $ax^2 + by^2 = cz^2$ is ***squarefree***, that is, has no square factors greater than 1.

If all three coefficients a, b, c have a common prime factor p we can of course divide the equation by p to get a simpler equation. Repeating this step, we may assume no prime p divides all three coefficients. If p divides two of the coefficients, say $a = pa'$ and $b = pb'$, we can still simplify the equation by multiplying it by p to get $a'(px)^2 + b'(py)^2 = pcz^2$ which can be written as $a'x^2 + b'y^2 = pcz^2$ by absorbing p into x and y, and this is a simpler equation in that $|abc|$ has decreased by a factor of p. The new equation still has squarefree coefficients since we could assume that the divisor p of a and b was not also a divisor of c. By the same reasoning we can arrange also that a and c are coprime and b and c are coprime, with all three coefficients still squarefree.

Now we have Legendre's Theorem as described in Chapter 0:

Theorem 2.6. *An equation $ax^2 + by^2 = cz^2$ with a, b, and c squarefree coprime nonzero integers has an integer solution $(x, y, z) \neq (0, 0, 0)$ exactly when the following conditions are satisfied: ac is a square mod b, bc is a square mod a, $-ab$ is a square mod c, and a and b do not both have the opposite sign from c.*

A more symmetric statement could be obtained by changing the sign of c and writing the equation as $ax^2 + by^2 + cz^2 = 0$. Then the conditions would be that $-ac$ is a square mod b, $-bc$ is a square mod a, $-ab$ is a square mod c, and the three coefficients a, b, c do not all have the same sign.

Proof: First we show that these congruence conditions must be satisfied if a solution exists. Suppose that we have a solution $(x, y, z) \neq (0, 0, 0)$ of $ax^2 + by^2 = cz^2$. We can assume each pair of x, y, z is coprime since for example if a prime p divides x and y then p^2 divides $ax^2 + by^2$ hence it divides cz^2, which implies p divides z since c is squarefree. Then the solution (x, y, z) could be simplified by dividing by p.

The equation $ax^2 + by^2 = cz^2$ implies that $ax^2 \equiv cz^2 \bmod b$. After multiplying this congruence by c we get $acx^2 \equiv c^2z^2 \bmod b$. Now, x and b are coprime since any prime dividing both would divide $ax^2 + by^2 = cz^2$ and so would divide c or z, neither of which is possible since b and c are coprime and x and z are coprime. Since x is coprime to b it has a multiplicative inverse mod b. Multiplying the congruence $acx^2 \equiv c^2z^2 \bmod b$ by the square of this inverse, we conclude that ac is a square mod b. In the same way we see that bc is a square mod a and $-ab$ is a square mod c.

The converse is considerably harder to prove, so let us first outline what the strategy will be. We will use the more symmetric equation $ax^2 + by^2 + cz^2 = 0$. If the left side of this equation could be factored as

$$ax^2 + by^2 + cz^2 = (a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z)$$

with all coefficients integers, then finding a solution of $ax^2 + by^2 + cz^2 = 0$ would be rather easy since we would just have to solve the linear Diophantine equation obtained by setting either factor equal to 0. However, factorizations like this rarely exist. Instead we will show that the congruence conditions in the theorem guarantee that there is a factorization modulo a suitable number n, namely $n = abc$. What this means concretely is that if one multiplies out the product of the two linear factors on the right in the displayed equation above, then the coefficients of the x^2, y^2, and z^2 terms will be congruent to a, b, and c mod n and the coefficients of the xy, yz, and xz terms will be 0 mod n. A solution of either congruence $a_ix + b_iy + c_iz \equiv 0 \bmod abc$, say $a_1x + b_1y + c_1z \equiv 0 \bmod abc$, will then give a solution of the congruence $ax^2 + by^2 + cz^2 \equiv 0 \bmod abc$.

The next step in the proof will be to show that a solution (x, y, z) of the congruence $a_1x + b_1y + c_1z \equiv 0 \bmod abc$ can be chosen so that the value of $ax^2 + by^2 + cz^2$ is a fairly small multiple of abc, in fact either 0 or $\pm abc$. The last step in the proof will then be a rather subtle trick to convert a solution of $ax^2 + by^2 + cz^2 = \pm abc$ into a solution of $ax^2 + by^2 + cz^2 = 0$.

Now we begin to fill in details. To factor $ax^2 + by^2 + cz^2$ mod abc we first factor it mod a, b, and c separately. To factor it mod a we just need to factor $by^2 + cz^2$ mod a. Multiplying $by^2 + cz^2$ by b gives $b^2y^2 + bcz^2$. We are assuming that $-bc$ is a square mod a so we have $-bc \equiv r^2$ mod a for some integer r. Then $b^2y^2 + bcz^2 \equiv b^2y^2 - r^2z^2$ mod a with $b^2y^2 - r^2z^2$ factoring as $(by + rz)(by - rz)$. Since b is coprime to a it has an inverse b^{-1} mod a so after multiplying the congruence $b^2y^2 + bcz^2 \equiv (by + rz)(by - rz)$ mod a by b^{-1} we have the desired factorization $by^2 + cz^2 \equiv (y + b^{-1}rz)(by - rz)$ mod a. Thus there is a factorization mod a of $ax^2 + by^2 + cz^2$ as a product $(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z)$ where the coefficients a_1 and a_2 happen to be 0, but this will not be significant for the rest of the argument.

In the same way there are similar factorizations of $ax^2 + by^2 + cz^2$ mod b and mod c, with possibly different coefficients $a_1, b_1, c_1, a_2, b_2, c_2$ of the linear factors. The Chinese Remainder Theorem, applied once for each of the six coefficients, implies that there is a single choice for the coefficients that works mod a, b, and c simultaneously. Since a, b, and c are coprime, the factorization then holds mod abc.

We will be interested in triples (x, y, z) of integers satisfying three inequalities

$$0 \le x < \alpha \qquad 0 \le y < \beta \qquad 0 \le z < \gamma \qquad (*)$$

for positive real numbers α, β, and γ that are not necessarily integers. To count how many triples (x, y, z) satisfy $(*)$ let $\lambda(\alpha)$ be the number of integers x with

$0 \le x < \alpha$, so $\lambda(\alpha) = \alpha$ if α is an integer and $\lambda(\alpha) = 1 + \lfloor\alpha\rfloor$ if α is not an integer, where $\lfloor\alpha\rfloor$ is the largest integer less than or equal to α. Thus $\lambda(\alpha) > \alpha$ if α is not an integer. The number of triples (x, y, z) satisfying $(*)$ is then $\lambda(\alpha)\lambda(\beta)\lambda(\gamma)$.

If $\lambda(\alpha)\lambda(\beta)\lambda(\gamma) > |abc|$ there must exist two different triples (x', y', z') and (x'', y'', z'') satisfying $(*)$ such that $a_1x' + b_1y' + c_1z' \equiv a_1x'' + b_1y'' + c_1z''$ mod abc. The triple $(x, y, z) = (x' - x'', y' - y'', z' - z'') \neq (0, 0, 0)$ will then satisfy $a_1x + b_1y + c_1z \equiv 0$ mod abc. The triple $(|x|, |y|, |z|)$ will also satisfy $(*)$ so $x^2 < \alpha^2$, $y^2 < \beta^2$, and $z^2 < \gamma^2$.

For the triple (x, y, z) we have $ax^2 + by^2 + cz^2 \equiv 0$ mod abc from the factorization of $ax^2 + by^2 + cz^2$ mod abc. Since a, b, and c do not all have the same sign, we can assume two are positive and one is negative by multiplying the equation by -1 if necessary. After a possible permutation of the coefficients we can assume that $a > 0$, $b > 0$, and $c < 0$. Since $x^2 < \alpha^2$, $y^2 < \beta^2$, and $z^2 < \gamma^2$ we then have:

$$c\gamma^2 < cz^2 \le ax^2 + by^2 + cz^2 \le ax^2 + by^2 < a\alpha^2 + b\beta^2$$

If we choose $\alpha = \sqrt{|bc|}$, $\beta = \sqrt{|ac|}$, and $\gamma = \sqrt{|ab|}$ then these inequalities give the inequalities $-|abc| < ax^2 + by^2 + cz^2 < 2|abc|$. Since $ax^2 + by^2 + cz^2 \equiv 0$ mod $|abc|$ we must therefore have either $ax^2 + by^2 + cz^2 = 0$ or $ax^2 + by^2 + cz^2 = |abc|$. The chosen values for α, β, and γ also give $\alpha\beta\gamma = |abc|$ so the earlier hypothesis $\lambda(\alpha)\lambda(\beta)\lambda(\gamma) > |abc|$ becomes $\lambda(\alpha)\lambda(\beta)\lambda(\gamma) > \alpha\beta\gamma$ which is satisfied unless α, β, and γ are all integers. Since a, b, and c are coprime and squarefree, α, β, and γ are all integers only when a, b, and c are ± 1, but in this case the equation $ax^2 + by^2 + cz^2 = 0$ is just $x^2 + y^2 - z^2 = 0$ which has obvious integer solutions.

All that remains is to deal with the possibility $ax^2 + by^2 + cz^2 = |abc|$, so $ax^2 + by^2 + cz^2 = -abc$. Rewriting this equation as $ax^2 + by^2 + c(z^2 + ab) = 0$, we would like to convert it into an equation of the form $aX^2 + bY^2 + cZ^2 = 0$. This suggests that we multiply the equation by $z^2 + ab$ to get a term $cZ^2 = c(z^2 + ab)^2$. Multiplying $ax^2 + by^2$ by $z^2 + ab$, we have:

$$\begin{aligned}(ax^2 + by^2)(z^2 + ab) &= ax^2z^2 + a^2bx^2 + by^2z^2 + ab^2y^2 \\ &= a(xz + by)^2 + b(yz - ax)^2\end{aligned}$$

Thus we have a solution of $aX^2 + bY^2 + cZ^2 = 0$, and this is not the trivial solution $(0, 0, 0)$ since $Z = z^2 + ab > 0$. □

To apply Legendre's Theorem one needs to be able to determine which numbers are squares modulo a given number n. The brute force approach is just to compute all the possible squares. For example for $n = 15$ the numbers mod 15 are $0, \pm1, \pm2, \pm3, \pm4, \pm5, \pm6$, and ±7 so the squares mod 15 are obtained by squaring these to get $0, 1, 4, 9, 16 \equiv 1, 25 \equiv 10, 36 \equiv 6$, and $49 \equiv 4$. Thus only six of the fifteen congruence classes mod 15 are squares mod 15, namely $0, 1, 4, 6, 9$, and 10. This approach becomes tedious for large values of n, but in Section 6.2 we will develop more efficient methods for determining whether a number m is a square mod n, which turns out to be quite a subtle question.

Exercises

1. (a) Find all integer solutions of the equations $40x + 89y = 1$ and $40x + 89y = 5$. (b) Find another equation $ax + by = 1$ with integer coefficients a and b that has an integer solution in common with $40x + 89y = 1$. *Hint*: Use the Farey diagram.

2. Find all integers x satisfying the congruence $31x \equiv 1 \bmod 71$, and then do the same for the congruence $31x \equiv 10 \bmod 71$. Are the solutions unique mod 71, i.e., unique up to adding multiples of 71?

3. Find all integer solutions of the equation $9x + 12y + 20z = 4$, and do this more generally for $9x + 12y + 20z = n$.

4. Find all solutions of the simultaneous congruences $x \equiv 6 \bmod 13$ and $x \equiv 7 \bmod 18$.

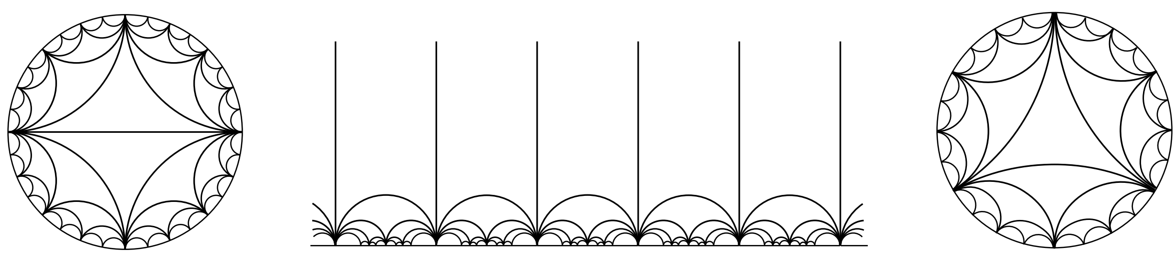
5. Show that for the Euler phi function the values $\varphi(n)$ approach infinity as n approaches infinity. In other words, show that for each number $N > 0$ there are only finitely many numbers n with $\varphi(n) < N$.

6. For each $n \leq 10$ determine which numbers are squares mod n by direct calculation.

7. Determine which curves $ax^2 + by^2 = c$ contain rational points for each triple of coprime integers a, b, c chosen from the numbers $1, 2, 3, 5$. When rational points exist, find a specific one.

3 Symmetries of the Farey Diagram

A notable feature of the various versions of the Farey diagram is their symmetry. For the circular Farey diagram the symmetries are the reflections across the horizontal and vertical axes and the 180 degree rotation about the center. For the upper half-plane Farey diagram there are symmetries that translate the diagram by any integer distance to the left or the right, as well as reflections across certain vertical lines, the vertical lines through an integer or half-integer point on the x-axis. The Farey diagram could also be drawn to have 120 degree rotational symmetry and three reflectional symmetries.



Our purpose in this chapter is to study all possible symmetries of the Farey diagram, where we interpret the word "symmetry" in a broader sense than the familiar meaning from Euclidean geometry. For our purposes, symmetries will be invertible transformations that take vertices to vertices, edges to edges, and triangles to triangles. There are simple algebraic formulas for these more general symmetries, and these formulas lead to effective means of calculation. An application in this chapter will be to computing the values of periodic or eventually periodic continued fractions, and symmetries of the diagram will play key roles in later chapters as well.

3.1 Linear Fractional Transformations

Our first goal will be to find formulas for all the symmetry transformations of the Farey diagram. The formulas will specify where each vertex is sent so they will have the form $T(x/y) = x'/y'$. It is easy to write down such formulas for some of the simpler symmetries. Reflection of the circular Farey diagram across the vertical axis sends a fraction x/y to y/x so it is the transformation $T(x/y) = y/x$. Reflection across the horizontal axis is $T(x/y) = -x/y$. Composing these two transformations in either order gives a 180 degree rotation of the Farey diagram about its centerpoint, the transformation $T(x/y) = {-y}/x$. For the upper halfplane Farey diagram the horizontal translation to the right by n units is $T(x/y) = x/y + n = {x+ny}/y$, while a leftward translation is $T(x/y) = x/y - n = {x-ny}/y$. All these formulas work equally well for the fraction $x/y = {\pm 1}/0$ with the exception of $x/y \pm n$, where the alternative forms ${x+ny}/y$ and ${x-ny}/y$ are preferable and give $T({\pm 1}/0) = {\pm 1}/0$.

In these examples the transformations have the form $T(x/y) = {ax+by}/{cx+dy}$ for integers a, b, c, d. Another notation is to let $z = x/y$ and then we have:

$$T(z) = T\left(\frac{x}{y}\right) = \frac{ax+by}{cx+dy} = \frac{a(x/y)+b}{c(x/y)+d} = \frac{az+b}{cz+d}$$

A transformation of the type $T(x/y) = {ax+by}/{cx+dy}$ or $T(z) = {az+b}/{cz+d}$ is called a ***linear fractional transformation*** since it is defined by a fraction whose numerator and denominator are linear functions. Fractions x/y, including ${\pm 1}/0$, correspond to pairs (x, y) and from this point of view linear fractional transformations $T(x/y) = {ax+by}/{cx+dy}$ correspond to linear transformations $T(x, y) = (ax+by, cx+dy)$. In matrix notation this becomes $T\binom{x}{y} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\binom{x}{y} = \binom{ax+by}{cx+dy}$.

Linear fractional transformations $T(x/y) = {ax+by}/{cx+dy}$ that give symmetries of the Farey diagram must take vertices to vertices and edges to edges, so let us see what this means for the coefficients a, b, c, d, which we will always assume are integers. Vertices of the Farey diagram are fractions p/q in lowest terms, including ${\pm 1}/0$, with p/q determining the same vertex as ${-p}/{-q}$. This ambiguity causes no problem for linear fractional transformations $T(x/y) = {ax+by}/{cx+dy}$ since ${ax+by}/{cx+dy} = {-ax-by}/{-cx-dy}$ so $T(p/q) = T({-p}/{-q})$. For T to take vertices to vertices means that for a fraction p/q in lowest terms we would like $T(p/q) = {ap+bq}/{cp+dq}$ to be in lowest terms as well. For T to take edges to edges means that if $\langle p/q, r/s\rangle$ is an edge we want $\langle {ap+bq}/{cp+dq}, {ar+bs}/{cr+ds}\rangle$ to be an edge also. In matrix terms this last condition is saying that if $\begin{pmatrix} p & r \\ q & s \end{pmatrix}$ has determinant ± 1 then $\begin{pmatrix} ap+bq & ar+bs \\ cp+dq & cr+ds \end{pmatrix}$, which is the product $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} p & r \\ q & s \end{pmatrix}$, should also have determinant ± 1. It is a general fact that the determinant of the product of two matrices is the product of the determinants of the two matrices. (For 2×2 matrices this is easy to check by a direct calculation.) Thus for $\begin{pmatrix} ap+bq & ar+bs \\ cp+dq & cr+ds \end{pmatrix}$ to have determinant ± 1 when $\begin{pmatrix} p & r \\ q & s \end{pmatrix}$ has determinant ± 1 the exact condition we need is that $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ should have determinant ± 1.

Proposition 3.1. *If the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer entries has determinant ± 1 then the associated linear fractional transformation $T(x/y) = {ax+by}/{cx+dy}$ takes vertices in the Farey diagram to vertices in the diagram and it takes each pair of vertices that are joined by an edge to another pair of vertices that are joined by an edge.*

It follows that T must take triangles in the diagram to triangles in the diagram since triangles correspond to sets of three vertices, any two of which are the endpoints of an edge.

Proof: We have shown that if $\langle p/q, r/s\rangle$ is an edge of the Farey diagram then so is $\langle T(p/q), T(r/s)\rangle$ when the matrix of T has determinant ± 1. This implies that T takes vertices to vertices since each vertex p/q is an endpoint of some edge $\langle p/q, r/s\rangle$, so $T(p/q)$ is an endpoint of the edge $\langle T(p/q), T(r/s)\rangle$ and therefore the fraction $T(p/q) = ap+bq/cp+dq$ is in lowest terms. □

We will use the notation $LF(\mathbb{Z})$ for the set of all linear fractional transformations $T(x/y) = ax+by/cx+dy$ with coefficients a, b, c, d in $\mathbb{Z}$ such that the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ has determinant ± 1. (Here $\mathbb{Z}$ is the set of integers.)

Changing a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ to its negative $-\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} -a & -b \\ -c & -d \end{smallmatrix}\right)$ produces the same linear fractional transformation since $-ax-by/-cx-dy = ax+by/cx+dy$. This is in fact the only way that different matrices with integer entries and determinant ± 1 can give the same linear fractional transformation, by the following argument. The transformation $T(x/y) = ax+by/cx+dy$ takes $1/0$ to a/c and $0/1$ to b/d so T determines each column of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ up to a sign. Changing the sign of both columns gives the same transformation so the only question is whether changing the sign of one column could give the same transformation. Changing the sign of the first column has the same effect as changing the sign of the second column since changing the sign of both columns gives the same transformation. So suppose that we change the sign of the second column, changing $ax+by/cx+dy$ to $ax-by/cx-dy$. If we apply these two transformations to $1/1$ we get $a+b/c+d$ and $a-b/c-d$. These fractions are in lowest terms by the previous proposition, so if they give the same vertex of the Farey diagram we would have either $a+b = a-b$ and $c+d = c-d$, hence $b = 0$ and $d = 0$, or we would have $a+b = b-a$ and $c+d = d-c$, hence $a = 0$ and $c = 0$. In either case the condition $ad - bc = \pm 1$ is violated. Thus we see that changing the sign of only one column of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ gives a different transformation, finishing the argument.

If we are given two linear fractional transformations $T(x/y) = ax+by/cx+dy$ and $S(x/y) = ex+fy/gx+hy$ then we can compose them to get another linear fractional transformation:

$$T\big(S(x/y)\big) = \frac{a(ex+fy)+b(gx+hy)}{c(ex+fy)+d(gx+hy)} = \frac{(ae+bg)x+(af+bh)y}{(ce+dg)x+(cf+dh)y}$$

The matrix of this transformation is just the product $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} e & f \\ g & h \end{smallmatrix}\right) = \left(\begin{smallmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{smallmatrix}\right)$, so composition of linear fractional transformations corresponds to matrix multiplication. It follows that if T and S are in $LF(\mathbb{Z})$ then so is their composition TS, which is also referred to as their product.

A transformation T in $LF(\mathbb{Z})$ has an inverse T^{-1} in $LF(\mathbb{Z})$ because the inverse of a 2×2 matrix is given by the formula

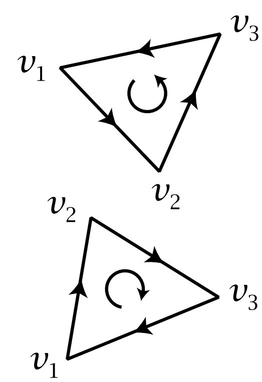
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

so if a, b, c, d are integers with $ad - bc = \pm 1$ then the inverse matrix also has integer entries and determinant ± 1. When computing the inverse of a transformation in $LF(\mathbb{Z})$ the factor $1/_{ad-bc}$ can be ignored since it is ± 1 and replacing a matrix by its negative gives the same linear fractional transformation, as we observed above.

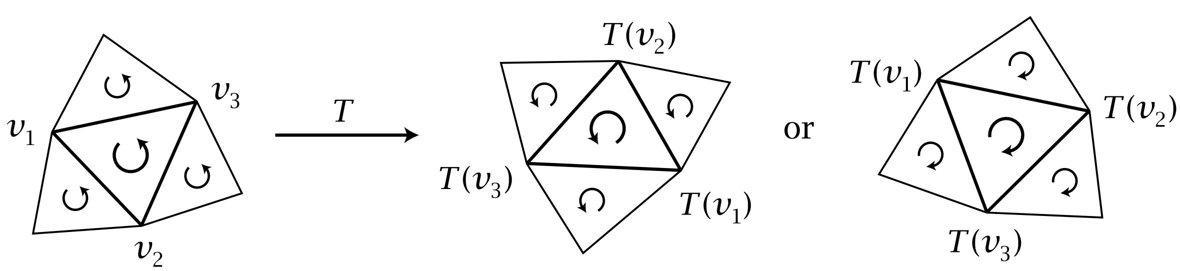
For a matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ the key property of its inverse A^{-1} is that both products AA^{-1} and $A^{-1}A$ are equal to the identity matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, corresponding to the identity transformation $I(x/y) = x/y$. Thus for any transformation T in $LF(\mathbb{Z})$ we have $TT^{-1} = I$ and $T^{-1}T = I$. The formula $T^{-1}T = I$ implies that T gives a one-to-one transformation of vertices since if two vertices v_1 and v_2 have the same image $T(v_1) = T(v_2)$ then we must have $T^{-1}(T(v_1)) = T^{-1}(T(v_2))$ so $v_1 = v_2$ and hence T cannot send two different vertices to the same vertex, which means it is one-to-one as a transformation from vertices to vertices. Also, the formula $TT^{-1} = I$ implies that every vertex v_1 is the image $T(v_2)$ of some vertex v_2 since we can write $v_1 = T(T^{-1}(v_1))$ and let $v_2 = T^{-1}(v_1)$. The same reasoning applies not just for vertices but also for edges and triangles. Thus T can never send two edges to the same edge or two triangles to the same triangle, and every edge or triangle is the image of some edge or triangle.

Orientations

Transformations in $LF(\mathbb{Z})$ can be divided into two types according to whether they preserve or reverse the orientations of triangles. A triangle in the Farey diagram can be oriented either clockwise or counterclockwise by choosing either a clockwise or counterclockwise ordering of its three vertices. Thus if the vertices are v_1, v_2, v_3 then this ordering of the vertices determines one orientation as in the figures at the right. This is the same orientation as when the vertices are ordered v_2, v_3, v_1 or v_3, v_1, v_2. The other three orderings determine the opposite orientation.



A transformation T in $LF(\mathbb{Z})$ takes each triangle to another triangle in a way that either preserves the two possible orientations or reverses them:



If a transformation preserves the orientation of one triangle, it has to preserve the orientation of the three adjacent triangles, and then of the triangles adjacent to these, and so on for all the triangles. Similarly, if the orientation of one triangle is reversed, then the orientations of all triangles are reversed. For example, reflection of the circular Farey diagram across its horizontal or vertical axis reverses the orientation of all triangles, while a 180 degree rotation of the diagram preserves the orientation of all triangles. A translation of the upper halfplane diagram by any number of units left or right preserves orientations of triangles while a reflection across a vertical line through an integer or half-integer point on the x-axis reverses orientations.

As we have seen, the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ corresponding to a linear fractional transformation ${}^{ax+by}/_{cx+dy}$ is unique up to multiplication by -1. The determinant $ad-bc$ does not change when each of a, b, c, d is changed to its negative, so each transformation in $LF(\mathbb{Z})$ has a well-defined determinant, either $+1$ or -1. The sign has a geometric interpretation:

Proposition 3.2. *An orientation-preserving transformation in $LF(\mathbb{Z})$ has determinant $+1$ and an orientation-reversing transformation has determinant -1.*

Proof: Consider a transformation $T({}^{x}/_{y}) = {}^{ax+by}/_{cx+dy}$ in $LF(\mathbb{Z})$ associated to a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. If we multiply one column of the matrix by -1 this changes the sign of the determinant. Let us check that it also changes whether T preserves or reverses orientation. Changing the sign of one column changes where T takes the triangle $\langle {}^{1}/_{0}, {}^{0}/_{1}, {}^{1}/_{1} \rangle$ from $\langle {}^{a}/_{c}, {}^{b}/_{d}, {}^{a+b}/_{c+d} \rangle$ to $\langle {}^{a}/_{c}, {}^{b}/_{d}, {}^{a-b}/_{c-d} \rangle$. These two triangles are different as we saw earlier, so they lie on opposite sides of the edge $\langle {}^{a}/_{c}, {}^{b}/_{d} \rangle$ and hence have opposite orientations. Thus the validity of the proposition for the transformation T is unaffected by changing one column of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ to its negative.

Applying this fact, we can arrange that $c \ge 0$ and $d \ge 0$ by multiplying columns by -1 if necessary. If $c = 0$ the condition $ad - bc = \pm 1$ implies $a = \pm 1$ and $d = 1$ (since $d \ge 0$), and then by multiplying the first column by -1 if necessary we can arrange that $a = 1$ so the matrix is $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$. This matrix has determinant $+1$ and the associated transformation sends the triangle $\langle {}^{1}/_{0}, {}^{0}/_{1}, {}^{1}/_{1} \rangle$ to $\langle {}^{1}/_{0}, {}^{b}/_{1}, {}^{b+1}/_{1} \rangle$ so it preserves orientation. Similarly, if $d = 0$ we can assume the matrix is $\left(\begin{smallmatrix} a & 1 \\ 1 & 0 \end{smallmatrix}\right)$ with determinant -1 and the associated transformation takes $\langle {}^{1}/_{0}, {}^{0}/_{1}, {}^{1}/_{1} \rangle$ to $\langle {}^{a}/_{1}, {}^{1}/_{0}, {}^{a+1}/_{1} \rangle$ so it reverses orientation.

Thus we have reduced to the case that $c > 0$ and $d > 0$. The transformation T takes the triangle $\langle {}^{1}/_{0}, {}^{0}/_{1}, {}^{1}/_{1} \rangle$ to $\langle {}^{a}/_{c}, {}^{b}/_{d}, {}^{a+b}/_{c+d} \rangle$ whose third vertex is the mediant of the first two. The edge $\langle {}^{a}/_{c}, {}^{b}/_{d} \rangle$ lies in either the upper or lower half of the circular Farey diagram, and in either case the orientation of $\langle {}^{a}/_{c}, {}^{b}/_{d}, {}^{a+b}/_{c+d} \rangle$ given by the ordering of its vertices is the same as the orientation of $\langle {}^{1}/_{0}, {}^{0}/_{1}, {}^{1}/_{1} \rangle$ exactly when ${}^{a}/_{c} > {}^{b}/_{d}$. Since $c > 0$ and $d > 0$ the inequality ${}^{a}/_{c} > {}^{b}/_{d}$ is equivalent to $ad - bc > 0$. Thus T is orientation-preserving exactly when $ad - bc = +1$. □

In what follows, when we say that a transformation T in $LF(\mathbb{Z})$ takes a triangle $\langle {}^{p}/_{q}, {}^{r}/_{s}, {}^{t}/_{u} \rangle$ to a triangle $\langle {}^{p'}/_{q'}, {}^{r'}/_{s'}, {}^{t'}/_{u'} \rangle$ we will mean that $T({}^{p}/_{q}) = {}^{p'}/_{q'}$, $T({}^{r}/_{s}) = {}^{r'}/_{s'}$, and $T({}^{t}/_{u}) = {}^{t'}/_{u'}$ so T preserves the order of the vertices. Similarly, when we say that T takes an edge $\langle p/q, r/s\rangle$ to an edge $\langle p'/q', r'/s'\rangle$ we will mean that $T(p/q) = p'/q'$ and $T(r/s) = r'/s'$.

Proposition 3.3. (a) *For any two triangles $\langle p/q, r/s, t/u\rangle$ and $\langle p'/q', r'/s', t'/u'\rangle$ in the Farey diagram there is a unique transformation in $LF(\mathbb{Z})$ taking $\langle p/q, r/s, t/u\rangle$ to $\langle p'/q', r'/s', t'/u'\rangle$.*
(b) *For any two edges $\langle p/q, r/s\rangle$ and $\langle p'/q', r'/s'\rangle$ there is a unique orientation-preserving transformation in $LF(\mathbb{Z})$ taking $\langle p/q, r/s\rangle$ to $\langle p'/q', r'/s'\rangle$.*

Proof: For a given pair of edges $\langle p/q, r/s\rangle$ and $\langle p'/q', r'/s'\rangle$ let T_1 be the transformation with matrix $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)$ and let T_2 be the transformation with matrix $\left(\begin{smallmatrix} p' & r' \\ q' & s' \end{smallmatrix}\right)$, so T_1 takes $\langle 1/0, 0/1\rangle$ to $\langle p/q, r/s\rangle$ and T_2 takes $\langle 1/0, 0/1\rangle$ to $\langle p'/q', r'/s'\rangle$. The composition $T = T_2T_1^{-1}$ then takes $\langle p/q, r/s\rangle$ to $\langle p'/q', r'/s'\rangle$. Hence T takes $\langle p/q, r/s, t/u\rangle$ to either $\langle p'/q', r'/s', t'/u'\rangle$ or the other triangle $\langle p'/q', r'/s', t''/u''\rangle$ having $\langle p'/q', r'/s'\rangle$ as an edge. For one of these two possibilities T is orientation-preserving and for the other T is orientation-reversing. We can change whether T preserves or reverses orientation by changing the signs in one column of the matrix of T_1 or T_2. Thus we can arrange that T takes $\langle p/q, r/s, t/u\rangle$ to $\langle p'/q', r'/s', t'/u'\rangle$.

A transformation in $LF(\mathbb{Z})$ taking $\langle p/q, r/s, t/u\rangle$ to $\langle p'/q', r'/s', t'/u'\rangle$ is unique since it must take the three triangles sharing an edge with $\langle p/q, r/s, t/u\rangle$ to the three triangles sharing the corresponding edges with $\langle p'/q', r'/s', t'/u'\rangle$, and then this determines where the next layer of six triangles sharing an edge with the three triangles adjacent to $\langle p/q, r/s, t/u\rangle$ are sent, and so on until all triangles are accounted for.

For part (b) we have found a product $T_2T_1^{-1}$ taking $\langle p/q, r/s\rangle$ to $\langle p'/q', r'/s'\rangle$, and if this product is orientation-reversing, we can make it orientation-preserving by changing the sign of one column of the matrix of T_1 or T_2. An orientation-preserving transformation taking $\langle p/q, r/s\rangle$ to $\langle p'/q', r'/s'\rangle$ is unique since if it preserves orientations this determines where it sends the two triangles adjacent to $\langle p/q, r/s\rangle$ and then uniqueness follows from the uniqueness in part (a). □

Reflections, Rotations, and Pivoting Transformations

In the remainder of this section we will describe five fairly simple types of symmetries of the Farey diagram given by elements of $LF(\mathbb{Z})$. Two other slightly more complicated types of symmetries will be described in the next section where they arise in connection with continued fractions.

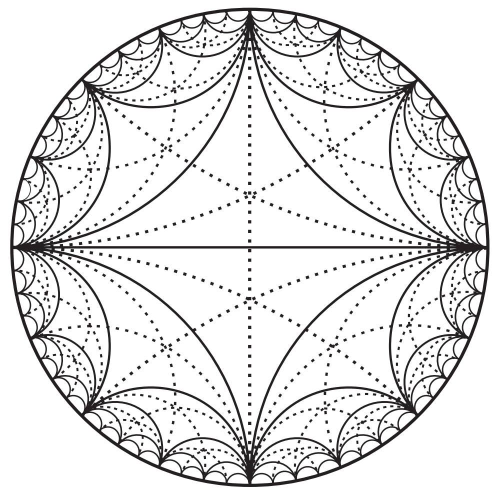
(1) The diagram can be reflected across any of its edges, leaving this edge fixed and interchanging the two triangles adjacent to it. This then determines where all the other triangles are sent. The simplest case is reflection across $\langle 1/0, 0/1\rangle$, the transformation $T(x/y) = -x/y$. To obtain a reflection across an arbitrary edge $\langle a/b, c/d\rangle$, let S be the transformation with matrix $\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right)$. The composition STS^{-1} sends $\langle a/b, c/d\rangle$ first to $\langle 1/0, 0/1\rangle$ by S^{-1}, then T leaves this edge fixed, then S sends it back to $\langle a/b, c/d\rangle$.

Thus STS^{-1} leaves $\langle a/b, c/d\rangle$ fixed so STS^{-1} is either the identity transformation or reflection across $\langle a/b, c/d\rangle$. The transformations S and S^{-1} either both preserve orientation or both reverse orientation, while T reverses orientation, so STS^{-1} reverses orientation and is therefore reflection across the edge $\langle a/b, c/d\rangle$. Its matrix can easily be computed:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} -a & c \\ -b & d \end{pmatrix}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} -ad-bc & 2ac \\ -2bd & ad+bc \end{pmatrix}$$

For example, the matrix giving reflection across $\langle 1/1, 1/2\rangle$ is $\begin{pmatrix} -3 & 2 \\ -4 & 3 \end{pmatrix}$. This can be checked by noting that its determinant is -1 and it fixes $1/1$ and $1/2$.

(2) The diagram can also be reflected across an arc perpendicular to any of its edges, any of the dotted arcs in the figure at the right. Each of the two triangles this arc crosses is then sent to itself by a reflection that interchanges the two vertices at the ends of the given edge and fixes the two vertices at the endpoints of the dotted arc crossing the edge. A special case is reflection across the vertical axis of the circular Farey diagram, $T(x/y) = y/x$. Reflection across an arc perpendicular to an edge $\langle a/b, c/d\rangle$ can be realized as STS^{-1} with S having matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ as before since STS^{-1} then interchanges a/b and c/d and is orientation-reversing. It is not hard to compute the matrix of STS^{-1} and we leave this as an exercise.



(3) The diagram can be rotated 180 degrees about the midpoint of any edge, interchanging the two adjacent triangles. This rotation is the composition of the reflection across this edge and the reflection across the arc perpendicular to the edge. Rotation about the midpoint of $\langle 1/0, 0/1\rangle$ is $T(x/y) = -y/x$ so rotation about the midpoint of an edge $\langle a/b, c/d\rangle$ is STS^{-1} with the same S as before since STS^{-1} interchanges the endpoints of $\langle a/b, c/d\rangle$ and is orientation-preserving.

(4) The diagram can be rotated by 120 degrees in either direction about the centerpoint of any triangle, the point of intersection of the three dotted arcs that cross the triangle in the figure above. In particular this rotates the triangle itself about its centerpoint. A simple case is the rotation of the triangle $\langle 1/0, 0/1, 1/1\rangle$ by 120 degrees counterclockwise. This is given by the transformation $T(x/y) = y/(y-x)$ with matrix $\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$ which has determinant 1 and takes the edge $\langle 1/0, 0/1\rangle$ to $\langle 0/1, 1/1\rangle$. For rotation of an arbitrary triangle $\langle a/b, c/d, e/f\rangle$ we may assume its vertices have been ordered to give it a counterclockwise orientation, so the transformation S with matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ takes $\langle 1/0, 0/1, 1/1\rangle$ to this triangle. Then STS^{-1} rotates $\langle a/b, c/d, e/f\rangle$ by

120 degrees counterclockwise since it is orientation-preserving and takes $\langle {}^a\!/_b, {}^c\!/_d \rangle$ to $\langle {}^c\!/_d, {}^e\!/_f \rangle$. Again the matrix for STS^{-1} could easily be computed.

(5) The diagram can be pivoted about any vertex v. If the vertices joined to v by edges are labeled v_i for all integers i, with v_i joined to v_{i+1} by an edge, then there is a pivoting transformation T sending each triangle $\langle v, v_i, v_{i+1} \rangle$ to the next triangle $\langle v, v_{i+1}, v_{i+2} \rangle$. The powers T^n are then also pivoting transformations sending $\langle v, v_i, v_{i+1} \rangle$ to $\langle v, v_{i+n}, v_{i+n+1} \rangle$ where n can be any nonzero integer, positive or negative. (When $n = 0$ one just has the identity transformation sending each vertex to itself.) For example, horizontal translation of the upper halfplane Farey diagram by any number of units to the right or left amounts to pivoting about the vertex ${}^1\!/_0$. The transformation T_n pivoting n steps counterclockwise about ${}^1\!/_0$ has matrix $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. For an arbitrary vertex ${}^a\!/_b$, if S is an orientation-preserving transformation taking ${}^1\!/_0$ to ${}^a\!/_b$ then S takes the infinite fan of triangles containing ${}^1\!/_0$ to the infinite fan containing ${}^a\!/_b$, so ST_nS^{-1} will pivot n steps counterclockwise about ${}^a\!/_b$. The different choices for S have matrices $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ with $ad - bc = 1$, so ST_nS^{-1} has matrix

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} a & na + c \\ b & nb + d \end{pmatrix}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} 1 - nab & na^2 \\ -nb^2 & 1 + nab \end{pmatrix}$$

where for the last equality we use the fact that $ad - bc = 1$. Note that c and d do not appear in the final answer, reflecting the fact that the pivoting transformation only depends on the pivot vertex ${}^a\!/_b$ and n. For example when ${}^a\!/_b = {}^0\!/_1$ we get the matrix $\begin{pmatrix} 1 & 0 \\ -n & 1 \end{pmatrix}$ for pivoting n steps counterclockwise about ${}^0\!/_1$.

Exercises

1. Find a formula for the linear fractional transformation that rotates the triangle $\langle {}^0\!/_1, {}^1\!/_2, {}^1\!/_1 \rangle$ to $\langle {}^1\!/_1, {}^0\!/_1, {}^1\!/_2 \rangle$.

2. Find the two orientation-reversing linear fractional transformations that take the edge $\langle {}^1\!/_2, {}^1\!/_3 \rangle$ to itself, possibly interchanging its two ends.

3. Find a formula for the linear fractional transformation that reflects the upper halfplane version of the Farey diagram across the vertical line $x = {}^3\!/_2$.

4. Compute the matrix of the transformation that reflects the Farey diagram across an arc perpendicular to an edge $\langle {}^a\!/_b, {}^c\!/_d \rangle$. Do the same for the 180 degree rotation about the centerpoint of this edge, and for the 120 degree rotation of a triangle $\langle {}^a\!/_b, {}^c\!/_d, {}^e\!/_f \rangle$.

5. Express the transformation $T({}^x\!/_y) = {}^{-y}\!/_x$ in four different ways as a composition of three pivoting transformations about ${}^1\!/_0$ or ${}^0\!/_1$.

6. (a) Find all the transformations in $LF(\mathbb{Z})$ that fix the vertex ${}^1\!/_0$, that is, take this vertex to itself.

(b) Find all the transformations in $LF(\mathbb{Z})$ that fix $0/1$.

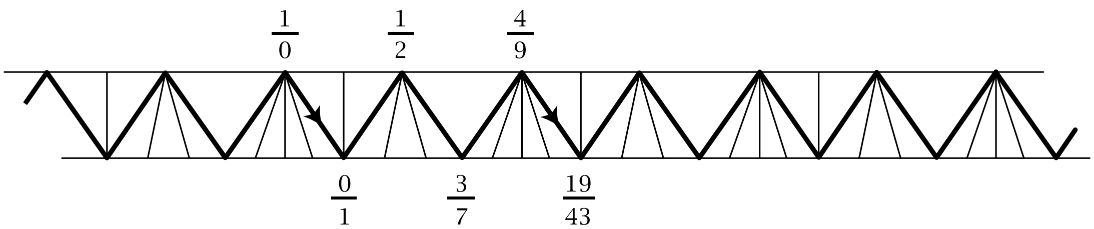
(c) Determine which of the transformations in (a) and (b) are reflections and describe these reflections.

(d) Show that if the transformation T fixes x/y then STS^{-1} fixes $S(x/y)$.

(e) Find all the transformations in $LF(\mathbb{Z})$ that fix $1/1$. Check that $T(x/y) = y/x$ is among the transformations you have found.

3.2 Translations and Glide Reflections

Linear fractional transformations can be used to compute the values of periodic or eventually periodic infinite continued fractions, and to see that these values are always quadratic irrational numbers. To illustrate this, consider the periodic continued fraction $\overline{1/2 + 1/3 + 1/1 + 1/4}$. The associated periodic strip in the Farey diagram can be extended to give an infinite strip that is periodic in both directions:



We would like to find a linear fractional transformation that gives the rightward translation of this strip that exhibits the periodicity. The only possibility is the transformation with matrix $\begin{pmatrix} 4 & 19 \\ 9 & 43 \end{pmatrix}$ since this sends the edge $\langle 1/0, 0/1 \rangle$ to $\langle 4/9, 19/43 \rangle$ and is orientation-preserving since the matrix has determinant 1 in view of the inequality $4/9 > 19/43$. This inequality can be verified either by a calculation or by visualizing how the strip lies inside the circular Farey diagram, with the part of the strip to the right of the edge $\langle 1/0, 0/1 \rangle$ lying in the upper half of the diagram.

To see that the transformation T with matrix $\begin{pmatrix} 4 & 19 \\ 9 & 43 \end{pmatrix}$ really does translate the strip along itself we can argue as follows. Let us label the ten triangles between the edges $\langle 1/0, 0/1 \rangle$ and $\langle 4/9, 19/43 \rangle$ as $t_1, t_2, \cdots, t_{10}$ from left to right, and then continue this labeling with the subsequent triangles $t_{11}, t_{12}, \cdots$ to the right. We can build the part of the strip to the right of the edge $\langle 1/0, 0/1 \rangle$ by starting with this edge and first adding the vertex v_1 just to the right of $1/0$ to form the triangle t_1, then adding the vertex v_2 to form t_2, and so on repeatedly, adding successive vertices v_i on one border of the strip or the other to form the successive triangles t_i. Since T is orientation-preserving and takes $\langle 1/0, 0/1 \rangle$ to $\langle 4/9, 19/43 \rangle$ it must take the triangle t_1 to the triangle t_{11} just to the right of the edge $\langle 4/9, 19/43 \rangle$, so T takes v_1 to v_{11}. The triangle t_2 must then be taken to t_{12} so v_2 is taken to v_{12}. In the same way we have $T(t_i) = t_{i+10}$ and $T(v_i) = v_{i+10}$ for all $i \geq 1$ so T translates the right half of the strip along itself. For the left half of the strip we can apply similar reasoning to T^{-1}. Thus T^{-1} sends t_{10} to the triangle just to the left of $\langle 1/0, 0/1\rangle$, then it sends t_9 to the second triangle to the left of $\langle 1/0, 0/1\rangle$, and so on. We conclude from all this that T is indeed a translation of the strip along itself.

The fractions labeling the vertices along the zigzag path in the strip moving toward the right are the convergents to $\overline{{}^1\!\nearrow_2 + {}^1\!\nearrow_3 + {}^1\!\nearrow_1 + {}^1\!\nearrow_4}$. Call these convergents $z_1, z_2, \cdots$ and their limit z. When we apply the translation T we are taking each convergent to a later convergent in the sequence, so both the sequence $\{z_n\}$ and the sequence $\{T(z_n)\}$ converge to z. On the other hand the sequence $\{T(z_n)\}$ converges to $T(z)$ since this is just saying that $\frac{4z_n+19}{9z_n+43}$ converges to $\frac{4z+19}{9z+43}$ as z_n converges to z. Thus we have $T(z) = z$.

In summary, what we have just argued is that the value z of the periodic continued fraction $\overline{{}^1\!\nearrow_2 + {}^1\!\nearrow_3 + {}^1\!\nearrow_1 + {}^1\!\nearrow_4}$ satisfies the equation $T(z) = z$, which is saying that z is a ***fixed point*** of the transformation T. Since $T(z) = \frac{4z+19}{9z+43}$ the equation $T(z) = z$ becomes $\frac{4z+19}{9z+43} = z$ which simplifies to $9z^2 + 39z - 19 = 0$. The roots of this equation are given by the quadratic formula:

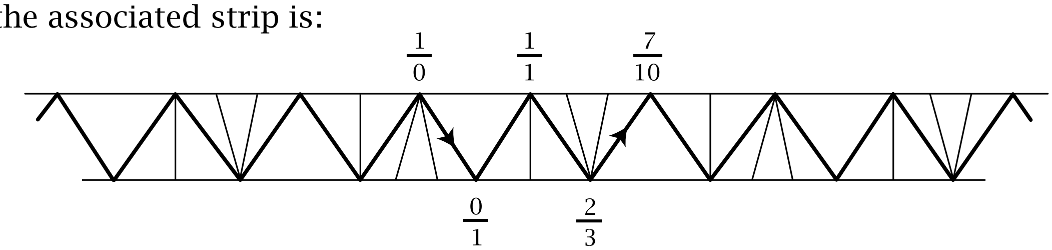
$$z = \frac{-39 \pm \sqrt{39^2 + 4\cdot 9\cdot 19}}{18} = \frac{-39 \pm 3\sqrt{13^2 + 4\cdot 19}}{18} = \frac{-13 \pm \sqrt{245}}{6} = \frac{-13 \pm 7\sqrt{5}}{6}$$

The positive root is the one that the right half of the infinite strip converges to, so we have determined the value of the continued fraction:

$$\overline{{}^1\!\nearrow_2 + {}^1\!\nearrow_3 + {}^1\!\nearrow_1 + {}^1\!\nearrow_4} = \frac{-13 + 7\sqrt{5}}{6}$$

The other root $(-13 - 7\sqrt{5})/6$ has an interpretation in terms of the diagram as well: It is the limit of the numbers labeling the vertices of the zigzag path moving off to the left rather than to the right. This follows by the same sort of argument as above.

A periodic continued fraction with period of odd length has an associated infinite strip with a different type of symmetry. As an example, consider $\overline{{}^1\!\nearrow_1 + {}^1\!\nearrow_2 + {}^1\!\nearrow_3}$. Here the associated strip is:

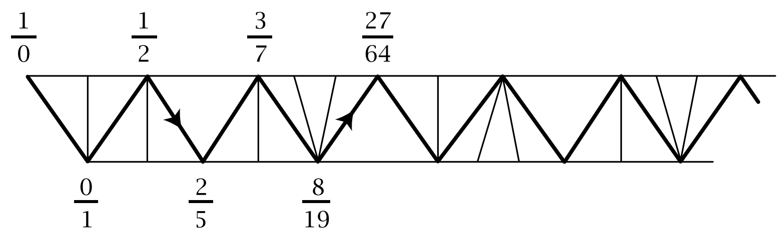


This strip is taken to itself by a transformation that takes $\langle 1/0, 0/1\rangle$ to $\langle 2/3, 7/10\rangle$ by combining a translation along the strip with reflection across the horizontal axis of the strip. A transformation of this type is called a ***glide reflection.*** The only linear fractional transformation that could realize this glide reflection is the transformation with matrix $\begin{pmatrix} 2 & 7 \\ 3 & 10 \end{pmatrix}$ since this takes $\langle 1/0, 0/1\rangle$ to $\langle 2/3, 7/10\rangle$ and is orientation-reversing as its determinant is -1. To check that this transformation gives a glide reflection of the strip one can argue as in the preceding example that each successive triangle to the right or left of $\langle 1/0, 0/1\rangle$ is moved along the strip in the same way that the glide reflection moves it, keeping in mind that orientations are now being reversed by both the glide reflection and the linear fractional transformation. This reasoning shows more generally that the translation or glide reflection symmetry of any periodic infinite strip in the Farey diagram can be realized by a linear fractional transformation.

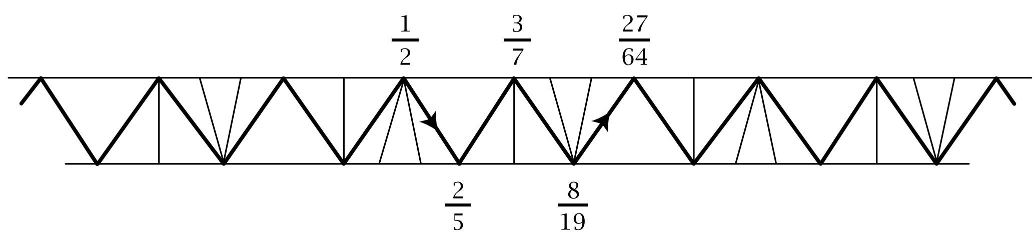
Just as in the preceding example the value of the continued fraction can be determined by solving the equation $T(z) = z$ where T is now the glide reflection. Thus we have $\frac{2z+7}{3z+10} = z$ which simplifies to $3z^2 + 8z - 7 = 0$ with roots $(-4 \pm \sqrt{37})/3$. The positive root gives the value of the continued fraction:

$$\overline{\tfrac{1}{1} + \tfrac{1}{2} + \tfrac{1}{3}} = \frac{-4+\sqrt{37}}{3}$$

Continued fractions that are only eventually periodic can be treated in a similar fashion. For example, consider $\tfrac{1}{2} + \tfrac{1}{2} + \overline{\tfrac{1}{1} + \tfrac{1}{2} + \tfrac{1}{3}}$. The corresponding infinite strip is:



In this case if we discard the triangles corresponding to the initial nonperiodic part of the continued fraction, $\tfrac{1}{2} + \tfrac{1}{2}$, and then extend the remaining periodic part in both directions, we obtain a periodic strip that is carried to itself by the glide reflection T taking $\langle 1/2, 2/5\rangle$ to $\langle 8/19, 27/64\rangle$:



We can compute T as a composition of two transformations realizing the two-step combination $\langle 1/2, 2/5\rangle \to \langle 1/0, 0/1\rangle \to \langle 8/19, 27/64\rangle$. Thus we consider the product

$$\begin{pmatrix} 8 & 27 \\ 19 & 64 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} 8 & 27 \\ 19 & 64 \end{pmatrix}\begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -14 & 11 \\ -33 & 26 \end{pmatrix}$$

so we have $T(z) = \frac{-14z+11}{-33z+26}$. This transformation has determinant -1 so it is the glide reflection we want. Now we solve $T(z) = z$, or $\frac{-14z+11}{-33z+26} = z$, which reduces to $33z^2 - 40z + 11 = 0$ with roots $z = (20 \pm \sqrt{37})/33$. Both roots are positive, and we want the smaller one, $(20 - \sqrt{37})/33$, because along the top edge of the periodic strip the numbers decrease as we move to the right approaching the smaller root and they increase as we move to the left approaching the larger root. Thus we have:

$$\tfrac{1}{2} + \tfrac{1}{2} + \overline{\tfrac{1}{1} + \tfrac{1}{2} + \tfrac{1}{3}} = \frac{20-\sqrt{37}}{33}$$

Notice that $\sqrt{37}$ occurs in both this example and the preceding one where we computed the value of $\overline{{}^{1}\!\!\nearrow_{1} + {}^{1}\!\!\nearrow_{2} + {}^{1}\!\!\nearrow_{3}}$. The explanation for this is that to get from $\overline{{}^{1}\!\!\nearrow_{1} + {}^{1}\!\!\nearrow_{2} + {}^{1}\!\!\nearrow_{3}}$ to ${}^{1}\!\!\nearrow_{2} + {}^{1}\!\!\nearrow_{2} + \overline{{}^{1}\!\!\nearrow_{1} + {}^{1}\!\!\nearrow_{2} + {}^{1}\!\!\nearrow_{3}}$ one adds 2 and inverts, then adds 2 and inverts again, and each of these operations of adding an integer or taking the reciprocal takes place within the set $\mathbb{Q}(\sqrt{37})$ of all numbers of the form $a + b\sqrt{37}$ with a and b rational. More generally, this argument shows that any eventually periodic continued fraction whose periodic part is $\overline{{}^{1}\!\!\nearrow_{1} + {}^{1}\!\!\nearrow_{2} + {}^{1}\!\!\nearrow_{3}}$ has as its value some number in $\mathbb{Q}(\sqrt{37})$. However, not all irrational numbers in $\mathbb{Q}(\sqrt{37})$ have eventually periodic continued fractions with periodic part $\overline{{}^{1}\!\!\nearrow_{1} + {}^{1}\!\!\nearrow_{2} + {}^{1}\!\!\nearrow_{3}}$. For example, the continued fraction for $\sqrt{37}$ itself is $6 + \overline{{}^{1}\!\!\nearrow_{12}}$, with a different periodic part. (This can be checked by computing the value of this continued fraction by the method above.)

The procedure we have used in these examples works in general for any irrational number z whose continued fraction is eventually periodic. From the periodic part of the continued fraction one constructs a periodic infinite strip in the Farey diagram, where the periodicity is given by a transformation $T(z) = {}^{az+b}\!/_{cz+d}$ in $LF(\mathbb{Z})$, with T either a translation or a glide reflection of the strip. As we argued in the first example, the number z satisfies the equation $T(z) = z$. This becomes the quadratic equation $az + b = cz^2 + dz$ with integer coefficients, or in more standard form, $cz^2 + (d-a)z - b = 0$. We would like to apply the quadratic formula to find the roots of this equation, but in order to do this the coefficient c must be nonzero. Suppose on the contrary that c was zero. Then the determinant condition $ad - bc = \pm 1$ would force a to be ± 1, and then from the first column of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \pm 1 & b \\ 0 & d \end{pmatrix}$ we see that T would take the vertex ${}^{\pm 1}\!/_{0}$ of the Farey diagram to itself. However a translation or glide reflection symmetry of a periodic infinite strip cannot take any vertex to itself since no vertex along the strip is taken to itself, and the other vertices lie in the complement of the strip which consists of disjoint pieces, each containing all the vertices lying on one side of an edge in the border of the strip, and a translation or glide reflection of the strip takes each of these pieces to a different piece.

Knowing that c is nonzero, we can apply the quadratic formula to deduce that the roots of the equation $cz^2 + (d-a)z - b = 0$ have the form $A + B\sqrt{n}$ with A and B rational numbers and n an integer. We know that the real number z defined by the given continued fraction is a root of the equation so n cannot be negative, and it cannot be a square since z is irrational.

Thus we have an argument that proves one half of Lagrange's Theorem:

Proposition 3.4. *A real number whose continued fraction is periodic or eventually periodic is a quadratic irrational.*

The converse statement that the continued fraction for every quadratic irrational is periodic or eventually periodic will be proved in Proposition 4.1 and Theorem 5.2.

As we saw above, the equation $T(z) = z$ for a fixed point of a transformation $T(z) = {}^{az+b}/_{cz+d}$ in $LF(\mathbb{Z})$ is $cz^2 + (d-a)z - b = 0$. This has roots $z = \frac{a-d\pm\sqrt{(d-a)^2+4bc}}{2c}$. If we let $ad - bc = \varepsilon = \pm 1$ then $bc = ad - \varepsilon$ and the roots can be rewritten as $z = \frac{a-d\pm\sqrt{(a+d)^2-4\varepsilon}}{2c}$. The discriminant $\delta = (a+d)^2 - 4\varepsilon$ determines the nature of the roots. If $\delta > 0$ there are two real roots, the situation we have been considering for translations and glide reflections. If $\delta = 0$ the two roots coalesce to a single root, the rational number ${}^{a-d}/_{2c}$. And if $\delta < 0$ there are no real roots, only complex roots. Thus the numbers $a+d$ and ε determine how many fixed points there are.

The number $a+d$ is called the *trace* of the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. A matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and its negative $-\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ determine the same transformation in $LF(\mathbb{Z})$, and changing a matrix to its negative changes the sign of the trace, so only the absolute value of the trace is well defined for elements of $LF(\mathbb{Z})$. We will usually assume the sign of a matrix is chosen to make the trace nonnegative.

Proposition 3.5. *The various types of transformations in* $LF(\mathbb{Z})$ *are distinguished by their determinants and traces according to the following table*:

	determinant	trace
180 degree rotation	$+1$	0
120 degree rotation	$+1$	1
pivot	$+1$	2
translation	$+1$	> 2
reflection	-1	0
glide reflection	-1	> 0

Proof: A general fact about the trace is that $\mathrm{trace}(AB) = \mathrm{trace}(BA)$ for matrices A and B. This can be checked by a direct calculation which we leave to the reader. A consequence is that $\mathrm{trace}(ABA^{-1}) = \mathrm{trace}(B)$ since the traces of $(AB)A^{-1}$ and $A^{-1}(AB)$ are equal.

We can apply this to get four of the six rows in the table as follows. As we saw in the previous section, every 180 degree rotation can be expressed as STS^{-1} for $T({}^{x}/_{y}) = {}^{-y}/_{x}$ and S some element of $LF(\mathbb{Z})$. The matrix of T is $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ with trace 0 so this is also the trace of STS^{-1}. This gives the first row of the table. For the second row we argue in the same way using $T({}^{x}/_{y}) = {}^{y}/_{y-x}$ with matrix $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right)$ of trace 1. This is a 120 degree rotation counterclockwise, and for the 120 degree rotation in the opposite direction we use the matrix $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ which has the same trace. For pivoting transformations we use a matrix $\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$ of trace 2. For the two kinds of reflections we use $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ of trace 0.

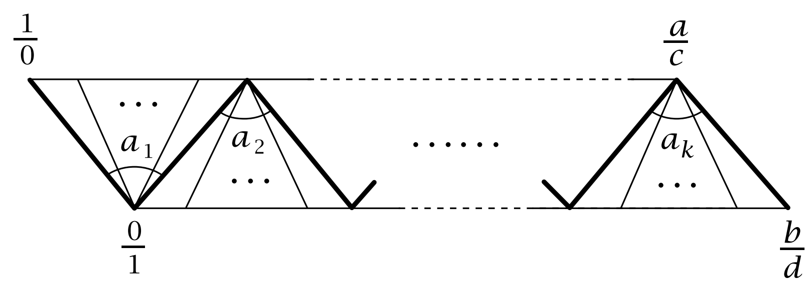
Translations have two distinct fixed points so the discriminant $(a+d)^2-4$ must be positive, which means $|a+d|>2$. Since we are taking traces to be nonnegative this condition becomes $a+d>2$. Glide reflections have discriminant $(a+d)^2+4$ which is always positive, but $a+d$ cannot be 0, otherwise the discriminant would be 4, a square, so fixed points would be rational, contradicting the fact that the fixed points of a glide reflection are irrational since their continued fractions are infinite. Thus $a+d>0$ for a glide reflection. $\square$

All combinations of trace and determinant can be realized using the simple matrices $\begin{pmatrix} 0 & -1 \\ 1 & n \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix}$ of trace n and determinant ± 1. An exercise at the end of this section is to determine exactly what these transformations of the Farey diagram look like.

It is a fact that every symmetry of the Farey diagram other than the identity transformation is of one of the six types listed in the previous table. From this it follows that one can determine the type of any given transformation just by computing its determinant and trace. It is not too difficult to prove that there are no other kinds of symmetries, but this fact will not be needed later in the book so we will not digress to give a proof here.

Factoring Translations and Glide Reflections

Let us show how translations and glide reflections can be realized as products of simpler transformations. Consider a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad-bc=\pm 1$ and all four entries a,b,c,d positive integers. There is then a strip in the upper half of the circular Farey diagram connecting the edge $\langle 1/0, 0/1\rangle$ to the edge $\langle a/c, b/d\rangle$. One possible configuration for this strip is the following:



Here the first fan in the strip opens upward and the last fan opens downward, but there are three other possibilities depending on whether the first and last fans open upward or downward. When $a/c > b/d$ as in the figure, then $ad-bc=+1$ so the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ defines an orientation-preserving transformation in $LF(\mathbb{Z})$ taking the edge $\langle 1/0, 0/1\rangle$ to $\langle a/c, b/d\rangle$. This is a translation of the infinite periodic strip obtained by extending the finite strip from $\langle 1/0, 0/1\rangle$ to $\langle a/c, b/d\rangle$ periodically in both directions.

We can move the edge $\langle 1/0, 0/1\rangle$ to $\langle a/c, b/d\rangle$ by a sequence of pivoting transformations, one for each fan. One first pivots the edge $\langle 1/0, 0/1\rangle$ across a fan of a_1 triangles to the second edge of the zigzag path, then this edge is pivoted across the

a_2 triangles in the second fan to the next edge of the zigzag path, and so on until we reach the right edge $\langle a/c, b/d\rangle$. These pivotings are alternately in the clockwise and counterclockwise direction, and the simplest pivotings of these two types are given by matrices $\begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ with $n > 0$, pivoting n steps clockwise about $0/1$ or counterclockwise about $1/0$ in the two cases. For the configuration of fans shown in the figure, let us consider the following product:

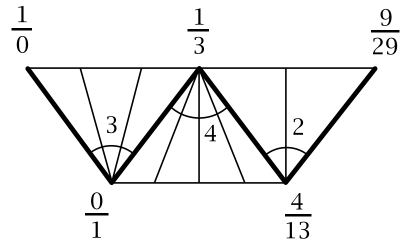
$$\begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix}\begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ a_3 & 1 \end{pmatrix}\begin{pmatrix} 1 & a_4 \\ 0 & 1 \end{pmatrix}\cdots\begin{pmatrix} 1 & 0 \\ a_{k-1} & 1 \end{pmatrix}\begin{pmatrix} 1 & a_k \\ 0 & 1 \end{pmatrix}$$

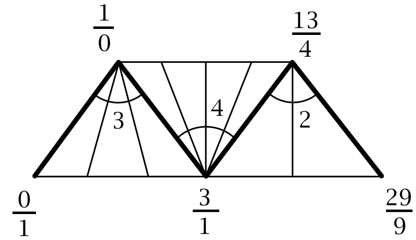
These matrices determine pivoting transformations that alternate between clockwise and counterclockwise as they should, with the number of steps being $a_1, a_2, \cdots, a_k$ as we want. However there seem to be two things wrong with this product. First, the order of the terms appears to be backwards since when we compose transformations we proceed from right to left, so this product would first pivot a_k steps, then a_{k-1} steps, and so on, whereas we want to move the edge $\langle 1/0, 0/1\rangle$ across the strip by first pivoting a_1 steps, then a_2 steps, and so on. The other problem is that each pivoting transformation in the product is pivoting about either $0/1$ or $1/0$ whereas the pivotings that move the $\langle 1/0, 0/1\rangle$ edge across the strip are pivoting about a sequence of different vertices.

Surprisingly enough, these two problems cancel each other out, and the product displayed above is actually correct and does equal $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. To see why, suppose we superimpose a copy of the strip on top of the circular Farey diagram, but with the right edge $\langle a/c, b/d\rangle$ lying on top of the edge $\langle 1/0, 0/1\rangle$ and each triangle in the rest of the strip lying exactly on top of a corresponding triangle in the lower half of the diagram. If we apply the last matrix of the product to this repositioned strip, this pivots the strip so that the next-to-last edge of the zigzag path lies on top of $\langle 1/0, 0/1\rangle$. Then applying the next-to-last matrix in the product to the newly positioned strip pivots it so that the third-to-last edge of the zigzag path lies on top of $\langle 1/0, 0/1\rangle$. Continuing in this way, we end up with the left edge of the strip lying on top of $\langle 1/0, 0/1\rangle$. This means that the product of all the matrices takes the strip back to its original position, so the product takes $\langle 1/0, 0/1\rangle$ to the right edge of the strip, as we wanted.

The other three possibilities for whether the first and last fans open upward or downward are treated in a similar fashion. For each fan opening upward one uses a matrix $\begin{pmatrix} 1 & 0 \\ a_i & 1 \end{pmatrix}$ giving a pivoting transformation about $0/1$ and for each fan opening downward one uses a matrix $\begin{pmatrix} 1 & a_i \\ 0 & 1 \end{pmatrix}$ pivoting about $1/0$.

As an example consider the matrix $\begin{pmatrix} 9 & 4 \\ 29 & 13 \end{pmatrix}$ which has determinant 1 and corresponds to the edge $\langle 9/29, 4/13\rangle$ with $9/29 > 4/13$. The corresponding strip in the Farey diagram is obtained by computing the continued fraction $9/29 = 1/\!\!\nearrow 3 + 1/\!\!\nearrow 4 + 1/\!\!\nearrow 2$ as in the first figure below:





From this we can read off that $\left(\begin{smallmatrix}9&4\\29&13\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&0\\3&1\end{smallmatrix}\right)\left(\begin{smallmatrix}1&4\\0&1\end{smallmatrix}\right)\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$. Similarly, for $\left(\begin{smallmatrix}13&29\\4&9\end{smallmatrix}\right)$ we have $29/9 = 3 + 1/\!\!\nearrow\!4 + 1/\!\!\nearrow\!2$ as in the second figure so $\left(\begin{smallmatrix}13&29\\4&9\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&3\\0&1\end{smallmatrix}\right)\left(\begin{smallmatrix}1&0\\4&1\end{smallmatrix}\right)\left(\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right)$. In both these cases the first and last fans in the strip open in the same direction, so if we extend the strip to an infinite periodic strip, this would produce adjacent fans with three and two triangles opening in the same direction, and each of these pairs of fans could be combined to give a single fan with five triangles.

Glide reflection symmetries of infinite periodic strips cannot be expressed as products of pivoting transformations since pivotings are orientation-preserving, but glide reflections can be expressed as products of simple glide reflections that, like pivotings, move an edge across a single fan but are orientation-reversing. An example is the transformation with matrix $\left(\begin{smallmatrix}0&1\\1&n\end{smallmatrix}\right)$ for an integer $n > 0$. This transformation takes the edge $\langle 1/0, 0/1\rangle$ to $\langle 0/1, 1/n\rangle$ and is orientation-reversing, a glide reflection symmetry of an infinite strip in which each fan has n triangles. A transformation with matrix $\left(\begin{smallmatrix}n&1\\1&0\end{smallmatrix}\right)$ has similar behavior, taking $\langle 1/0, 0/1\rangle$ to $\langle n/1, 1/0\rangle$.

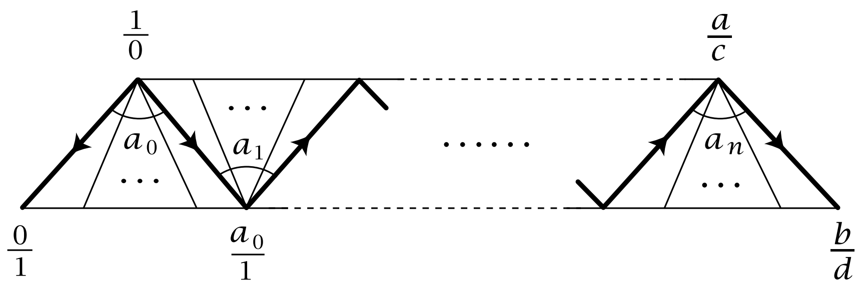
For example, the matrix $\left(\begin{smallmatrix}4&9\\13&29\end{smallmatrix}\right)$ of determinant -1 gives a glide reflection taking the left edge of the first strip in the preceding figure to the right edge. This glide reflection is a symmetry of the infinite strip obtained by first applying the glide reflection to the given strip to get a strip twice as long, then taking the periodic extension of this doubled strip in both directions. The corresponding factorization of $\left(\begin{smallmatrix}4&9\\13&29\end{smallmatrix}\right)$ is $\left(\begin{smallmatrix}4&9\\13&29\end{smallmatrix}\right) = \left(\begin{smallmatrix}0&1\\1&3\end{smallmatrix}\right)\left(\begin{smallmatrix}0&1\\1&4\end{smallmatrix}\right)\left(\begin{smallmatrix}0&1\\1&2\end{smallmatrix}\right)$, as one can check by the method we used in the case of translations of a periodic strip, placing a copy of the strip on top of the Farey diagram with the right edge of the strip on top of the edge $\langle 1/0, 0/1\rangle$, but with the copy flipped over since we are now dealing with a glide reflection.

More generally, for any matrix $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ of positive integers with determinant ± 1 there is an associated strip from the edge $\langle 1/0, 0/1\rangle$ to $\langle a/c, b/d\rangle$, and we can express this matrix as a product of the basic matrices $\left(\begin{smallmatrix}1&n\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&1\\1&n\end{smallmatrix}\right)$, or $\left(\begin{smallmatrix}n&1\\1&0\end{smallmatrix}\right)$, by putting arrows on the edges of the zigzag path in the strip to indicate orientations of the edges, with the left edge oriented from $1/0$ to $0/1$ and the right edge oriented from a/c to b/d and the intermediate edges oriented arbitrarily. These orientations, together with the directions that the fans open, determine the factors $\left(\begin{smallmatrix}1&n\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&1\\1&n\end{smallmatrix}\right)$, or $\left(\begin{smallmatrix}n&1\\1&0\end{smallmatrix}\right)$ in the product representing $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$.

As an example, in the proof of Theorem 2.1 we made use of the following product:

$$\begin{pmatrix}1&a_0\\0&1\end{pmatrix}\begin{pmatrix}0&1\\1&a_1\end{pmatrix}\begin{pmatrix}0&1\\1&a_2\end{pmatrix}\cdots\begin{pmatrix}0&1\\1&a_n\end{pmatrix}$$

The corresponding strip is



with the last fan on the right opening either downward as shown or possibly upward, depending on whether n is even or odd. The first fan has both its left and right edges oriented downward so the first matrix in the product gives the corresponding pivoting transformation, but all the other fans have both edges oriented to the right so they correspond to glide reflections, the other matrices in the product. If $a_0 = 0$ the first matrix is the identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ so it can be omitted along with the first fan.

Exercises

1. Compute the value of each of the following continued fractions by first drawing the associated infinite strip of triangles, then finding a linear fractional transformation T in $LF(\mathbb{Z})$ that gives the periodicity in the strip, then solving $T(z) = z$.

(a) $\overline{\frac{1}{2} + \frac{1}{5}}$

(b) $\overline{\frac{1}{2} + \frac{1}{1} + \frac{1}{1}}$

(c) $\overline{\frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{2}}$

(d) $2 + \overline{\frac{1}{1} + \frac{1}{1} + \frac{1}{4}}$

(e) $2 + \overline{\frac{1}{1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{4}}$

(f) $\frac{1}{1} + \frac{1}{1} + \overline{\frac{1}{2} + \frac{1}{3}}$

2. Find an infinite periodic strip of triangles in the Farey diagram such that the transformation $\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ is a glide reflection along this strip and $\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ is a translation along the strip.

3. In an example in this section we computed the value of the continued fraction $\overline{\frac{1}{1} + \frac{1}{2} + \frac{1}{3}}$ to be $(-4+\sqrt{37})/3$ using the infinite periodic strip of triangles associated to this continued fraction. Use the same periodic strip to compute the continued fraction for $(-4-\sqrt{37})/3$ at the opposite end of the strip.

4. Draw pictures showing how the transformations with matrices $\begin{pmatrix} 0 & -1 \\ 1 & n \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix}$ act on the Farey diagram for all $n \geq 0$. For example, when the transformation is a translation draw the periodic strip.

5. Express the following transformations as compositions of pivot transformations:

(a) $T(x/y) = \frac{13x+3y}{69x+16y}$

(b) $T(x/y) = \frac{10x+33y}{33x+109y}$

4 Quadratic Forms

Finding Pythagorean triples is answering the question of when the sum of two squares is equal to a square. This leads naturally to the broader question of exactly which numbers are sums of two squares. Thus one asks, when does an equation $x^2 + y^2 = n$ have integer solutions, and how can one find these solutions? The brute force approach of simply plugging in values for x and y leads to the following list of all solutions for $n \le 50$ (apart from interchanging x and y):

$$\mathbf{1} = 1^2 + 0^2,\ \mathbf{2} = 1^2 + 1^2,\ \mathbf{4} = 2^2 + 0^2,\ \mathbf{5} = 2^2 + 1^2,\ \mathbf{8} = 2^2 + 2^2,\ \mathbf{9} = 3^2 + 0^2,$$
$$\mathbf{10} = 3^2 + 1^2,\ \mathbf{13} = 3^2 + 2^2,\ \mathbf{16} = 4^2 + 0^2,\ \mathbf{17} = 4^2 + 1^2,\ \mathbf{18} = 3^2 + 3^2,$$
$$\mathbf{20} = 4^2 + 2^2,\ \mathbf{25} = 5^2 + 0^2 = 4^2 + 3^2,\ \mathbf{26} = 5^2 + 1^2,\ \mathbf{29} = 5^2 + 2^2,\ \mathbf{32} = 4^2 + 4^2,$$
$$\mathbf{34} = 5^2 + 3^2,\ \mathbf{36} = 6^2 + 0^2,\ \mathbf{37} = 6^2 + 1^2,\ \mathbf{40} = 6^2 + 2^2,\ \mathbf{41} = 5^2 + 4^2,$$
$$\mathbf{45} = 6^2 + 3^2,\ \mathbf{49} = 7^2 + 0^2,\ \mathbf{50} = 5^2 + 5^2 = 7^2 + 1^2$$

Notice that in some cases there is more than one way to write n as a sum of two squares. Our first goal will be to describe a more efficient way to find the integer solutions of $x^2+y^2 = n$ and to display them graphically in a way that helps illuminate their structure. The technique for doing this will work not just for the function x^2+y^2 but also for any function $Q(x,y) = ax^2 + bxy + cy^2$, where a, b, and c are integer constants. Such a function $Q(x,y)$ with at least one of the coefficients a, b, c nonzero is called a ***quadratic form***, or more briefly, just a ***form***.

Solving $x^2 + y^2 = n$ amounts to representing n as the sum of two squares. More generally, solving $Q(x,y) = n$ is called *representing n by the form $Q(x,y)$*. So the overall goal is to solve the *representation problem*: Which numbers n are represented by a given form $Q(x,y)$, and how does one find such representations? Since every quadratic form $Q(x,y)$ has $Q(0,0) = 0$, the pair $(x,y) = (0,0)$ is not very interesting, so we will always assume implicitly that $(x,y) \neq (0,0)$, as we did for the list of solutions of $x^2 + y^2 = n$ above.

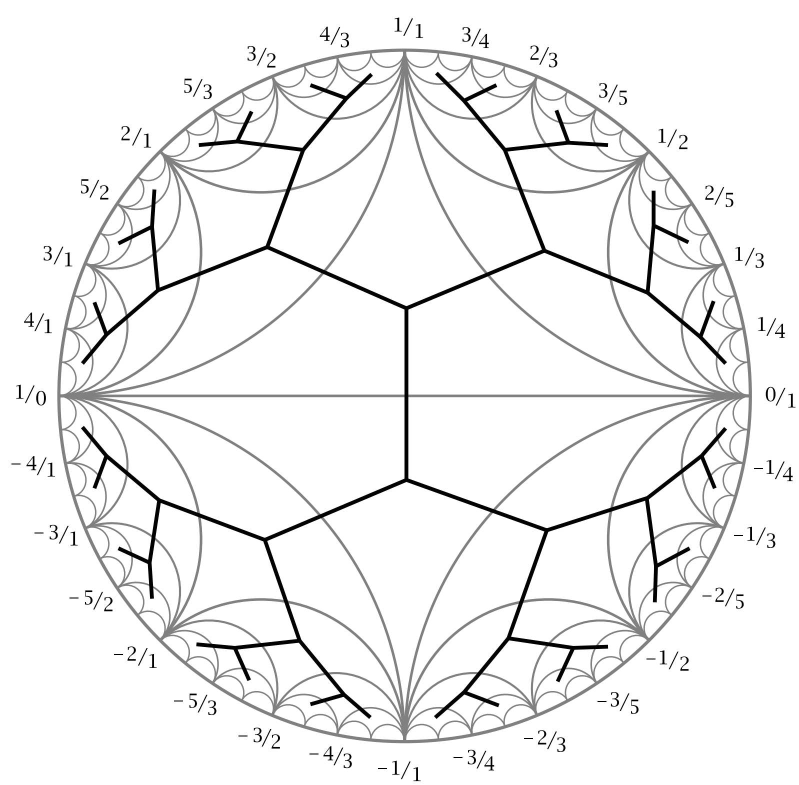
Before starting to describe the method for displaying the values of a quadratic form graphically, let us make a preliminary observation: If the greatest common divisor of two integers x and y is d, then we can write $x = dx'$, $y = dy'$, and $Q(x,y) = d^2Q(x',y')$ where the greatest common divisor of x' and y' is 1. Hence it suffices to find the values of Q on ***primitive*** pairs (x,y), the pairs whose greatest common divisor is 1, and then multiply these values by arbitrary squares d^2.

In a similar way, if the coefficients a, b, c of a form $Q(x,y) = ax^2 + bxy + cy^2$ have greatest common divisor d, so $a = da'$, $b = db'$, and $c = dc'$ for integers a', b', c' whose greatest common divisor is 1, then $Q(x,y) = d(a'x^2 + b'xy + c'y^2) = dQ'(x,y)$ for the form $Q'(x,y) = a'x^2 + b'xy + c'y^2$. Multiplying all the values of a form by a constant d is a fairly trivial operation, so for most purposes it suffices to restrict attention to forms for which the greatest common divisor of the coefficients is 1. Such forms are called ***primitive forms***.

Primitive pairs (x,y) correspond almost exactly to fractions x/y in lowest terms, the only ambiguity being that both (x,y) and $(-x,-y)$ correspond to the same fraction $x/y = -x/-y$. However, this ambiguity does not affect the value of a quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ since $Q(x,y) = Q(-x,-y)$. This means that we can regard $Q(x,y)$ as being essentially a function $f(x/y)$. Notice that we are not excluding the possibilities $(x,y) = (1,0)$ and $(x,y) = (-1,0)$ which correspond to the "fractions" $1/0$ and $-1/0$. There will be no need to distinguish between $1/0$ and $-1/0$ since $Q(1,0) = Q(-1,0)$.

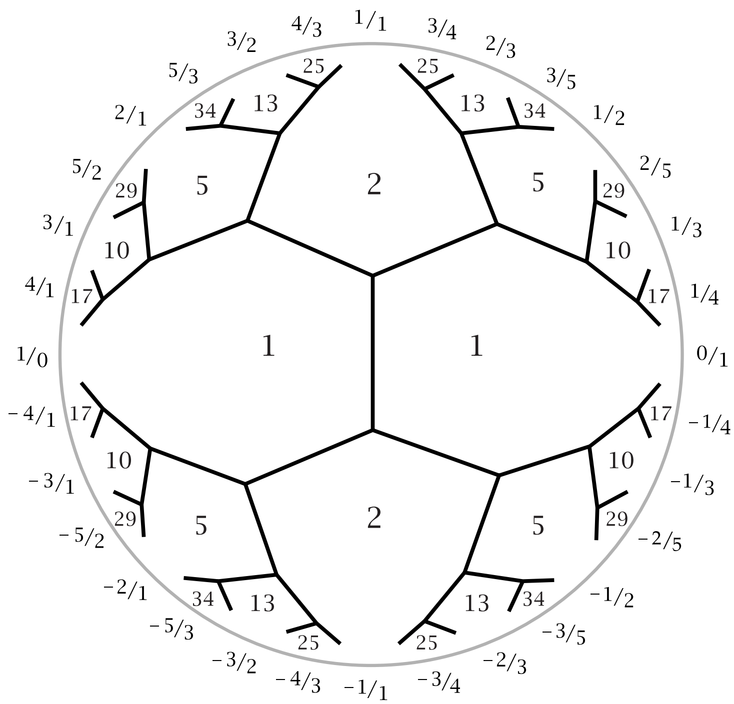
4.1 The Topograph

We already have a nice graphical representation of rational numbers x/y along with $\pm 1/0$ as the vertices in the Farey diagram. Here is a picture of the Farey diagram with the so-called *dual tree* superimposed:



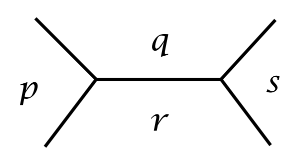
The dual tree has a vertex in the center of each triangle of the Farey diagram, and it has an edge crossing each edge of the Farey diagram. As with the Farey diagram, we can only draw a finite part of the dual tree. The actual dual tree has branching that repeats infinitely often with smaller and smaller branches.

The tree divides the interior of the large circle into regions, each of which is adjacent to one vertex of the original diagram. We can write the value $Q(x,y)$ in the region adjacent to the vertex x/y. This is shown in the figure below for the quadratic form $Q(x,y) = x^2+y^2$, where to unclutter the picture we no longer draw the triangles of the original Farey diagram.



For example the 13 in the region adjacent to the fraction $2/3$ represents the value 2^2+3^2, and the 29 in the region adjacent to $5/2$ represents the value 5^2+2^2.

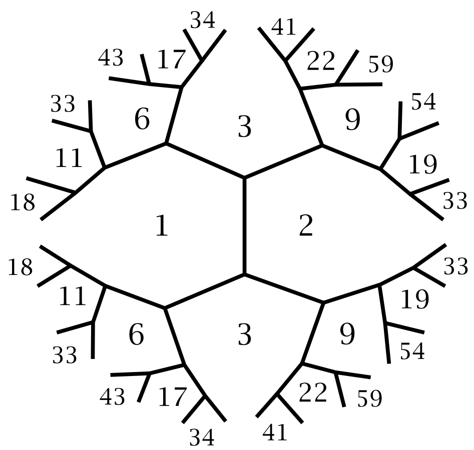
For a quadratic form Q this picture showing the values $Q(x,y)$ is called the ***topograph*** of Q. It turns out that there is a very simple method for computing the topograph from just a very small amount of initial data. This method is based on the following ***arithmetic progression rule***: If the values of $Q(x,y)$ in the four regions surrounding an edge in the tree are p, q, r, and s as indicated in the figure at the right, then the three numbers p, $q+r$, s form an arithmetic progression.



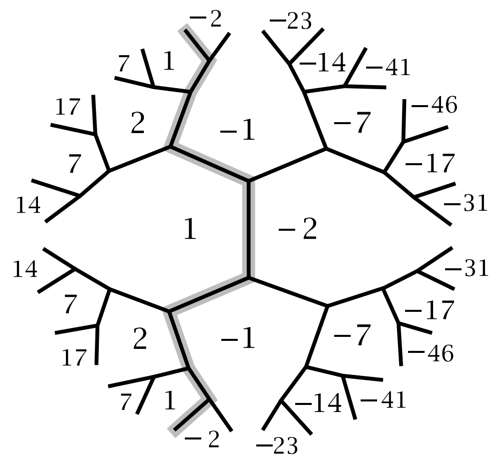
We can check this in the topograph of x^2+y^2 shown above. Consider for example one of the edges separating the values 1 and 2. The values in the four regions surrounding this edge are $1,1,2,5$ and the arithmetic progression is $1, 1+2, 5$. For an edge separating the values 1 and 5 the arithmetic progression is $2, 1+5, 10$. For an edge separating the values 5 and 13 the arithmetic progression is $2, 5+13, 34$. And similarly for all the other edges.

The arithmetic progression rule implies that the values of Q in the three regions surrounding a single vertex of the tree determine the values in all other regions, by starting at the vertex where the three adjacent values are known and working one's way outward in the dual tree. The easiest place to start for a quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ is with the three values $Q(1,0) = a$, $Q(0,1) = c$, and $Q(1,1) = a + b + c$ for the three fractions $1/0$, $0/1$, and $1/1$. Here are two examples:

$Q(x,y) = x^2 + 2y^2$



$Q(x,y) = x^2 - 2y^2$



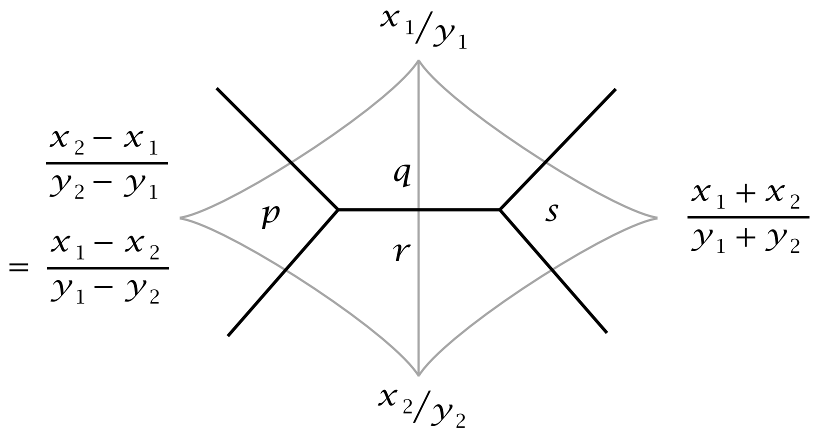
In the first case we start with the values 1 and 2 together with the 3 just above them. These determine the value 9 above the 2 via the arithmetic progression 1, $2+3$, 9. Similarly the 6 above the 1 is determined by the arithmetic progression 2, $1+3$, 6. Next one can fill in the 19 next to the 9 we just computed, using the arithmetic progression 3, $2+9$, 19, and so on for as long as one likes.

The procedure for the other form $x^2 - 2y^2$ is just the same, but here there are negative as well as positive values. The edges that separate positive values from negative values will be important later, so we have indicated these edges by special shading.

Perhaps the most noticeable thing in both the examples $x^2 + 2y^2$ and $x^2 - 2y^2$ is the fact that the values in the lower half of the topograph are the same as those in the upper half. We could have predicted in advance that this would happen because $Q(x,y) = Q(-x,y)$ whenever $Q(x,y) = ax^2 + cy^2$, with no xy term. The topograph for $x^2 + y^2$ has even more symmetry since the values of $x^2 + y^2$ are unchanged when x and y are switched, so the topograph has left-right symmetry as well.

Given any three integers a, b, and c which are not all zero, there is always a quadratic form whose topograph has these three numbers surrounding a vertex since the form $ax^2 + (c-a-b)xy + by^2$ takes the values a, b, and c for (x,y) equal to $(1,0)$, $(0,1)$, and $(1,1)$.

Now let us prove the arithmetic progression rule. Let the two vertices of the Farey diagram corresponding to the values q and r have labels x_1/y_1 and x_2/y_2 as in the following figure:



Then by the mediant rule for labeling vertices, the labels on the p and s regions are the fractions shown. Note that these labels are correct even when $x_1/y_1 = 1/0$ and $x_2/y_2 = 0/1$. For a quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ we then have:

$$s = Q(x_1 + x_2, y_1 + y_2) = a(x_1 + x_2)^2 + b(x_1 + x_2)(y_1 + y_2) + c(y_1 + y_2)^2$$
$$= \underbrace{ax_1^2 + bx_1y_1 + cy_1^2}_{Q(x_1,y_1) = q} + \underbrace{ax_2^2 + bx_2y_2 + cy_2^2}_{Q(x_2,y_2) = r} + (\cdots)$$

Similarly, we have:

$$p = Q(x_1 - x_2, y_1 - y_2) = \underbrace{ax_1^2 + bx_1y_1 + cy_1^2}_{Q(x_1,y_1) = q} + \underbrace{ax_2^2 + bx_2y_2 + cy_2^2}_{Q(x_2,y_2) = r} - (\cdots)$$

The omitted terms in $(\cdots)$ are the same in both cases, namely the terms involving both subscripts 1 and 2. If we compute $p + s$ by adding the two formulas together, the terms $(\cdots)$ will cancel, leaving just $p + s = (q + r) + (q + r)$. This equation can be rewritten as $(q + r) - p = s - (q + r)$, which just says that $p, q + r, s$ is an arithmetic progression. □

Exercises

1. Draw the topograph for the form $Q(x,y) = 2x^2 + 5y^2$, showing all the values of $Q(x,y) \le 60$ in the topograph, with the associated fractional labels x/y. If there is symmetry in the topograph, you only need to draw one half of the topograph and state that the other half is symmetric.

2. Do the same for the form $Q(x,y) = 2x^2 + xy + 2y^2$, in this case displaying all values $Q(x,y) \le 40$ in the topograph.

3. Do the same for the form $Q(x,y) = x^2 - y^2$, showing all the values between $+30$ and -30 in the topograph, but omitting the labels x/y this time.

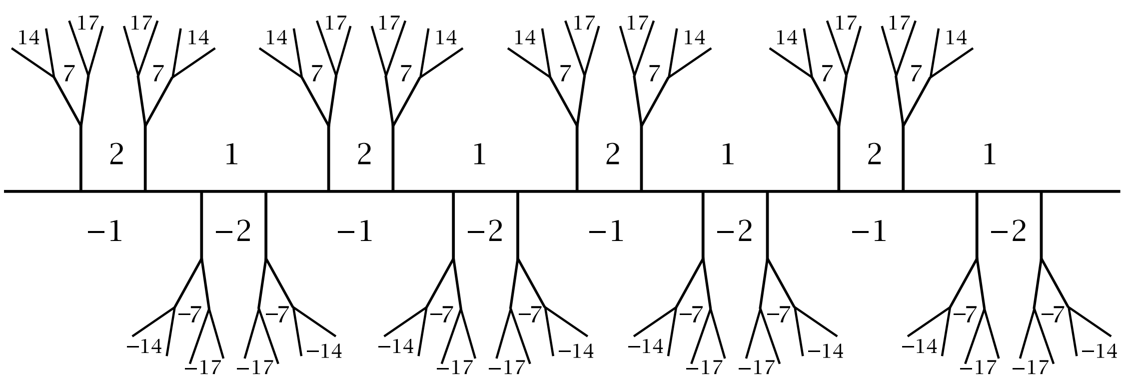
4. For the form $Q(x,y) = 2x^2 - xy + 3y^2$ do the following:
(a) Draw the topograph, showing all the values $Q(x,y) \le 30$ in the topograph, and including the labels x/y.
(b) List all the values $Q(x,y) \le 30$ in order, including the values when the pair (x,y) is not primitive.
(c) Find all the integer solutions of $Q(x,y) = 24$, both primitive and nonprimitive. (And do not forget that quadratic forms always satisfy $Q(x,y) = Q(-x,-y)$.)

5. Find the quadratic form $Q(x,y)$ for which $Q(3,5) = Q(4,7) = Q(7,12) = 1$ by first drawing a strip in the Farey diagram containing the triangles $\langle 1/0, 0/1, 1/1 \rangle$ and $\langle 3/5, 4/7, 7/12 \rangle$ (this can be done using the continued fraction for $7/12$), then adding the edges of the dual tree that meet these triangles, then filling in values of the topograph starting with the given values.

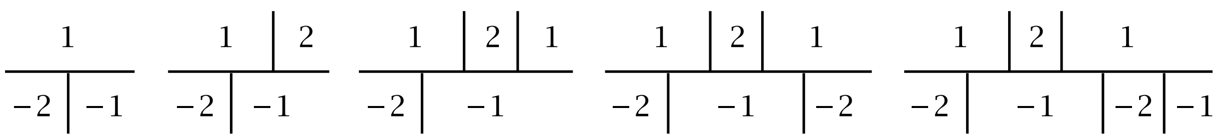
4.2 Periodicity

For most quadratic forms that take on both positive and negative values, such as $x^2 - 2y^2$, there is another way of drawing the topograph that reveals some hidden and unexpected properties. Looking back at the topograph we drew for $x^2 - 2y^2$ we see a zigzag path of edges separating the positive and negative values, and if we straighten this path out to be a line, called the ***separator line***, what we see is the following infinitely repeated pattern:

$Q(x,y) = x^2 - 2y^2$

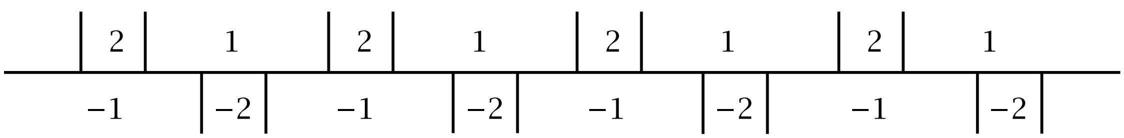


To construct this, one can first build the separator line starting with the three values $Q(1,0) = 1$, $Q(0,1) = -2$, and $Q(1,1) = -1$. Place these as shown in part (a) of the figure below, with a horizontal line segment separating the positive from the negative values.



(a) (b) (c) (d) (e)

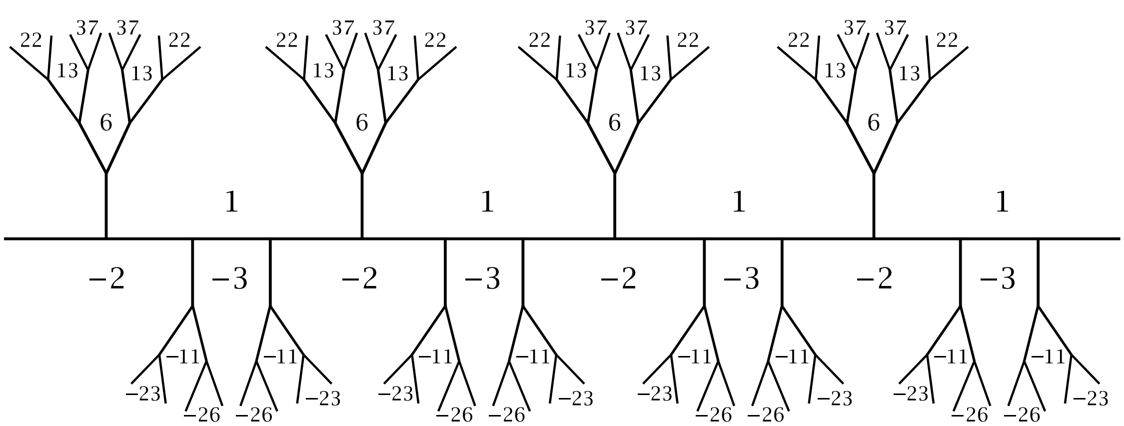
To extend the separator line one step farther to the right, apply the arithmetic progression rule to compute the next value 2 using the arithmetic progression $-2, 1-1, 2$. Since this value 2 is positive, we place it above the horizontal line and insert a vertical edge to separate this 2 from the 1 to the left of it, as in (b) of the figure. Now we repeat the process with the next arithmetic progression $1, 2-1, 1$ and put the new 1 above the horizontal line with a vertical edge separating it from the preceding 2, as shown in (c). At the next step we compute the next value -2 and place it below the horizontal line since it is negative, giving (d). One more step produces (e) where we see that further repetitions will produce a pattern that repeats periodically as we move to the right. The arithmetic progression rule also implies that it repeats periodically to the left, so it is periodic in both directions:



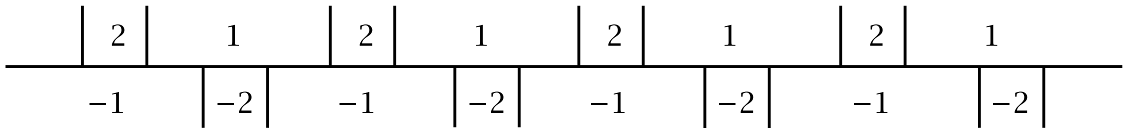
Thus we have the periodic separator line. To get the rest of the topograph we can then work our way upward and downward from the separator line, as shown in the original figure. As one moves upward from the separator line, the values of Q become larger and larger, approaching $+\infty$ monotonically, and as one moves downward, the values approach $-\infty$ monotonically. The reason for this will become clear in Section 5.1 when we discuss something called the Monotonicity Property.

An interesting property of this form x^2-2y^2 that is evident from its topograph is that its negative values are exactly the negatives of its positive values. This would have been hard to predict from the formula x^2-2y^2. Indeed, the similar-looking form x^2-3y^2 no longer has this sign symmetry property, as one can see in its straightened-out topograph:

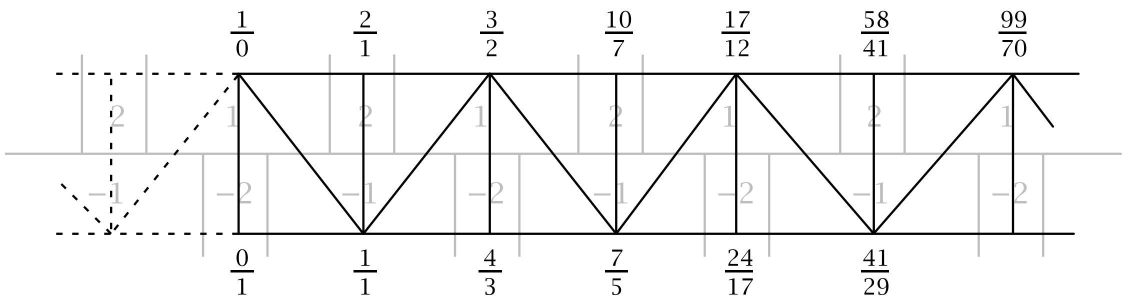
$Q(x,y) = x^2 - 3y^2$



There is a close connection between the separator line in the topograph of a quadratic form x^2-dy^2 and the infinite continued fraction for $\sqrt{d}$ when d is a positive integer that is not a square. In fact, we will see that the topograph can be used to compute the continued fraction for $\sqrt{d}$. As an example let us look at the case $d = 2$. The relevant portion of the topograph for $x^2 - 2y^2$ is the strip along the line separating the positive and negative values:

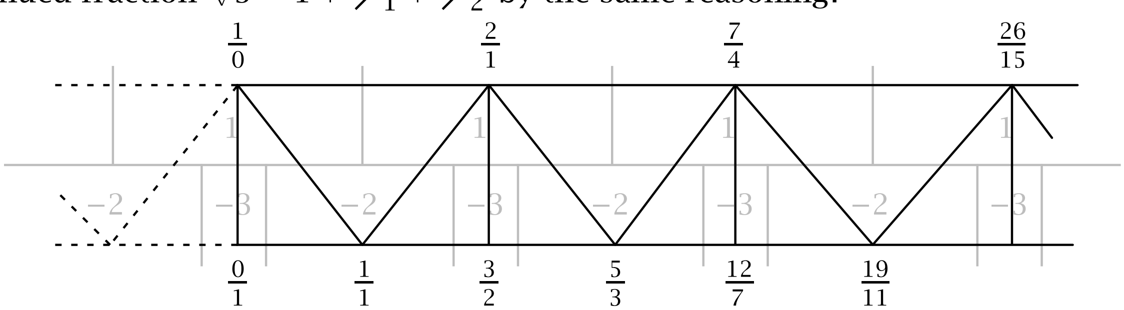


This is a part of the dual tree of the Farey diagram. If we superimpose the triangles of the Farey diagram corresponding to this part of the dual tree, we obtain an infinite strip of triangles:

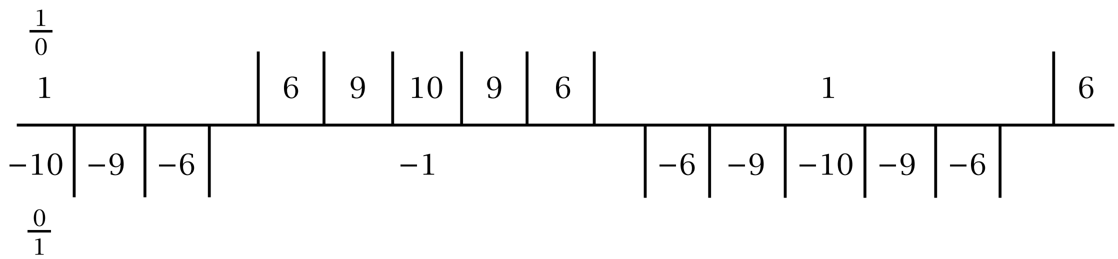


Ignoring the dotted triangles to the left, the infinite strip of triangles corresponds to the infinite continued fraction $1 + \overline{1/2}$. We saw how to compute the value of this continued fraction in Chapter 2, but there is an easier way using the quadratic form $x^2 - 2y^2$. For fractions x/y labeling the vertices along the infinite strip, the corresponding values $n = x^2 - 2y^2$ are either ± 1 or ± 2. We can rewrite the equation $x^2 - 2y^2 = n$ as $(x/y)^2 = 2 + n/y^2$. As we go farther and farther to the right in the infinite strip, both x and y are getting larger and larger while n only varies through finitely many values, namely ± 1 and ± 2, so the quantity n/y^2 is approaching 0. The equation $(x/y)^2 = 2 + n/y^2$ then implies that $(x/y)^2$ is approaching 2, so x/y is approaching $\sqrt{2}$. Since these fractions x/y are the convergents for the infinite continued fraction $1 + \overline{1/2}$ that corresponds to the infinite strip, this implies that the value of the continued fraction $1 + \overline{1/2}$ is $\sqrt{2}$.

As another example, the quadratic form $x^2 - 3y^2$ can be used to compute the continued fraction $\sqrt{3} = 1 + \overline{1/1 + 1/2}$ by the same reasoning:

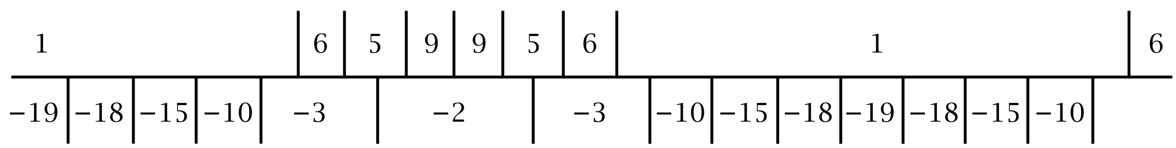


One can see in these two examples that it is not really necessary to draw the strip of triangles, and one can just read off the continued fraction directly from the periodic separator line. Let us illustrate this by considering the separator line for the form $x^2 - 10y^2$ shown below:



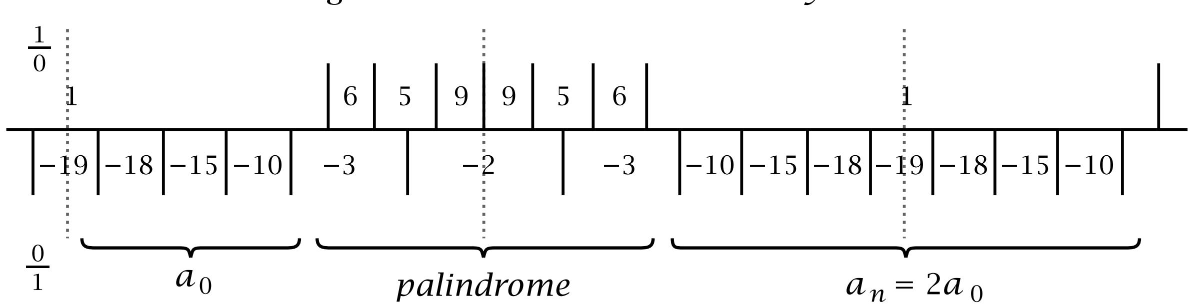
If one moves toward the right along the separator line starting at a point in the edge separating the $1/0$ region from the $0/1$ region, one first encounters three edges leading off to the right (downward), then six edges leading off to the left (upward), then six edges leading off to the right, and thereafter six edges leading off to the left and right alternately. This means that the continued fraction for $\sqrt{10}$ is $3 + \overline{1/6}$.

Here is a more complicated example showing how to compute the continued fraction for $\sqrt{19}$ from the form $x^2 - 19y^2$:



From this we read off that $\sqrt{19} = 4 + \overline{1/2 + 1/1 + 1/3 + 1/1 + 1/2 + 1/8}$.

In Section 5.1 we will prove that the topographs of forms $x^2 - dy^2$ always have a periodic separator line when d is a positive integer that is not a square. As in the examples above, this separator line always includes the edge of the topograph separating the $1/0$ and $0/1$ regions since the form takes the positive value $+1$ at $1/0$ and the negative value $-d$ at $0/1$. In addition to being periodic, the separator line also has mirror symmetry with respect to reflection across the vertical line through the $1/0$ and $0/1$ regions. This is because the form $x^2 - dy^2$ has no xy term, so replacing x/y by $-x/y$ does not change the value of the form. Replacing x/y by $-x/y$ reflects the circular Farey diagram across the horizontal edge from $1/0$ to $0/1$, and this reflects the periodic separator line across the vertical line through the $1/0$ and $0/1$ regions. Once the separator line has symmetry with respect to this vertical line, the periodicity forces it to have mirror symmetry with respect to an infinite sequence of vertical lines, the dotted lines in the figure below for the form $x^2 - 19y^2$.



The reflection lines are the translates of the initial symmetry line L by all the powers

T^n of the periodicity transformation T, along with all the lines halfway between these lines $T^n(L)$. These midlines are lines of mirror symmetry since each individual period has mirror symmetry, as reflection across L takes the left half of the period between L and $T(L)$ to the right half of the period between L and $T^{-1}(L)$.

Because of all these mirror symmetries along the separator line for $x^2 - dy^2$ it follows that the continued fraction for $\sqrt{d}$ has the form

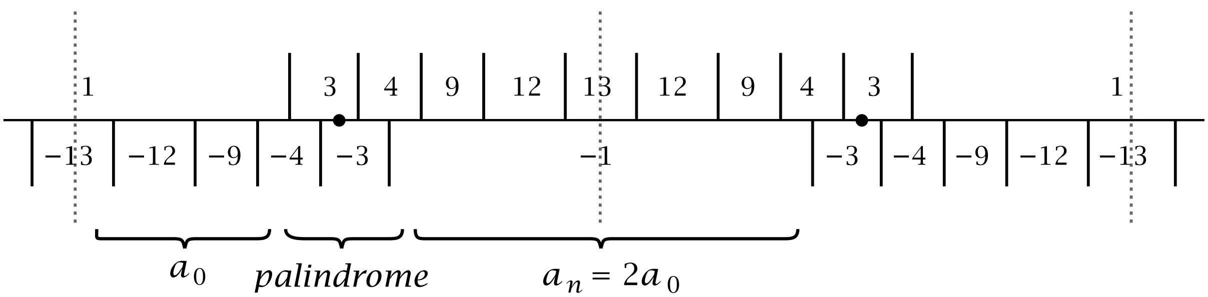
$$\sqrt{d} = a_0 + \overline{{}^1\!\nearrow_{a_1} + {}^1\!\nearrow_{a_2} + \cdots + {}^1\!\nearrow_{a_n}}$$

with two further special properties:

- $a_n = 2a_0$.
- The intermediate terms $a_1, a_2, \cdots, a_{n-1}$ form a palindrome, reading the same forward as backward.

Thus in $\sqrt{19} = 4 + \overline{{}^1\!\nearrow_2 + {}^1\!\nearrow_1 + {}^1\!\nearrow_3 + {}^1\!\nearrow_1 + {}^1\!\nearrow_2 + {}^1\!\nearrow_8}$ the final 8 is twice the initial 4, and the intermediate terms $2, 1, 3, 1, 2$ form a palindrome. These special properties held also in the earlier examples, but were less apparent because there were fewer terms in the repeated part of the continued fraction.

In some cases there is an additional kind of symmetry along the separator line, as illustrated for the form $x^2 - 13y^2$:



As before there is a horizontal translation giving the periodicity and there are mirror symmetries across vertical lines, but now there is an extra glide reflection along the strip that interchanges the positive and negative values of the form. Performing this glide reflection twice in succession gives the translational periodicity. There are also 180 degree rotational symmetries about the points marked with dots on the separator line, and these rotations account for the palindromic middle part of the continued fraction:

$$\sqrt{13} = 3 + \overline{{}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_6}$$

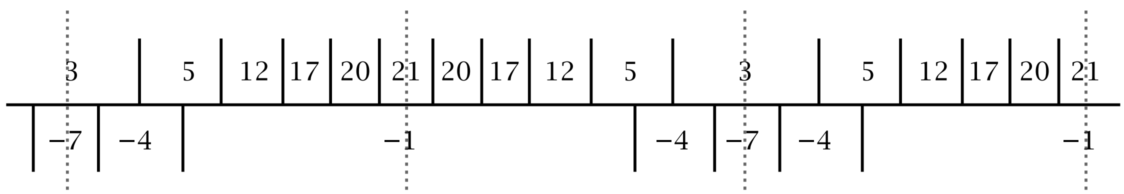
The fact that the periodic part has odd length corresponds to the separator line having the glide reflection symmetry. We could rewrite the continued fraction to have a periodic part of even length by doubling the period:

$$\sqrt{13} = 3 + \overline{{}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_6 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_1 + {}^1\!\nearrow_6}$$

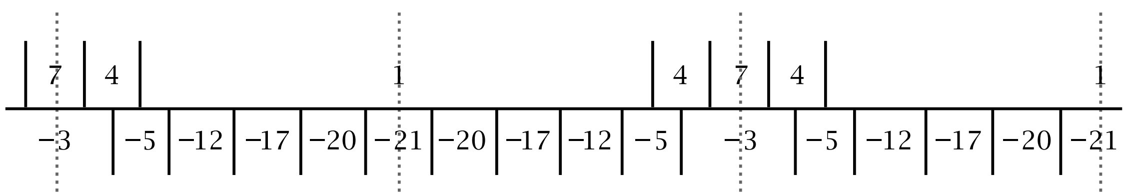
This corresponds to ignoring the glide reflection and just considering the translational periodicity.

We have been using quadratic forms $x^2 - dy^2$ to compute the continued fractions for irrational numbers $\sqrt{d}$, but everything works just the same for irrational numbers $\sqrt{p/q}$ using the quadratic form $qx^2 - py^2$ in place of $x^2 - dy^2$. Following the same reasoning as before, if the equation $qx^2 - py^2 = n$ is rewritten as $q(x/y)^2 = p + n/y^2$ then we see that as we move out along the periodic separator line the numbers x and y approach infinity while n cycles through finitely many values, so the term n/y^2 approaches 0 and the fractions x/y approach a number z satisfying $qz^2 = p$, so $z = \sqrt{p/q}$. This argument depends of course on the existence of a periodic separator line, and we will prove in the next chapter that forms $qx^2 - py^2$ always have a periodic separator line if p and q are positive and the roots $\pm\sqrt{p/q}$ of $qz^2 - p = 0$ are irrational.

Here are some examples. For the first one we use the form $3x^2 - 7y^2$ to compute the continued fraction for $\sqrt{7/3}$:

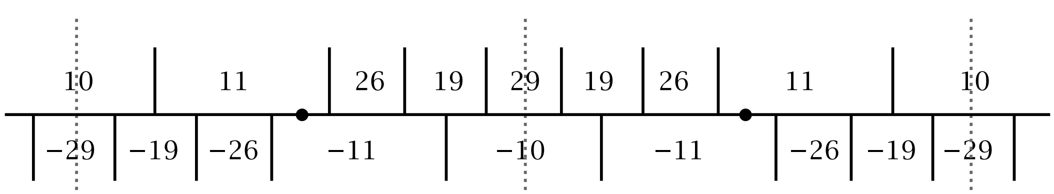


This gives $\sqrt{7/3} = 1 + \overline{1/1 + 1/1 + 1/8 + 1/1 + 1/1 + 1/2}$. For comparison, we can compute the continued fraction for $\sqrt{3/7}$ from the topograph of $7x^2 - 3y^2$:



The separator line here is obtained from the previous one by reflecting across a horizontal axis and changing the sign of the labels. These modifications correspond to changing $3x^2 - 7y^2$ to $3y^2 - 7x^2$ by first interchanging x and y which reflects the Farey diagram and hence also the topograph, and then changing the sign of the resulting form $3y^2 - 7x^2$ to get $7x^2 - 3y^2$. From the separator line for $7x^2 - 3y^2$ we then read off the continued fraction $1/1 + \overline{1/1 + 1/1 + 1/8 + 1/1 + 1/1 + 1/2}$ for $\sqrt{3/7}$. This is the reciprocal of the previous continued fraction since $\sqrt{3/7}$ is the reciprocal of $\sqrt{7/3}$.

For the next example we use $10x^2 - 29y^2$ to compute the continued fraction for $\sqrt{29/10}$ from the separator line:



This gives $\sqrt{29/10} = 1 + \overline{{}^1\!\nearrow_1 + {}^1\!\nearrow_2 + {}^1\!\nearrow_2 + {}^1\!\nearrow_1 + {}^1\!\nearrow_2}$. The period of odd length here corresponds to the existence of the glide reflection and 180 degree rotation symmetries.

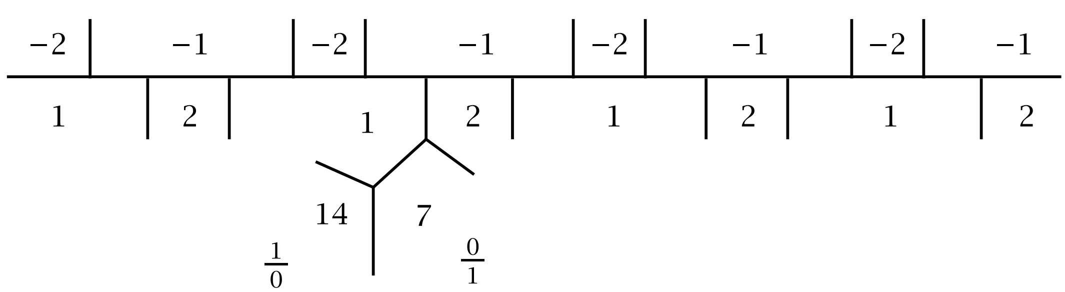
As we see in these examples there are two cases:

$$\sqrt{p/q} = a_0 + \overline{{}^1\!\nearrow_{a_1} + {}^1\!\nearrow_{a_2} + \cdots + {}^1\!\nearrow_{a_n}} \quad \text{if } {}^p\!/_q > 1$$
$$\sqrt{p/q} = {}^1\!\nearrow_{a_0} + \overline{{}^1\!\nearrow_{a_1} + {}^1\!\nearrow_{a_2} + \cdots + {}^1\!\nearrow_{a_n}} \quad \text{if } {}^p\!/_q < 1$$

The palindrome property and the relation $a_n = 2a_0$ that we observed in the continued fraction for $\sqrt{d}$ still hold for irrational numbers $\sqrt{p/q}$. The key point is that the form $qx^2 - py^2$ is unchanged when the sign of x is changed, so its topograph has mirror symmetry with respect to reflection across a line through the ${}^1\!/_0$ and ${}^0\!/_1$ regions, and this symmetry implies the special properties of the continued fraction.

One might ask whether the irrational numbers $\sqrt{p/q}$ are the only numbers having a continued fraction $a_0 + \overline{{}^1\!\nearrow_{a_1} + \cdots + {}^1\!\nearrow_{a_n}}$ or ${}^1\!\nearrow_{a_0} + \overline{{}^1\!\nearrow_{a_1} + \cdots + {}^1\!\nearrow_{a_n}}$ satisfying the palindrome property and the relation $a_n = 2a_0$. Here we should restrict attention only to positive irrational numbers since the numbers $a_0, a_1, \cdots, a_n$ must all be positive. The answer is Yes, as we will see later in this section.

More generally, quadratic forms can be used to compute the continued fractions for all quadratic irrationals. To illustrate the general method let us find the continued fraction for $(10 + \sqrt{2})/14$ which is a root of the equation $14z^2 - 20z + 7 = 0$. The associated quadratic form is $14x^2 - 20xy + 7y^2$, obtained by setting $z = {}^x\!/_y$ and then multiplying by y^2. We would like to find a periodic separator line in the topograph of this form. To do this we start with the three values at ${}^1\!/_0$, ${}^0\!/_1$, and ${}^1\!/_1$, which are the positive numbers 14, 7, and 1, and we then use the arithmetic progression rule to move in a direction that leads to negative values since the separator line separates positive and negative values of the form. In this way we are led to a separator line which is indeed periodic:

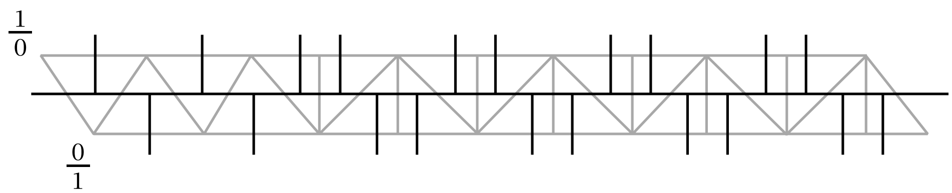


This figure lies in the upper half of the circular Farey diagram where the fractions ${}^x\!/_y$ labeling the regions in the topograph are positive. If we follow the separator line out to either end, the labels ${}^x\!/_y$ have both x and y increasing monotonically and approaching infinity, as a consequence of the mediant rule for labeling vertices of the Farey diagram. Hence the values

$$14z^2 - 20z + 7 = 14({}^x\!/_y)^2 - 20({}^x\!/_y) + 7 = (14x^2 - 20xy + 7y^2)/y^2$$

are approaching zero since the values of the numerator $14x^2 - 20xy + 7y^2$ on the right just cycle through a finite set of numbers repeatedly, the values of the form along the separator line, while the denominators y^2 approach infinity. Thus the labels ${}^x\!/_y$ are approaching the roots of the equation $14z^2 - 20z + 7 = 0$. Since we are in the upper half of the Farey diagram, the smaller of the two roots, which is $(10 - \sqrt{2})/14$, is the limit toward the right along the separator line and the larger root $(10 + \sqrt{2})/14$ is the limit toward the left.

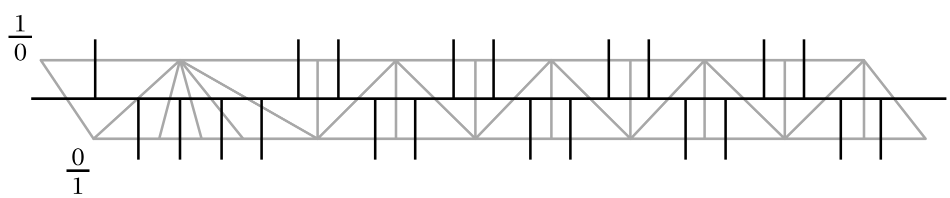
To get the continued fraction for the smaller root, we follow the path in the dual tree of the topograph that starts with the edge between ${}^1\!/_0$ and ${}^0\!/_1$, then zigzags up to the separator line, then goes out this line to the right. If we straighten this path out it looks like the following:



The continued fraction is therefore:

$$\frac{10-\sqrt{2}}{14} = {}^1\!\nearrow\!_1 + {}^1\!\nearrow\!_1 + {}^1\!\nearrow\!_1 + {}^1\!\nearrow\!_1 + \overline{{}^1\!\nearrow\!_2}$$

It is not actually necessary to redraw the straightened-out path since in the original form of the topograph we can read off the sequence of left and right “side roads” as we go along the path, the sequence $LRLR\overline{LLRR}$ where L denotes a side road to the left and R a side road to the right. This sequence determines the continued fraction. For the other root $(10 + \sqrt{2})/14$ the straightened-out path has the following shape:



The sequence of side roads is $LRRRR\overline{LLRR}$ so the continued fraction is

$$\frac{10+\sqrt{2}}{14} = {}^1\!\nearrow\!_1 + {}^1\!\nearrow\!_4 + \overline{{}^1\!\nearrow\!_2}$$

In this example the periodic parts of the continued fractions for both roots are the same, but in general the periodic part for one root is the reverse of the periodic part for the other root since one is moving along the separator line in opposite directions to get to the two roots.

We will show that the procedure in the preceding example works for all quadratic irrational numbers, and this will prove the harder half of Lagrange’s Theorem:

Proposition 4.1. *The continued fraction for every quadratic irrational is eventually periodic.*

The proof will involve associating a quadratic form to each quadratic irrational, and we will need to use the fact that the quadratic forms arising in this way all have periodic separator lines. This will be proved in the next chapter, so the proof will not be officially complete until then.

Proof: Quadratic irrationals are the numbers $\alpha = A + B\sqrt{n}$ for which A and B are rational, B is nonzero, and n is a positive integer that is not a square. The first step in the proof will be to find a quadratic equation with integer coefficients having α as a root. From the quadratic formula we know the other root will have to be the conjugate $\overline{\alpha} = A - B\sqrt{n}$, with $\overline{\alpha} \neq \alpha$ since $B \neq 0$. A quadratic equation having α and $\overline{\alpha}$ as roots is then $(z-\alpha)(z-\overline{\alpha}) = 0$. Multiplied out, this becomes $z^2-(\alpha+\overline{\alpha})z+\alpha\overline{\alpha} = z^2 - 2Az + (A^2 - B^2n) = 0$ which has rational coefficients since A and B are rational. After multiplying by a common denominator for the coefficients, this becomes an equation $az^2 + bz + c = 0$ with integer coefficients having α and $\overline{\alpha}$ as roots. Here $a > 0$ since it is the common denominator we multiplied by.

The polynomial az^2+bz+c determines a quadratic form $ax^2+bxy+cy^2$. This form has two special properties:

- Its topograph contains both positive and negative values. This is because the polynomial $az^2+bz+c = a(z-\alpha)(z-\overline{\alpha})$ takes negative values when z is between the two roots α and $\overline{\alpha}$, where the two factors in parentheses have opposite sign, and positive values when z is greater than both roots or less than both roots, so the two parenthetical factors have the same sign. Thus there are rational numbers $z = {}^x/_y$ where the left side of the equation

$$a({}^x/_y)^2 + b({}^x/_y) + c = (ax^2 + bxy + cy^2)/y^2$$

  has both signs, hence the same is true for the numerator on the right.
- The topograph does not contain the value 0. To see why, suppose there is a pair $(x,y) \neq (0,0)$ with $ax^2 + bxy + cy^2 = 0$. We cannot have $y = 0$, otherwise x would also be 0 since $a \neq 0$. Then since $y \neq 0$, the displayed equation above would say that ${}^x/_y$ was a rational root of $az^2+bz+c = 0$, contradicting the fact that its roots α and $\overline{\alpha}$ are irrational.

We will show in Theorem 5.2 that every form $ax^2 + bxy + cy^2$ with these two properties has a periodic separator line in its topograph. This corresponds to an infinite periodic strip in the Farey diagram.

Lemma 4.2. *The ends of the periodic strip in the topograph of a hyperbolic form $ax^2 + bxy + cy^2$ are at the roots α and $\overline{\alpha}$ of the equation $az^2 + bz + c = 0$.*

Proof: Consider the labels ${}^x/_y$ on the vertices along the strip. Since the denominators y approach infinity as we go out to either end of the strip while the values of the form $ax^2 + bxy + cy^2$ cycle through finitely many values, it follows that the values of the right side of the equation

$$a({}^x\!/\!_y)^2 + b({}^x\!/\!_y) + c = (ax^2 + bxy + cy^2)/y^2$$

are approaching zero. This means that the vertex labels ${}^x\!/\!_y$ are approaching a root of the equation $az^2 + bz + c = 0$. In our discussion of infinite strips in Section 2.2 we saw that the two ends of any infinite strip are at two different irrational numbers, so the two ends of the periodic strip for the form $ax^2 + bxy + cy^2$ are at the two roots α and $\overline{\alpha}$ of the equation $az^2 + bz + c = 0$. □

With this lemma we can finish the proof of Proposition 4.1 by comparing two infinite strips with an end at the root α of the equation $az^2 + bz + c = 0$. One infinite strip is the strip given by the continued fraction for α. This strip consists of all the triangles in the upper halfplane Farey diagram that meet the vertical line through α. This strip starts at the vertex ${}^1\!/\!_0$ at the top and then moves downward through an infinite sequence of triangles approaching α. The other infinite strip is the one corresponding to the separator line for the form $ax^2 + bxy + cy^2$, which has an end at α by the lemma. The ends of both strips at α eventually coincide since the analysis of infinite strips in Section 2.2 showed that the ends of every infinite strip eventually consist of the triangles meeting the vertical lines through the irrational numbers at the ends of the strip. Thus the continued fraction for α is eventually periodic since the periodic strip for $ax^2 + bxy + cy^2$ is periodic. □

We are now able to answer a question raised earlier in this section:

Proposition 4.3. *The numbers $\sqrt{p/q}$ are the only quadratic irrationals having continued fractions $a_0 + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$ or ${}^1\!/\!_{a_0} + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$ satisfying the palindrome property and the relation $a_n = 2a_0$.*

Proof: Consider first a continued fraction $a_0 + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$ satisfying the palindrome property and the relation $a_n = 2a_0$. The initial a_0 in this continued fraction must be positive since it is half of the positive number a_n. The reciprocal of the continued fraction $a_0 + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$ is ${}^1\!/\!_{a_0} + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$ and the reciprocal of $\sqrt{p/q}$ is $\sqrt{q/p}$ so it will suffice to prove the proposition just for continued fractions of the type ${}^1\!/\!_{a_0} + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$.

Let ${}^1\!/\!_{a_0} + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$ be a continued fraction satisfying the palindrome condition and the relation $a_n = 2a_0$. We may assume n, the length of the period, is even since doubling the period gives the continued fraction

$${}^1\!/\!_{a_0} + \overline{{}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n} + {}^1\!/\!_{a_1} + \cdots + {}^1\!/\!_{a_n}}$$

which again satisfies the palindrome condition and the "$a_n = 2a_0$" condition, where the new palindrome is $a_1 \cdots a_{n-1} a_n a_1 \cdots a_{n-1}$ which is a palindrome if $a_1 \cdots a_{n-1}$ is a palindrome.

The strip in the upper half of the circular Farey diagram corresponding to the continued fraction $\overline{1/a_0 + 1/a_1 + \cdots + 1/a_n}$ starts at the $\langle 1/0, 0/1 \rangle$ edge and converges to the value α of the continued fraction at the other end of the strip. Combining the strip with its reflection across the $\langle 1/0, 0/1 \rangle$ edge gives an infinite strip with mirror symmetry across the $\langle 1/0, 0/1 \rangle$ edge. This doubled strip is periodic along its entire length by the palindrome condition and the condition $a_n = 2a_0$. The other end of the doubled strip converges to $\overline{\alpha}$ since we have seen that the two endpoints of a periodic strip satisfy a single quadratic equation $T(z) = z$ where T is the periodicity transformation. The two roots α and $\overline{\alpha}$ of this equation are conjugates and they are also negatives of each other by the mirror symmetry across the edge $\langle 1/0, 0/1 \rangle$, so we have $\overline{\alpha} = -\alpha$. Writing α as $A + B\sqrt{m}$ with A and B rational, the equation $\overline{\alpha} = -\alpha$ becomes $A - B\sqrt{m} = -A - B\sqrt{m}$ which implies that $A = 0$. Since α is positive we then have $\alpha = B\sqrt{m}$ with $B > 0$. Thus α is the square root of the positive rational number $B^2 m$. $\square$

Another natural question one might ask is whether every periodic line in the dual tree of the Farey is realizable as the separator line of some form. A trivial sort of periodic line which cannot be realized is an infinite line in which all the abutting edges lie on one side of the line. This is dual to an infinite fan in the Farey diagram consisting of all the triangles containing a given vertex. When we say "periodic line" we will implicitly exclude trivial lines like this.

Proposition 4.4. *Every periodic line in the dual tree of the Farey diagram occurs as the separator line for some form.*

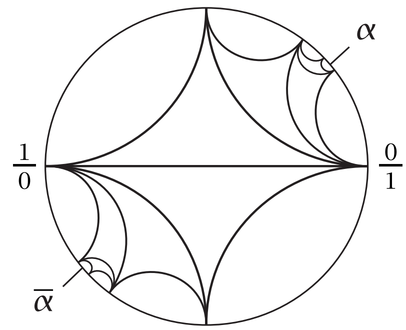
Proof: Given a periodic line, the periodicity of this line and of the corresponding infinite strip is realized by some linear fractional transformation T. As we have seen, the endpoints of the strip are the fixed points of T, the solutions of $T(z) = z$. This can be rewritten as a quadratic equation $az^2 + bz + c = 0$ with integer coefficients. The coefficient a must be nonzero, otherwise we would have an equation $bz + c = 0$ with only one root if $b \neq 0$, while if $b = 0$ the equation would have no roots if $c \neq 0$. If $c = 0$ as well as $a = 0$ and $b = 0$ the equation would degenerate to $0 = 0$, meaning that every z satisfied $T(z) = z$ so T would be the identity transformation rather than the periodicity transformation, a contradiction. Thus a must be nonzero, and we may assume that $a > 0$ by multiplying the equation by -1 if necessary.

We claim that the the periodic line we started with is a separator line in the topograph of the form $ax^2 + bxy + cy^2$. This just means that the values of the form at vertices along one edge of the associated periodic strip are all positive and the values along the other edge are all negative. To see why this is so let us factor $az^2 + bz + c$ as $a(z-\alpha)(z-\overline{\alpha})$ where α and $\overline{\alpha}$ are the roots of $az^2 + bz + c = 0$ at the ends of the strip. From this factorization and the fact that a is positive we see that the product $a(z-\alpha)(z-\overline{\alpha})$ is negative if z is between α and $\overline{\alpha}$ and positive if z is greater than both α and $\overline{\alpha}$ or less than both α and $\overline{\alpha}$. (We saw this previously in the proof of Proposition 4.1.) Taking z to be a rational number x/y, the equation

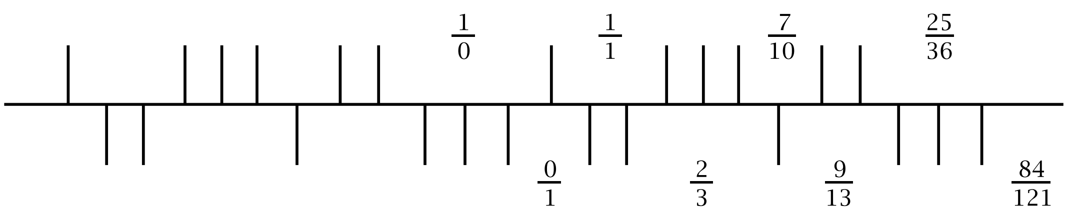
$$a(x/y)^2 + b(x/y) + c = (ax^2 + bxy + cy^2)/y^2$$

implies that the form $ax^2 + bxy + cy^2$ takes negative values for x/y in the interval between α and $\overline{\alpha}$ and positive values for x/y outside this interval, assuming $x/y \neq 1/0$ so we are not dividing by 0 in the equation above.

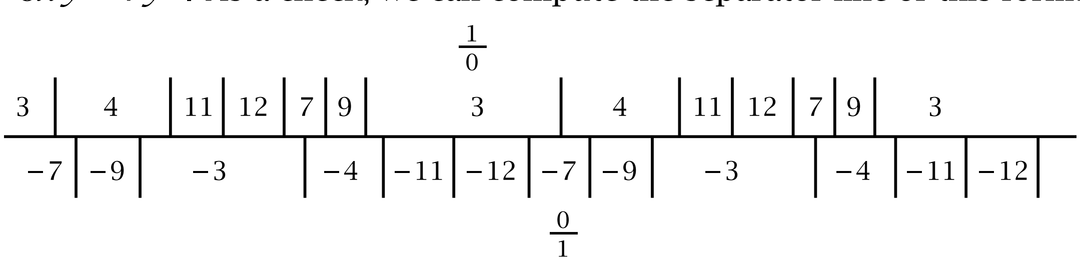
In terms of the circular Farey diagram the roots α and $\overline{\alpha}$ divide the boundary circle into two arcs, with the form taking positive values at vertices in one arc and negative values at vertices in the other arc, with the possible exception of the vertex $1/0$. However, this vertex is not actually exceptional since it lies in the arc with positive values and the form takes the value $a > 0$ when $x/y = 1/0$. This proves what we wanted since vertices along one edge of the strip lie in one arc and vertices along the other edge lie in the other arc. □



To illustrate the procedure in the preceding proof let us find a quadratic form whose periodic separator line is the following:



The fractional labels correspond to vertices of the underlying Farey diagram, and from these we see that the translation giving the periodicity sends $1/0$ to $25/36$ and $0/1$ to $84/121$. The matrix of this transformation is $\begin{pmatrix} 25 & 84 \\ 36 & 121 \end{pmatrix}$ so it is the transformation $T(z) = {}^{25z+84}/_{36z+121}$. The fixed points of T are determined by setting this equal to z. The resulting equation simplifies to $36z^2 + 96z - 84 = 0$ or just $3z^2 + 8z - 7 = 0$. The roots α and $\overline{\alpha}$ of this equation $az^2 + bz + c = 0$ are the fixed points, but we do not actually have to compute them since we showed in the preceding proof that the quadratic form we want is then $ax^2 + bxy + cy^2$ which in this example is just $3x^2 + 8xy - 7y^2$. As a check, we can compute the separator line of this form:



This provides a realization of the given periodic line as the separator line of a hyperbolic form. Any constant multiple of this form would also have the same separator line since we would just be multiplying all the labels along the line by the same constant.

We could have simplified the calculation in this example by observing that the periodic line we started with is taken to itself by a glide reflection that moves the line only half as far along itself as the translation T that we used. This glide reflection is $T'(z) = {}^{2z+7}/_{3z+10}$ and it has the same fixed points as T so we could use the equation $T'(z) = z$ instead of $T(z) = z$. Thus we have ${}^{2z+7}/_{3z+10} = z$ which simplifies more directly to $3z^2 + 8z - 7 = 0$, the same final equation as before.

Exercises

1. Determine the periodic separator line in the topograph for each of the following quadratic forms. (You do not need to include the fractional labels ${}^x/_y$.)
(a) $x^2 - 7y^2$
(b) $3x^2 - 4y^2$
(c) $x^2 + xy - y^2$

2. For the following quadratic forms, draw enough of the topograph, starting with the edge separating the ${}^1/_0$ and ${}^0/_1$ regions, to locate the periodic separator line, and include the separator line itself in your topograph.
(a) $x^2 + 3xy + y^2$
(b) $6x^2 + 18xy + 13y^2$
(c) $37x^2 - 104xy + 73y^2$

3. Using your answers in the first problem above, write down the continued fraction expansions for $\sqrt{7}$, $2\sqrt{3}/3$, and $(-1+\sqrt{5})/2$.

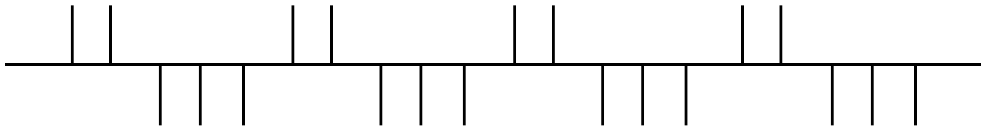
4. Use a quadratic form to compute continued fractions for the following pairs of numbers:
(a) $(3+\sqrt{6})/2$ and $(3-\sqrt{6})/2$
(b) $(11+\sqrt{13})/6$ and $(11-\sqrt{13})/6$
(c) $(14+\sqrt{7})/9$ and $(14-\sqrt{7})/9$

5. Compute the periodic separator line for the form $x^2 - 43y^2$ and use this to find the continued fraction for $\sqrt{43}$.

6. Use the form $x^2 - 2n^2y^2$ to compute the continued fraction for $n\sqrt{2}$ for $n = 1, 2, 3, 4, 5$.

7. Compute the continued fraction for $\sqrt{21}$ using the form $x^2 - 21y^2$. Can you explain the relationship between this continued fraction and the one for $\sqrt{7/3}$ computed in this section?

8. (a) Find a quadratic form whose periodic separator line has the following pattern:

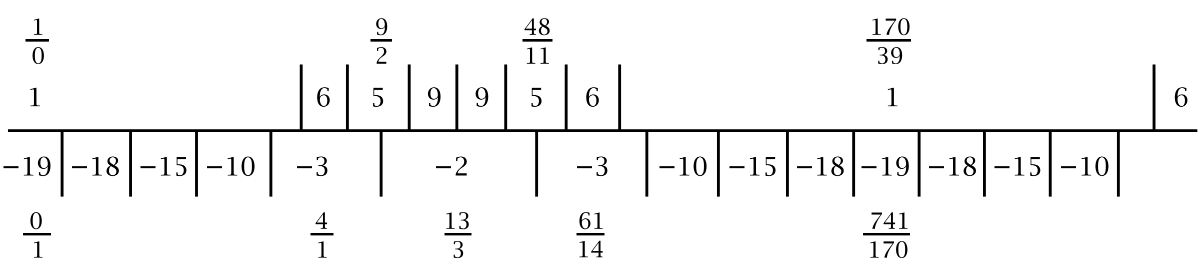


(b) Generalize part (a) by replacing each pair of upward edges with m upward edges and each triple of downward edges with n downward edges.

4.3 Pell's Equation

We encountered the equation $x^2 - dy^2 = 1$ briefly in Chapter 0. It is traditionally called Pell's equation, and the similar equation $x^2 - dy^2 = -1$ is sometimes called Pell's equation as well, or else the negative Pell's equation. If d is a square then the equations are not very interesting since in this case d can be incorporated into the y^2 term, so one is looking at the equations $x^2 - y^2 = 1$ and $x^2 - y^2 = -1$, which have only the trivial solutions $(x,y) = (\pm 1, 0)$ for the first equation and $(x,y) = (0, \pm 1)$ for the second equation since these are the only cases when the difference between two squares is ± 1. We will therefore assume that d is not a square in what follows. It will suffice to find the solutions with x and y positive since the signs of x and y do not affect the value of $x^2 - dy^2$.

As an example let us look at the equation $x^2 - 19y^2 = 1$. We drew a portion of the periodic separator line for the form $x^2 - 19y^2$ earlier, and here it is again with some of the fractional labels x/y shown as well:



Ignoring the label $741/170$ for the moment, the other fractional labels are the first few convergents for the continued fraction for $\sqrt{19}$ that we computed before, which is $4 + \overline{1/2 + 1/1 + 1/3 + 1/1 + 1/2 + 1/8}$. These fractional labels are the labels on the vertices of the zigzag path in the infinite strip of triangles in the Farey diagram, which we can imagine being superimposed on the separator line in the figure. The fractional label we are most interested in is the $170/39$ in the upper right because this is the label on a region where the value of the form $x^2 - 19y^2$ is 1. This means exactly that $(x,y) = (170, 39)$ is a solution of $x^2 - 19y^2 = 1$. In terms of continued fractions, the fraction $170/39$ is the value of the initial portion $4 + 1/2 + 1/1 + 1/3 + 1/1 + 1/2$ of the continued fraction for $\sqrt{19}$, with the final term of the period omitted.

Since the topograph of $x^2 - 19y^2$ is periodic along the separator line, there are infinitely many different solutions of $x^2 - 19y^2 = 1$ along the separator line. Going toward the left just gives the negatives $-x/y$ of the fractions x/y to the right, so since we are only interested in the positive solutions it will suffice to see what happens toward the right. One way to do this is to use the linear fractional transformation that gives the periodicity translation toward the right. This transformation sends the edge $\langle 1/0, 0/1 \rangle$ of the Farey diagram to the edge $\langle 170/39, 741/170 \rangle$. Here $741/170$ is the value of the continued fraction $4 + 1/\!\!\nearrow 2 + 1/\!\!\nearrow 1 + 1/\!\!\nearrow 3 + 1/\!\!\nearrow 1 + 1/\!\!\nearrow 2 + 1/\!\!\nearrow 4$ obtained from the continued fraction for $\sqrt{19}$ by replacing the final number 8 in the period by one-half of its value, 4. The figure above shows why this is the right thing to do. We then get an infinite sequence of larger and larger positive solutions of $x^2 - 19y^2 = 1$ by repeatedly applying the periodicity transformation with matrix $\left(\begin{smallmatrix} 170 & 741 \\ 39 & 170 \end{smallmatrix}\right)$ to go from one solution to the next. For example,

$$\begin{pmatrix} 170 & 741 \\ 39 & 170 \end{pmatrix} \begin{pmatrix} 170 \\ 39 \end{pmatrix} = \begin{pmatrix} 57799 \\ 13260 \end{pmatrix}$$

so the next solution of $x^2 - 19y^2 = 1$ after $(170, 39)$ is $(57799, 13260)$, and we could compute more solutions if we wanted. Obviously they are getting large rather quickly.

The two 170's in the matrix $\left(\begin{smallmatrix} 170 & 741 \\ 39 & 170 \end{smallmatrix}\right)$ can hardly be just a coincidence. Notice also that the entry 741 factors as $19 \cdot 39$ which hardly seems like it should be just a coincidence either. Let us check that these numbers had to occur. In general, for the form $x^2 - dy^2$ let us suppose that we have found the first solution $(x, y) = (p, q)$ after $(1, 0)$ for Pell's equation $x^2 - dy^2 = 1$, so $p^2 - dq^2 = 1$. Then based on the previous example we suspect that the periodicity transformation is:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p & dq \\ q & p \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} px + dqy \\ qx + py \end{pmatrix}$$

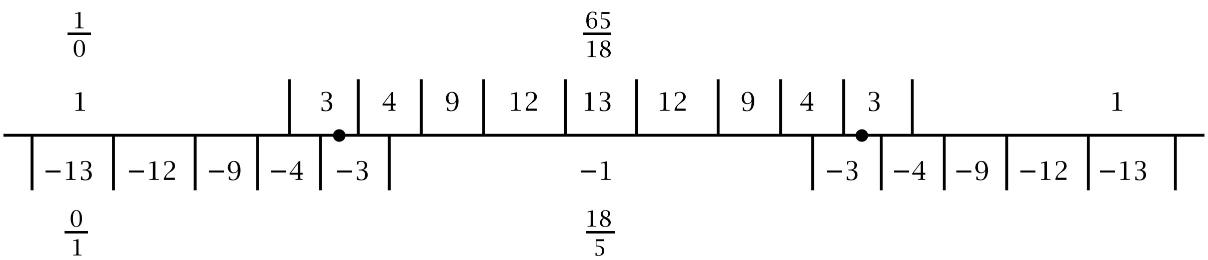
To check that this is correct, the main thing to verify is that T preserves the values of the quadratic form. Substituting $(px + dqy, qx + dy)$ for (x, y) in $x^2 - dy^2$ gives:

$$\begin{aligned} (px + dqy)^2 - d(qx + py)^2 & \\ &= p^2x^2 + 2pdqxy + d^2q^2y^2 - dq^2x^2 - 2pdqxy - dp^2y^2 \\ &= (p^2 - dq^2)x^2 - d(p^2 - dq^2)y^2 \\ &= x^2 - dy^2 \quad \text{since } p^2 - dq^2 = 1 \end{aligned}$$

So T does preserve the values of the form. In particular T takes regions in the topograph with positive values to other such regions, and similarly for regions with negative values, so the separator line is taken to itself. The determinant of $\left(\begin{smallmatrix} p & dq \\ q & p \end{smallmatrix}\right)$ is $p^2 - dq^2 = 1$ which is positive so T preserves orientation and hence it has to be a translation along the separator line. Since we chose (p, q) to be the first solution of $x^2 - dy^2 = 1$ after $(1, 0)$, it follows that T is the periodicity transformation and all occurrences of the label 1 along the separator line are images of the one at $1/0$ under positive or negative powers of T. (We have not actually proved yet that periodic separator lines always exist for forms $x^2 - dy^2$, but this will be shown in Theorem 5.2.)

There are no other solutions of $x^2 - 19y^2 = 1$ besides the ones along the separator line because, as we will see in Section 5.1, the values in a topograph with a separator line change in a monotonic fashion as one moves away from the separator line, steadily increasing toward $+\infty$ on the positive side of the separator line and steadily decreasing toward $-\infty$ on the negative side. Thus the value 1 can occur only along the separator line itself. The monotonicity property also implies that the value -1 never appears in the topograph of $x^2 - 19y^2$ since it does not appear along the separator line, so the negative Pell equation $x^2 - 19y^2 = -1$ has no integer solutions.

For an example where $x^2 - dy^2 = -1$ does have solutions, let us look again at the earlier example of $x^2 - 13y^2$:



The first positive solution $(x, y) = (p, q)$ of $x^2 - 13y^2 = -1$ corresponds to the value -1 in the middle of the figure. This is determined by the continued fraction $p/q = 3 + 1/1 + 1/1 + 1/1 + 1/1 = 18/5$, so we have $(p, q) = (18, 5)$. The matrix $\left(\begin{smallmatrix} p & dq \\ q & p \end{smallmatrix}\right)$ in this case is $\left(\begin{smallmatrix} 18 & 65 \\ 5 & 18 \end{smallmatrix}\right)$ with determinant $18^2 - 13 \cdot 5^2 = -1$ so this gives the glide reflection along the periodic separator line taking $1/0$ to $18/5$ and $0/1$ to $65/18$. The smallest positive solution of $x^2 - 13y^2 = +1$ is obtained by applying this glide reflection to $(18, 5)$, which gives:

$$\begin{pmatrix} 18 & 65 \\ 5 & 18 \end{pmatrix} \begin{pmatrix} 18 \\ 5 \end{pmatrix} = \begin{pmatrix} 324 + 325 \\ 90 + 90 \end{pmatrix} = \begin{pmatrix} 649 \\ 180 \end{pmatrix}$$

Repeated applications of the glide reflection will give solutions of $x^2 - 13y^2 = -1$ and $x^2 - 13y^2 = +1$ alternately.

Exercises

1. For the quadratic form $x^2 - 14y^2$ do the following things:

(a) Draw the separator line in the topograph and compute the continued fraction for $\sqrt{14}$.

(b) Find the smallest positive integer solutions of $x^2 - 14y^2 = 1$ and $x^2 - 14y^2 = -1$, if these equations have integer solutions.

(c) Find the linear fractional transformation that gives the periodicity translation along the separator line and use this to find a second positive solution of $x^2 - 14y^2 = 1$.

(d) Determine the integers n with $|n| \leq 12$ such that the equation $x^2 - 14y^2 = n$ has an integer solution. (Do not forget the possibility that there could be solutions (x, y) that are not primitive.)

2. For the quadratic form $x^2 - 29y^2$ do the following things:

(a) Draw the separator line and compute the continued fraction for $\sqrt{29}$.

(b) Find the smallest positive integer solution of $x^2 - 29y^2 = -1$.

(c) Find a glide reflection symmetry of the separator line and use this to find the smallest positive integer solution of $x^2 - 29y^2 = 1$.

3. Show that every positive integer that is not a square can be expressed as a quotient ${}^{n^2-1}/_{k^2}$ for a suitably chosen pair of integers n and k, and in fact there are infinitely many different choices for such a pair. Why did we exclude squares?

5 The Classification of Quadratic Forms

We can divide quadratic forms $Q(x,y) = ax^2 + bxy + cy^2$ with integer coefficients a, b, c into four broad classes according to the signs of the values $Q(x,y)$, where as usual we restrict x and y to be integers. We will always assume at least one of the coefficients is nonzero, so Q is not identically zero, and we will always assume (x,y) is not $(0,0)$. There are four possibilities:

(I) If $Q(x,y)$ takes on both positive and negative values but not 0 then we call Q a ***hyperbolic*** form.

(II) If $Q(x,y)$ takes on both positive and negative values and also the value 0 then we call Q a **0-*hyperbolic*** form.

(III) If $Q(x,y)$ takes on only positive values or only negative values then we call Q an ***elliptic*** form.

(IV) If Q takes on the value 0 and either positive or negative values, but not both, then Q is called a ***parabolic*** form.

The hyperbolic-elliptic-parabolic terminology is motivated in part by what the level curves $ax^2 + bxy + cy^2 = k$ are when we allow x and y to take on all real values so that one gets actual curves. The level curves are hyperbolas in cases (I) and (II), and ellipses in case (III). In case (IV), however, the level curves are not parabolas as one might guess, but straight lines. From the classical perspective of conic sections parabolas are the transitional case between hyperbolas and ellipses, but from another viewpoint one can pass from hyperbolas to ellipses through a transitional case of a pair of parallel lines as in the family of curves $x^2 - cy^2 = 1$ which are hyperbolas for $c > 0$, ellipses for $c < 0$, and a pair of parallel lines for $c = 0$. Parabolic forms are much simpler than the other types and we will not be spending much time on them.

As we will show later in the chapter, there is an easy way to distinguish the four types of forms $ax^2 + bxy + cy^2$ in terms of their discriminants $\Delta = b^2 - 4ac$:

(I) If Δ is positive but not a square then Q is hyperbolic.

(II) If Δ is positive and a square then Q is 0-hyperbolic.

(III) If Δ is negative then Q is elliptic.

(IV) If Δ is zero then Q is parabolic.

Discriminants play a central role in the theory of quadratic forms. A natural question to ask is whether every integer occurs as the discriminant of some form, and this is easy to answer. For a form $ax^2 + bxy + cy^2$ we have $\Delta = b^2 - 4ac$, and this is congruent to b^2 mod 4. A square such as b^2 is always congruent to 0 or 1 mod 4, so the discriminant of a form is always congruent to 0 or 1 mod 4. Conversely, for every integer Δ congruent to 0 or 1 mod 4 there exists a form whose discriminant is Δ. The simplest ones are:

$$x^2 - ky^2 \text{ with discriminant } \Delta = 4k$$

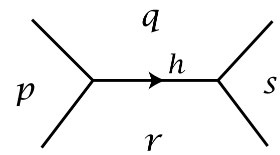
$$x^2 + xy - ky^2 \text{ with discriminant } \Delta = 4k + 1$$

Here k can be positive, negative, or zero. The forms $x^2 - ky^2$ and $x^2 + xy - ky^2$ are called the ***principal*** quadratic forms of these discriminants.

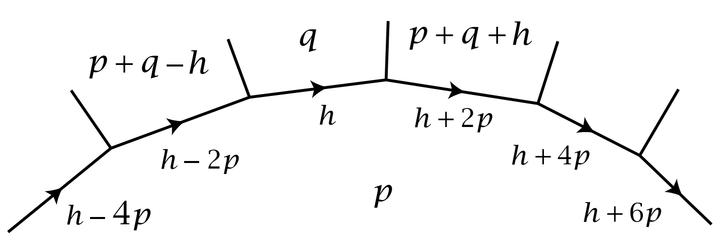
5.1 The Four Types of Forms

We will analyze each of the four types of forms in turn, but before doing this let us make a few preliminary observations that apply to all forms.

In the arithmetic progression rule controlling the labeling of the four regions surrounding an edge of the topograph, we can label the edge by the common increment $h = (q+r) - p = s - (q+r)$ as in the figure at the right. The edge can be oriented by an arrow showing the direction in which the progression increases by h. Changing the sign of h corresponds to changing the orientation of the edge. In the special case that h happens to be 0 the orientation of the edge is irrelevant and can be omitted.



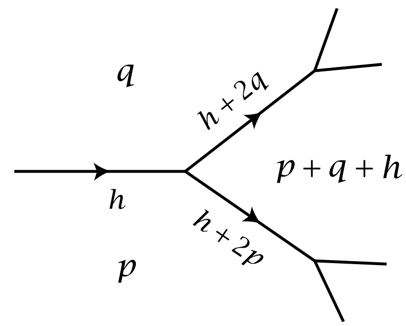
The values of the increment h along the boundary of a region in the topograph have the interesting property that they also form an arithmetic progression when all these edges are oriented in the same direction, and the amount by which h increases as we move from one edge to the next is $2p$ where p is the label on the region adjacent to all these edges:



We will call this property the ***second arithmetic progression rule.*** To see why it holds, start with the edge labeled h in the figure, with the adjacent regions labeled p and q. The original arithmetic progression rule then gives the value $p + q + h$ in the next region to the right. From this we can deduce that the label on the edge between the regions labeled p and $p+q+h$ must be $h+2p$ since this is the increment from q to

$p + (p + q + h)$. Thus the edge label increases by $2p$ when we move from one edge to the next edge to the right, so by repeated applications of this fact we see that we have an arithmetic progression of edge labels all along the border of the region labeled p.

Another thing worth noting at this point is something that we will refer to as the ***monotonicity property***. This says that in the figure at the right, if the three labels p, q, and h adjacent to an edge are all positive, then so are the three labels for the next two edges in front of this edge, and the new labels are larger than the old labels. It follows that when one continues forward going out this part of the topograph, all the labels become monotonically larger the farther one goes. Similarly, when the original three labels are negative, all the labels become larger and larger negative numbers.

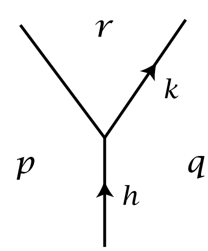


Next we have a very useful way to compute the discriminant of a form directly from its topograph:

Proposition 5.1. *If an edge in the topograph of a form $Q(x,y)$ is labeled h with adjacent regions labeled p and q, then the discriminant of $Q(x,y)$ is $h^2 - 4pq$.*

Note that the sign of h and the orientation of the edge are irrelevant here. The proposition implies that if the discriminant is known then any two of p, q, and $|h|$ determine the third.

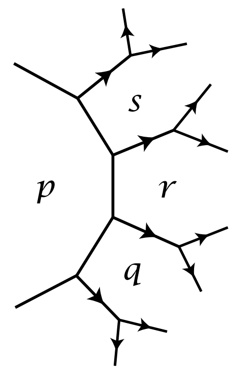
Proof: For the given form $Q(x,y) = ax^2 + bxy + cy^2$, the $1/0$ and $0/1$ regions in the topograph are labeled a and c, and the edge in the topograph separating these two regions has $h = b$ since the $1/1$ region is labeled $a + b + c$. So the statement of the proposition is correct for this edge. For other edges we proceed by induction, moving farther and farther out the tree. For the induction step suppose we have two adjacent edges labeled h and k as in the figure, and suppose inductively that the discriminant equals $h^2 - 4pq$. We have $r = p + q + h$, and from the second arithmetic progression rule we know that $k = h + 2q$. Then we have $k^2 - 4qr = (h+2q)^2 - 4q(p+q+h) = h^2 + 4hq + 4q^2 - 4pq - 4q^2 - 4hq = h^2 - 4pq$, which means that the result holds for the edge labeled k as well. □



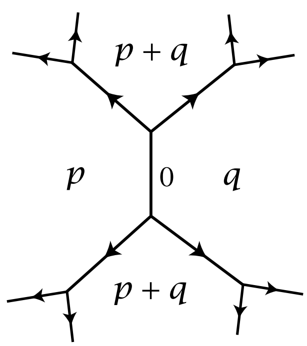
Elliptic Forms

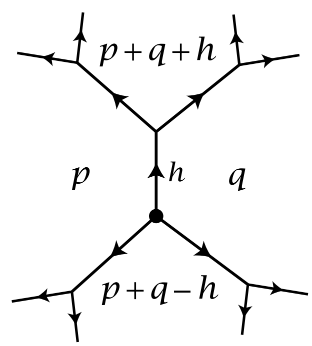
Elliptic forms have fairly simple qualitative behavior, so let us look at these forms first. Recall that we defined a form $Q(x,y)$ to be elliptic if it takes on only positive or only negative values at all integer pairs $(x,y) \neq (0,0)$. The positive and negative cases are equivalent since one can switch from one to the other just by putting a minus sign in front of Q. Thus it suffices to consider the case that Q takes on only positive values, and we will always assume we are in this case whenever we are dealing with elliptic forms. We will also generally assume when we look at topographs of elliptic forms that the orientations of the edges are chosen so as to give positive h-values, unless we state otherwise.

For a positive elliptic form Q let p be the minimum positive value taken on by Q, so $Q(x,y) = p$ for some $(x,y) \neq (0,0)$. Here (x,y) must be a primitive pair otherwise Q would take on a smaller positive value than p. Thus there is a region in the topograph of Q with the label p. All the edges having one endpoint at this region must be oriented away from the region, by the arithmetic progression rule and the assumption that p is the minimum value of Q. The monotonicity property then implies that all edges farther away from the p region are also oriented away from the region, and the values of Q increase steadily as one moves away from the p region.

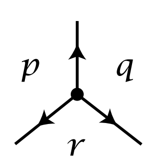


For the edges making up the border of the p region we know that the h-labels on these edges form an arithmetic progression with increment $2p$, provided that we temporarily re-orient these edges so that they all point in the same direction. If some edge bordering the p region has the label $h = 0$ then the topograph has the form shown in the first figure below, with the orientations on edges that give positive h-labels. An example of such a form is $px^2 + qy^2$. We call the 0-labeled edge a ***source edge*** since all other edges are oriented away from this edge.





The other possibility is that no edge bordering the p region has label $h = 0$. Then since the labels on these edges form an arithmetic progression, there must be some vertex where the terms in the progression change sign. Thus when we orient the edges to give positive h-labels, all three edges meeting at this vertex will be oriented away from the vertex, as in the second figure above. We call this a ***source vertex*** since all edges in the topograph are oriented away from this vertex.



If the three regions surrounding a source vertex are labeled p, q, r then the fact that the three edges leading from this vertex all point away from the vertex is equivalent to the three inequalities $p < q + r$, $q < p + r$, and $r < p + q$. These are called triangle inequalities since they are satisfied by the lengths of the three sides of any triangle. In the case of a source edge one of the inequalities becomes an equality, for example $r = p + q$ in the earlier figure with a source edge.

As we know, any three integers p, q, r can be realized as the three labels surrounding a vertex in the topograph of some form. If these are positive integers satisfying the triangle inequalities then this vertex is the source vertex of an elliptic form since these inequalities imply that the three edges at this vertex are oriented away from the vertex, so the monotonicity property guarantees that all values of the form are positive. The situation for source edges is simpler since any two positive integers p and q determine an elliptic form with a source edge having adjacent regions labeled p and q as in the earlier figure.

Hyperbolic Forms

The topographs of hyperbolic forms exhibit quite different behavior from the topographs of elliptic forms since they always have a ***periodic separator line*** of the sort that we saw in several of the examples in the previous chapter. Here is the general statement:

Theorem 5.2. *In the topograph of a hyperbolic form the edges for which the two adjacent regions are labeled by numbers of opposite sign form a line which is infinite in both directions, and the topograph is periodic along this line, with other edges of the topograph leading off the line on both sides.*

Proof: For a hyperbolic form Q all regions in the topograph have labels that are either positive or negative, never zero, and there must exist two regions of opposite sign. By moving along a path in the topograph joining two such regions we will somewhere encounter two adjacent regions of opposite sign. Thus there must exist edges whose two adjacent regions have opposite sign. Let us call these edges ***separating edges***.

At an end of a separating edge the value of Q in the next region must be either positive or negative since Q does not take the value 0:

This implies that exactly one of the two edges at each end of the first separating edge is also a separating edge. Repeating this argument, we see that each separating edge is part of a line of separating edges that is infinite in both directions, and the edges that lead off from this line are not separating edges.

The monotonicity property implies that as we move off this line of separating edges the values of Q are steadily increasing through positive integers on the positive side and steadily decreasing through negative integers on the negative side. In particular this means that there are no other separating edges that are not on the initial separator line, so there is only one separator line.

It remains to prove that the topograph is periodic along the separator line. We can assume all the edges along the separator line are oriented in the same direction by changing the signs of the h values if necessary. For an edge of the separator line labeled h with adjacent regions labeled p and $-q$ with $p > 0$ and $q > 0$, we know that $h^2 + 4pq$ is the discriminant Δ, by Proposition 5.1. The equation $\Delta = h^2 + 4pq$ with p and q positive implies that Δ is positive and furthermore that each of $|h|$, p, and q is less than Δ. Thus there are only finitely many possible values for h, p, and q along the separator line since Δ is a constant depending only on Q. It follows that there are only finitely many possible combinations of values h, p, and q at each edge on the separator line. Since the separator line is infinite, there must then be two edges on the line that have the same values of h, p, and q. Since the topograph is uniquely determined by the three labels h, p, q at a single edge, the translation of the line along itself that takes one edge to another edge with the same three labels must preserve all the labels on the line. This shows that the separator line is periodic.

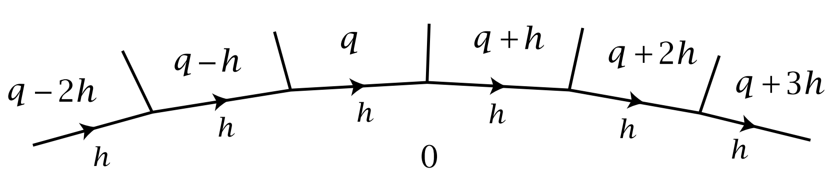
There must be edges leading away from the separator line on both the positive and the negative side, otherwise there would be just a single region on one side of the line, and then the second arithmetic progression rule would say that the h labels along the line formed an infinite arithmetic progression with nonzero increment $2p$ where p is the label on the region in question. However, this would contradict the fact that these h labels are periodic. $\square$

The qualitative behavior of the topograph of a hyperbolic form away from the separator line fits the pattern we have seen in examples. Since the separator line is periodic the whole topograph is periodic, consisting of repeating sequences of trees leading off from the separator line on each side, with monotonically increasing positive values of the form on each tree on the positive side of the separator line and monotonically decreasing negative values on the negative side, as a consequence of the monotonicity property.

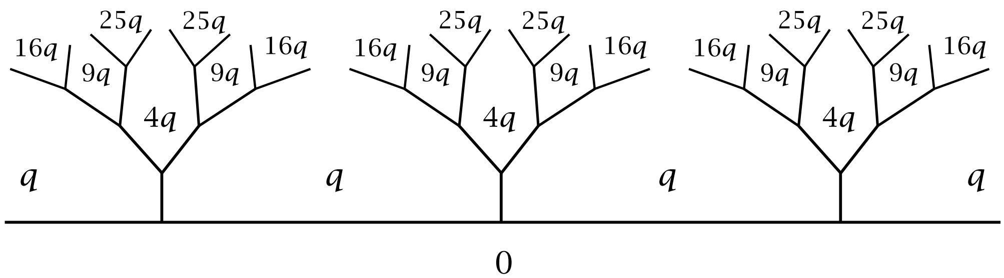
Parabolic and 0-Hyperbolic Forms

The remaining types of forms to consider are parabolic forms and 0-hyperbolic forms. These turn out to be less interesting, and they play only a minor role in the theory of quadratic forms.

Parabolic and 0-hyperbolic forms are the forms whose topograph contains at least one region labeled 0. By the second arithmetic progression rule, each edge adjacent to a 0 region has the same label h, and from this it follows that the labels on the regions adjacent to the 0 region form an arithmetic progression:

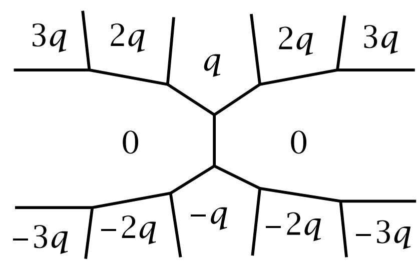


When $h = 0$ the topograph has the very simple pattern shown in the following figure:



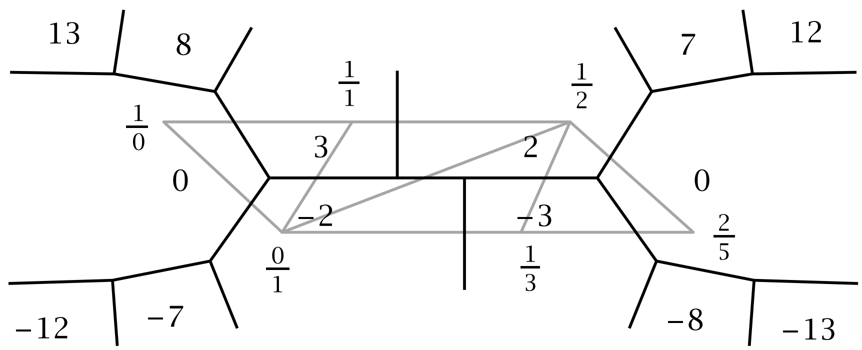
Thus the form is parabolic, taking on only positive or only negative values away from the 0 region, depending on the sign of q. We cannot have $q = 0$ since we are not allowing forms to be identically zero. An example of a form with this topograph is $Q(x,y) = qx^2$, with the 0 region at $x/y = 0/1$. The topograph is periodic along the 0 region since it consists of the same tree pattern repeated infinitely often.

The remaining case is that the label h on the edges bordering a 0 region is nonzero. The arithmetic progression of values of Q adjacent to the 0 region is then not constant, so it includes both positive and negative numbers, and hence Q is 0-hyperbolic. If the arithmetic progression includes the value 0, this gives a second 0 region adjacent to the first one, and the topograph is as shown at the right. An example of a form with this topograph is $Q(x,y) = qxy$, with the two 0 regions at $x/y = 1/0$ and $0/1$.



If the arithmetic progression of values of Q adjacent to the 0 region does not include 0, there will be an edge separating the positive from the negative values in the progression. We can extend this separating edge to a line of separating edges as we did with hyperbolic forms. If this extension does not eventually terminate with a second 0 region, the reasoning we used in the hyperbolic case would yield two edges along this line having the same h and the same positive and negative labels on the two adjacent regions, forcing the line to be periodic in the direction of this extension. This in turn would force it to be periodic in both directions by the arithmetic progression rule. But this is impossible since the line began with a 0 region at one end. Thus the topograph contains a finite separator line connecting two 0 regions.

An example of such a form is $Q(x,y) = qxy - py^2 = (qx - py)y$ which has the value 0 at $x/y = 1/0$ and at $x/y = p/q$ or the reduction of p/q to lowest terms if p and q are not coprime. Here we must have $|q| > 1$ for the two 0 regions to be nonadjacent. The separator line must follow the strip of triangles in the Farey diagram corresponding to the continued fraction for p/q since the separator line is dual to a finite strip of triangles with the vertices $1/0$ and p/q at its two ends. For example, for $p/q = 2/5$ the topograph of the form $5xy - 2y^2 = (5x - 2y)y$ is shown in the following figure:



General Conclusions

Having described the topographs of the four types of forms, we can now deduce the characterization of each type in terms of the discriminant:

Proposition 5.3. *The four types of forms are distinguished by their discriminants, which are negative for elliptic forms, positive nonsquares for hyperbolic forms, positive squares for* 0*-hyperbolic forms, and zero for parabolic forms.*

Proof: Consider first an elliptic form Q, which we may assume takes on only positive values since changing Q to $-Q$ does not change the discriminant. The topograph of Q contains either a source vertex or a source edge. For a source edge with the label $h = 0$ separating regions with positive labels p and q the discriminant is $\Delta = h^2 - 4pq = -4pq$, which is negative. For a source vertex with adjacent regions having positive labels p, q, r the edge between the p and q regions is labeled $h = p + q - r$ so the discriminant can be expressed in the following way:

$$\begin{aligned}\Delta = h^2 - 4pq &= (p+q-r)^2 - 4pq \\ &= p^2 + q^2 + r^2 - 2pq - 2pr - 2qr \\ &= p(p-q-r) + q(q-p-r) + r(r-p-q)\end{aligned}$$

In the last line the three quantities in parentheses are negative by the triangle inequalities, so Δ is again negative.

For a parabolic form the topograph contains a region labeled 0 bordered by edges labeled 0, so $\Delta = h^2 - 4pq = 0$. A 0-hyperbolic form has a region labeled 0 bordered by edges all having the same nonzero label h so $\Delta = h^2$, a positive square.

For an edge in the separator line for a hyperbolic form the adjacent regions have labels p and $-q$ with p and q positive so $\Delta = h^2 + 4pq$ is positive. To see that Δ is not a square, suppose the form is $ax^2 + bxy + cy^2$. Here a must be nonzero, otherwise the form would have the value 0 at $(x, y) = (1, 0)$, which is impossible for a hyperbolic form. If the discriminant was a square then the equation $az^2 + bz + c = 0$ would have a rational root $z = x/y$ with $y \neq 0$ by the familiar quadratic formula $z = (-b \pm \sqrt{b^2 - 4ac})/2a$. Thus we would have $a(x/y)^2 + b(x/y) + c = 0$ and hence $ax^2 + bxy + cy^2 = 0$, so the form would have the value 0 at a pair (x, y) with $y \neq 0$, which is again impossible for a hyperbolic form. $\square$

The presence or absence of periodicity in a topograph has the following consequence:

Proposition 5.4. *If an equation $Q(x,y) = n$ with $n \neq 0$ has one integer solution (x,y) then it has infinitely many integer solutions when Q is hyperbolic or parabolic, but only finitely many integer solutions when Q is elliptic or 0-hyperbolic.*

Proof: Consider first the hyperbolic and parabolic cases. Suppose (x,y) is a solution of $Q(x,y) = n$. If (x,y) is a primitive pair, then n appears in the topograph of Q so by periodicity it appears infinitely often, giving infinitely many solutions of $Q(x,y) = n$. If there is a nonprimitive solution (x,y) then it is d times a primitive pair (x',y') with $Q(x',y') = n/d^2$. The latter equation has infinitely many solutions (x',y') by what we just showed, hence $Q(x,y) = n$ has infinitely many solutions $(x,y) = (dx',dy')$.

For elliptic and 0-hyperbolic forms there is no periodicity, and the monotonicity property implies that each number appears in the topograph at most a finite number of times. Thus $Q(x,y) = n$ can have only finitely many primitive solutions. If it had infinitely many nonprimitive solutions, these would yield infinitely many primitive solutions of equations $Q(x,y) = m$ for certain divisors m of n. However, this is impossible since each equation $Q(x,y) = m$ for a fixed m can have only finitely many primitive solutions and n has only finitely many divisors since we assume it is nonzero. $\square$

Exercises

1. (a) Find two primitive elliptic forms $ax^2 + cy^2$ that have the same discriminant but take on different sets of values. Draw enough of the topographs of the two forms to make it apparent that they do not have exactly the same sets of values. (Remember that the topograph only shows the values $Q(x,y)$ for primitive pairs (x,y).)
(b) Do the same thing with hyperbolic forms $ax^2 + cy^2$.

2. (a) Show the quadratic form $Q(x,y) = 92x^2 - 74xy + 15y^2$ is elliptic by computing its discriminant.
(b) Find the source vertex or edge in the topograph of this form.
(c) Using the topograph of this form, find all the integer solutions of $92x^2 - 74xy + 15y^2 = 60$, and explain why your list of solutions is a complete list. (There are exactly four pairs of solutions $\pm(x,y)$, three of which will be visible in the topograph.)

3. Show that if a form takes the same value on two adjacent regions of its topograph, then these regions are both adjacent to the source vertex or edge when the form is elliptic, or both lie along the separator line when the form is hyperbolic.

4. Show that the minimum value of $|h|$ for all the edges in the border of a given region in the topograph of an elliptic or hyperbolic form occurs at an edge having an endpoint that achieves the minimum distance to the separator line or source vertex or edge of all vertices in the border of the given region.

5. (a) Show that if a quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ can be factored as a product $(Ax+By)(Cx+Dy)$ with A,B,C,D integers, then Q takes the value 0 at some pair of integers $(x,y) \neq (0,0)$, hence Q must be either 0-hyperbolic or parabolic. Show also, by a direct calculation, that the discriminant of this form is a square.

(b) Find a 0-hyperbolic form $Q(x,y)$ such that $Q(1,5)=0$ and $Q(7,2)=0$ and draw a portion of the topograph of Q that includes the two regions where $Q(x,y)=0$.

5.2 Equivalence of Forms

In the topographs we have drawn we often omit the fractional labels x/y for the regions in the topograph since the more important information is often just the values $Q(x,y)$ of the form. This leads to the idea of considering two quadratic forms to be equivalent if their topographs "look the same" when the labels x/y are disregarded. For a precise definition, one can say that quadratic forms Q_1 and Q_2 are ***equivalent*** if there is a vertex v_1 in the topograph of Q_1 and a vertex v_2 in the topograph of Q_2 such that the values of Q_1 in the three regions surrounding v_1 are equal to the values of Q_2 in the three regions surrounding v_2. For example if the values at v_1 are $2,2,3$ then the values at v_2 should also be $2,2,3$, in any order, but $2,3,3$ is regarded as different from $2,2,3$. Since the three values around a vertex determine all the other values in a topograph, having the same values at one vertex guarantees that the topographs look the same everywhere if the labels x/y are omitted.

An alternative definition of equivalence of forms would be to say that two forms are equivalent if there is a linear fractional transformation in $LF(\mathbb{Z})$ that takes the topograph of one form to the topograph of the other form. This is really the same as the first definition since there is a vertex of the topograph in the center of each triangle of the Farey diagram and we know that elements of $LF(\mathbb{Z})$ are determined by where they send a triangle, so if two topographs each have a vertex surrounded by the same triple of numbers, there is an element of $LF(\mathbb{Z})$ taking one topograph to the other, and conversely.

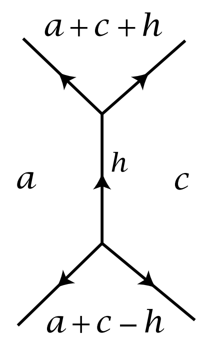
A topograph and its mirror image correspond to equivalent forms since the mirror image topograph has the same three labels around each vertex as at the corresponding vertex of the original topograph. For example, switching the variables x and y reflects the circular Farey diagram across its vertical axis and hence reflects the topograph of a form $Q(x,y)$ to the topograph of the equivalent form $Q(y,x)$. As another example, the forms $ax^2 + bxy + cy^2$ and $ax^2 - bxy + cy^2$ are always equivalent since they are related by changing (x, y) to $(-x, y)$, reflecting the Farey diagram across its horizontal axis, with a corresponding reflection of the topograph.

Equivalent forms have the same discriminant since the discriminant of a form is determined by the three numbers surrounding any vertex, as these three numbers determine the numbers p, q, h at each edge abutting the vertex and the discriminant is $h^2 - 4pq$ for any of these edges.

Our next goal will be to see how to compute all the different equivalence classes of forms of a given discriminant. The method for doing this will depend on which of the four types of forms we are dealing with.

Reduced Elliptic Forms

Let us look at elliptic forms first to see how to determine all the different equivalence classes for a given discriminant in this case. As usual it suffices to consider only the forms with positive values. At a source vertex or edge in the topograph of a positive elliptic form Q let the smaller two of the three adjacent values of Q be a and c with $a \le c$, and let the edge between them be labeled $h \ge 0$. The third of the three smallest values of Q is then $a + c - h$. The form Q is equivalent to the form $ax^2 + hxy + cy^2$ which has the values a, c, and $a + h + c$ for $(x, y) = (1, 0)$, $(0, 1)$, and $(1, 1)$. Since a and c are the smallest values of Q we have $a \le c \le a + c - h$, and the latter inequality is equivalent to $h \le a$. Summarizing, we have the inequalities $0 \le h \le a \le c$.



Thus every positive elliptic form is equivalent to a form $ax^2 + hxy + cy^2$ with $0 \le h \le a \le c$. An elliptic form satisfying these conditions is called ***reduced.*** Two different reduced elliptic forms with the same discriminant are never equivalent since a and c are the labels on the two regions in the topograph where the form takes its smallest values, and h is determined by a, c, and Δ via the formula $\Delta = h^2 - 4ac$ since we assume $h \ge 0$.

To avoid dealing with negative numbers let us set $\Delta = -D$ with $D > 0$, so the discriminant equation becomes $D = 4ac - h^2$. To find all equivalence classes of forms of discriminant $-D$ we therefore need to find all solutions of the equation

$$4ac = h^2 + D \qquad \text{with} \qquad 0 \le h \le a \le c$$

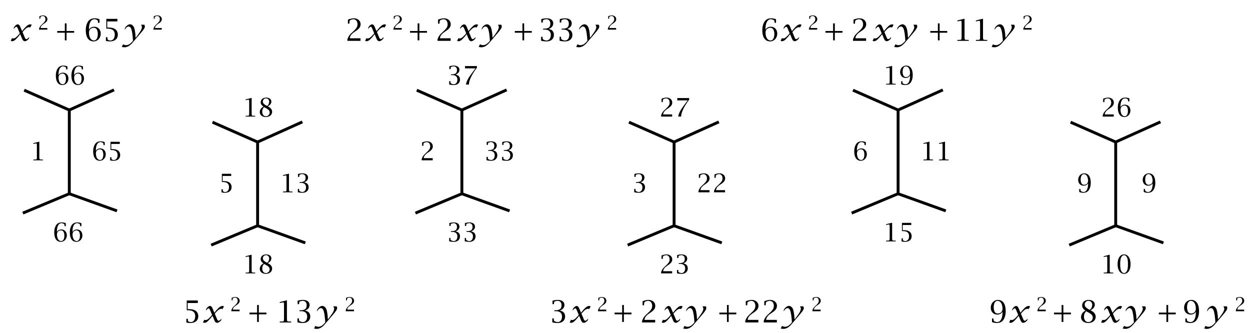
This equation implies that h must have the same parity as D, and we can bound the choices for h by the inequalities $4h^2 \le 4a^2 \le 4ac = D + h^2$ which imply $3h^2 \le D$, or $h^2 \le D/3$. This limits h to a finite number of possibilities, and for each of these values of h we just need to find all of the finitely many factorizations of $h^2 + D$ as $4ac$ with $a \le c$ and $h \le a$. In particular this shows that there are just finitely many equivalence classes of elliptic forms of a given discriminant.

As an example consider the case $\Delta = -260$, so $D = 260$. Since Δ is even, so is h, and we must have $h^2 \le 260/3$ so h must be $0, 2, 4, 6$, or 8. The corresponding values of a and c that are possible can then be computed from the equation $4ac = 260 + h^2$, always keeping in mind the requirement that $h \le a \le c$. The possibilities are shown in the following table:

h	ac	(a,c)
0	65	$(1,65)$, $(5,13)$
2	66	$(2,33)$, $(3,22)$, $(6,11)$
4	69	—
6	74	—
8	81	$(9,9)$

As a side comment, note that the values of ac increase successively by $1,3,5,7,\cdots$. This always happens when Δ is even, so the h values are $0,2,4,6,\cdots$. For odd Δ the values of h are $1,3,5,7,\cdots$ and the increments for ac are $2,4,6,8,\cdots$. (Let it be an exercise for the reader to figure out why these statements are true.)

From the table we see that every positive elliptic form of discriminant -260 is equivalent to one of the six reduced forms x^2+65y^2, $5x^2+13y^2$, $2x^2+2xy+33y^2$, $3x^2+2xy+22y^2$, $6x^2+2xy+11y^2$, or $9x^2+8xy+9y^2$, and no two of these reduced forms are equivalent to each other. Here are small parts of the topographs of these forms:



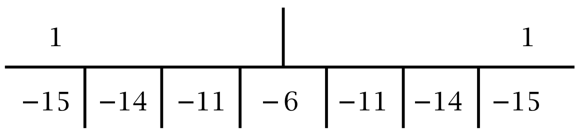
In the first two topographs the central edge is a source edge, and in the last four topographs the lower vertex is a source vertex.

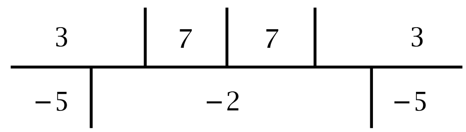
One might wonder what would happen if we continued the table with larger values of h not satisfying $h^2 \le {}^{260}\!/_3$. For example for $h = 10$ we would have $ac = 90$ so the condition $a \le c$ would force a to be 9 or less, violating the condition $h \le a$. Larger values of h would run into similar difficulties. The condition $h^2 \le {}^{D}\!/_3$ saves one the trouble of trying larger values of h.

Cycles of Hyperbolic Forms

Next we consider hyperbolic forms of a given discriminant $\Delta > 0$. The topograph of a hyperbolic form has a separator line, so for each edge in the separator line we have the edge label h with the adjacent regions labeled p and $-q$ for $p > 0$ and $q > 0$. We can assume $h \ge 0$ by reorienting the edge if necessary. The discriminant equation is $\Delta = h^2 + 4pq$. Since p and q are positive this implies $h^2 < \Delta$ so there are only finitely many possibilities for h along the separator lines of forms of the given discriminant Δ. For each h we then look at the factorizations $\Delta - h^2 = 4pq$. There can be only finitely many of these, so this means there are just finitely many possible combinations of labels $h, p, -q$ and hence only finitely many possible separator lines. Thus the number of equivalence classes of hyperbolic forms of a given discriminant is finite.

As an example, let us determine all the quadratic forms of discriminant 60, up to equivalence. Two obvious forms of discriminant 60 are $x^2 - 15y^2$ and $3x^2 - 5y^2$, whose separator lines consist of periodic repetitions of the following two patterns:





From the topographs it is apparent that these two forms are not equivalent, and also that the negatives of these two forms, $-x^2 + 15y^2$ and $-3x^2 + 5y^2$, give two more inequivalent forms, for a total of four equivalence classes so far. To see whether there are others we use the formula $\Delta = 60 = h^2 + 4pq$ relating the values p and $-q$ adjacent to an edge labeled h in the separator line, with $p > 0$ and $q > 0$. The various possibilities are listed in the table below. The equation $\Delta = h^2 + 4pq$ implies that h and Δ must have the same parity, just as in the elliptic case.

h	pq	(p,q)
0	15	$(1,15)$, $(3,5)$, $(5,3)$, $(15,1)$
2	14	$(1,14)$, $(2,7)$, $(7,2)$, $(14,1)$
4	11	$(1,11)$, $(11,1)$
6	6	$(1,6)$, $(2,3)$, $(3,2)$, $(6,1)$

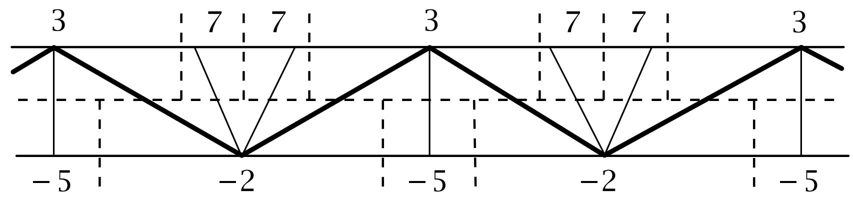
Each pair of values for (p,q) in the table occurs at some edge along the separator line in one of the two topographs shown above or the negatives of these topographs. Hence every form of discriminant 60 is equivalent to one of these four. If it had not been true that all the possibilities in the table occurred in the topographs of the forms we started with, we could have used these other possibilities for h, p, and q to generate new forms $px^2 + hxy - qy^2$ with new topographs, eventually exhausting all the finitely many possibilities.

The procedure in this example works for all hyperbolic forms. One makes a list of all the positive integer solutions of $\Delta = h^2 + 4pq$, then one constructs separator lines that realize all the resulting pairs (p,q). The different separator lines correspond exactly to the different equivalence classes of forms of discriminant Δ. Each solution (h,p,q) gives a form $px^2 + hxy - qy^2$. These are organized into *cycles* corresponding to the pairs $(p,-q)$ occurring along one of the periodic separator lines. Thus in the preceding example with $\Delta = 60$ the 14 pairs (p,q) in the table give rise to the four cycles along the four different separator lines.

A hyperbolic form $ax^2 + bxy + cy^2$ belongs to one of the cycles for the discriminant $\Delta = b^2 - 4ac$ exactly when $a > 0$ and $c < 0$ since a and c are the numbers p

and $-q$ lying on opposite sides of an edge of the separator line when $(x,y) = (1,0)$ and $(0,1)$.

If we superimpose the separator line of a hyperbolic form on the associated infinite strip in the Farey diagram, we see that the forms within a cycle correspond to the edges of the Farey diagram that lie in the strip and join one border of the strip to the other. For example, for the form $3x^2 - 5y^2$ we obtain the following picture, with fans of two triangles alternating with fans of three triangles:



The number of forms within a cycle can be fairly large in general. The situation can be improved somewhat by considering only the "most important" forms in the cycle, namely the forms that correspond to those edges in the strip that separate pairs of adjacent fans, indicated by heavier lines in the figure above. In terms of the topograph itself these are the edges in the separator line whose two endpoints have edges leading away from the separator line on opposite sides. The forms corresponding to these edges are traditionally called the *reduced* forms within the given equivalence class. In the example of discriminant 60 these are the forms with $(p,q) = (1,6)$, $(6,1)$, $(3,2)$, and $(2,3)$. These are the forms $x^2+6xy-6y^2$, $6x^2+6xy-y^2$, $3x^2+6xy-2y^2$, and $2x^2 + 6xy - 3y^2$. In this example there is just one reduced form for each cycle, but in more complicated examples there can be any number of reduced forms in a cycle. Note that the reduced forms do not necessarily give the simplest-looking forms, which in this example were the original forms $x^2 - 15y^2$ and $3x^2 - 5y^2$ along with their negatives $-x^2 + 15y^2$ and $-3x^2 + 5y^2$, or alternatively $15x^2 - y^2$ and $5x^2 - 3y^2$.

0-Hyperbolic and Parabolic Forms

For 0-hyperbolic forms it is rather easy to determine all the equivalence classes of forms of a fixed discriminant. As we saw in our initial discussion of 0-hyperbolic forms, their topographs contain two regions labeled 0, and the labels on the regions adjacent to each 0-region form an arithmetic progression with increment given by the label on the edges bordering the 0-region. Previously we called this edge label h but now let us change notation and call it q. We may assume q is positive by re-orienting the edges if necessary. The discriminant is $\Delta = q^2$ so both 0-regions must have the same edge label q. Either one of the two arithmetic progressions determines the form up to equivalence since two successive terms in the progression together with the 0 in the adjacent region give the three values of the form around a vertex in the topograph.

The form $qxy - py^2$ has discriminant q^2 and has $-p$ as one term of the arithmetic progression adjacent to the 0-region $x/y = 1/0$, namely in the region $x/y = 0/1$.

Thus every 0-hyperbolic form of discriminant q^2 is equivalent to one of these forms $qxy - py^2$. Arithmetic progressions with increment q can be thought of as congruence classes mod q, so only the mod q value of p affects the arithmetic progression and hence we may assume $0 \le p < q$. The number of equivalence classes of 0-hyperbolic forms of discriminant q^2 is therefore at most q, the number of congruence classes mod q. However, the number of equivalence classes could be smaller since each form has two 0 regions and hence two arithmetic progressions, which could be the same or different. Since either arithmetic progression determines the form, if the two progressions are the same then the topograph must have a mirror symmetry interchanging the two 0-regions. This always happens for example if the two 0-regions touch, which is the case $p = 0$ so the form is qxy and the mirror symmetry just interchanges x and y. If we let r denote the number of forms $qxy - py^2$ without mirror symmetry then the number of equivalence classes of 0-hyperbolic forms of discriminant q^2 is $q - r$ since each form without mirror symmetry has two different arithmetic progressions giving the same form.

For parabolic forms it is even easier to describe what all the different equivalence classes are since we have seen exactly what their topographs look like: There is a single region labeled 0 and all the regions adjacent to this have the same label q, which can be any nonzero integer, positive or negative. The integer q thus determines the equivalence class, so there is one equivalence class of parabolic forms for each nonzero integer q, with qx^2 being one form in this equivalence class. Parabolic forms all have discriminant 0, so in this case there are infinitely many different equivalence classes with the same discriminant. However, if we look only at primitive forms then there are just the two classes given by the forms $\pm x^2$.

Every parabolic form is equivalent to one of the forms qx^2 by a change of variables $T(x,y) = (sx+ty, ux+vy)$ with $sv-tu = \pm 1$, so every parabolic form factors as $q(sx+ty)^2$ for some pair of coprime integers s and t, with $q = \pm 1$ for primitive forms. Similarly, every 0-hyperbolic form is equivalent to a form $y(qx - py)$ so the form can be written as $(ux+vy)(q(sx+ty) - p(ux+vy))$ which can be simplified to a product $(Ax+By)(Cx+Dy)$ with A, B, C, D integers. Conversely, every form that factors as $(Ax+By)(Cx+Dy)$ with integer coefficients has the value 0 when $(x,y) = (-B,A)$ or $(-D,C)$ so the form must be parabolic or 0-hyperbolic. Parabolic forms are the case that the two linear factors are the same up to a constant multiple.

We have now shown how to compute all the equivalence classes of forms of a given discriminant for each of the four types of forms. In particular we have proved the following general fact:

Theorem 5.5. *There are only a finite number of equivalence classes of forms with a given nonzero discriminant.*

Exercises

1. (a) For positive elliptic forms of discriminant $\Delta = -D$, verify that the smallest value of D for which there are at least two inequivalent forms of discriminant $-D$ is $D = 12$.

(b) If we add the requirement that all forms under consideration are primitive, then what is the smallest D?

2. Determine all the equivalence classes of positive elliptic forms of discriminants -67, -104, and -347.

3. Find two elliptic forms that are not equivalent but take on the same three smallest values $a < b < c$.

4. Determine the number of equivalence classes of quadratic forms of discriminant $\Delta = 120$ and list one form from each equivalence class.

5. Do the same thing for $\Delta = 61$.

6. (a) Find the smallest positive nonsquare discriminant for which there is more than one equivalence class of forms of that discriminant. (In particular, show that all smaller discriminants have only one equivalence class.)

(b) Find the smallest positive nonsquare discriminant for which there are two inequivalent forms of that discriminant, neither of which is simply the negative of the other.

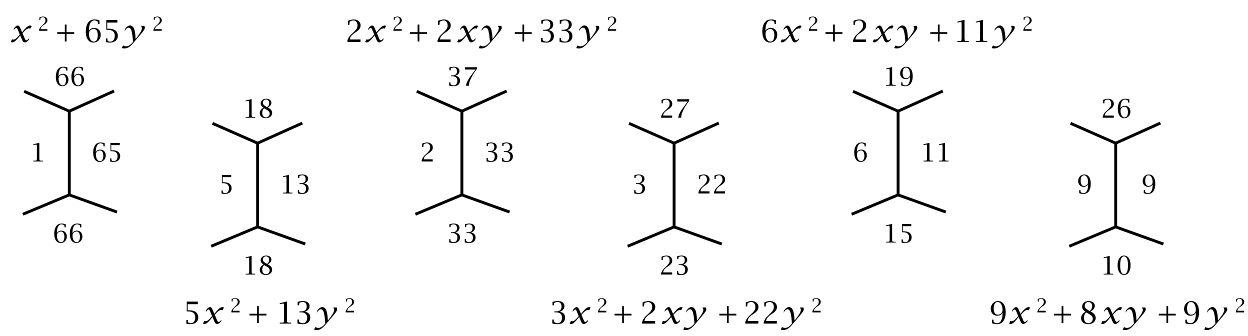
7. (a) Determine all the equivalence classes of 0-hyperbolic forms of discriminant 49.

(b) Determine which equivalence class in part (a) each of the forms $7xy - py^2$ for $p = 0, 1, 2, 3, 4, 5, 6$ belongs to.

5.3 The Class Number

When considering equivalence classes of forms of a given discriminant there are further refinements that turn out to be very useful. The first involves forms whose topographs are mirror images of each other. According to the definition we have given, two such forms are regarded as equivalent. However, there is a more refined notion of equivalence in which two forms are considered equivalent only if there is an orientation-preserving transformation in $LF(\mathbb{Z})$ taking the topograph of one form to the topograph of the other. In this case the forms are called ***properly equivalent.***

To illustrate the distinction between equivalence and proper equivalence, let us look at the earlier example of discriminant $\Delta = -260$ where we saw that there were six equivalence classes of forms:



In the first two topographs the central edge is a source edge and in the other four the lower vertex is a source vertex. Whenever there is a source edge the topograph has mirror symmetry across a line perpendicular to the source edge. When there is a source vertex there is mirror symmetry only when at least two of the three surrounding values of the form are equal, as in the third and sixth topographs above, but not the fourth or fifth topographs. Thus the mirror images of the fourth and fifth topographs correspond to two more quadratic forms which are not equivalent to them under any orientation-preserving transformation. With the more refined notion of proper equivalence there are therefore eight proper equivalence classes of forms of discriminant -260.

To obtain explicit formulas for the mirror image forms we can interchange the coefficients a and c in $ax^2+bxy+cy^2$, which corresponds to interchanging x and y, reflecting the topograph across a vertical line. Alternatively we could change the sign of b, which corresponds to changing the sign of either x or y and thus reflecting the topograph across a horizontal line.

For a general discriminant Δ each equivalence class of forms of discriminant Δ gives rise to two proper equivalence classes except when the class contains forms with mirror symmetry, in which case equivalence and proper equivalence amount to the same thing since every orientation-reversing equivalence can be converted into an orientation-preserving equivalence by composing with a mirror reflection. Here we are using the fact that the only linear fractional transformations that take a topograph to itself and reverse orientation are mirror reflections, as will be shown in Section 5.4 when we study symmetries of topographs in more detail.

Multiplying a form by an integer $d > 1$ does not change its essential features in any significant way, so it is reasonable when classifying forms to restrict attention just to primitive forms, the forms that are not proper multiples of other forms. In other words, one considers only the forms $ax^2 + bxy + cy^2$ for which a, b, and c have no common divisor greater than 1. The primitivity of a form is detectable just from the numbers appearing in its topograph since all the numbers in the topograph of a nonprimitive form are divisible by some number $d > 1$, and conversely if all numbers in the topograph of a form $ax^2+bxy+cy^2$ are divisible by d then in particular a, c, and $a+b+c$, the values at $(1,0)$, $(0,1)$, and $(1,1)$, are divisible by d which implies that b is also divisible by d so the whole form is divisible by d. Thus primitivity is a property of equivalence classes of forms. Multiplying a form by d multiplies its discriminant by d^2, so nonprimitive forms of discriminant Δ exist exactly when Δ is a square times another discriminant. For example, when $\Delta = -12 = 4(-3)$ one has the primitive form x^2+3y^2 as well as the nonprimitive form $2x^2+2xy+2y^2$ which is twice the form x^2+xy+y^2 of discriminant -3.

The number of proper equivalence classes of primitive forms of a given discriminant is called the ***class number*** for that discriminant, where in the case of elliptic forms one considers only the forms with positive values. The traditional notation for the class number for discriminant Δ is h_Δ. (This h has nothing to do with the h labels on edges in topographs.)

Since we have an algorithm for computing the finite set of equivalence classes of forms of a given nonzero discriminant, this leads to an algorithm for computing class numbers. When computing the table of triples (h,a,c) for elliptic forms or (h,p,q) for hyperbolic forms we omit the nonprimitive triples since these correspond to nonprimitive forms. Then we determine which of the remaining forms have mirror symmetry. For elliptic forms these are the cases when one or more of the inequalities $0 \le h \le a \le c$ is an equality, as we will see in the next section. For hyperbolic forms mirror symmetries can be detected in the separator line. Forms with mirror symmetry count once when computing the class number, and forms without mirror symmetry count twice. However, just having an algorithm to compute the class number h_Δ does not make it transparent how h_Δ depends on Δ, and indeed this is a very difficult question which is still only partially understood.

Of special interest are the discriminants for which all forms are primitive. These are called ***fundamental discriminants***. Thus a fundamental discriminant is one which is not a square times a smaller discriminant. For example, 8 is a fundamental discriminant even though it is divisible by a square, 4, since the other factor 2 is not the discriminant of any form, as it is not congruent to 0 or 1 mod 4. Technically 1 is a fundamental discriminant according to our definition, but we will exclude this trivial case. Thus fundamental discriminants are never squares, so fundamental discriminants appear only for elliptic and hyperbolic forms. With 1 excluded it is easy to check that the fundamental discriminants Δ with $|\Delta| < 40$ are 5, 8, 12, 13, 17, 20, 21, 24, 28, 29, 33, 37 and -3, -4, -7, -8, -11, -15, -19, -20, -23, -24, -31, -35, -39.

It is not hard to give a precise characterization of the discriminants Δ that are fundamental. First write $\Delta = 2^k n$ with $k \ge 0$ and n odd, possibly negative. If any odd square divides n then we can factor this out of Δ and still get a discriminant since odd squares are congruent to 1 mod 4 so multiplying by an odd square does not affect whether a number is 0 or 1 mod 4. The exponent k in 2^k can never be 1 since this would imply $\Delta \equiv 2$ mod 4. If $k \ge 4$ we can factor powers of 4 out of

Δ until we have k equal to 2 or 3 and still have a discriminant. If $k = 3$ we cannot factor a 4 out of Δ since this would give the excluded case $k = 1$. If $k = 2$ we can factor $4 = 2^k$ out of Δ exactly when $n \equiv 1 \bmod 4$. Finally, when $k = 0$ we have $\Delta = n$ so we must have $n \equiv 1 \bmod 4$. Thus the fundamental discriminants other than -4 and ± 8 are of three types:

- $\Delta = n$ with $|n|$ a product of distinct odd primes and $n \equiv 1 \bmod 4$.
- $\Delta = 4n$ with $|n|$ a product of distinct odd primes and $n \equiv 3 \bmod 4$.
- $\Delta = 8n$ with $|n|$ a product of distinct odd primes.

Every nonsquare discriminant can be factored uniquely as $\Delta = d^2\Delta'$ where Δ' is a fundamental discriminant and $d \geq 1$. The number d is called the ***conductor*** of Δ. Fundamental discriminants are those whose conductor is 1. Conductors will become important when we study the deeper properties of forms in later chapters. The class number h_Δ is always a multiple of $h_{\Delta'}$ and there is a not-too-complicated formula for what this multiple is, so the determination of class numbers reduces largely to the case of fundamental discriminants. However, we will not be going into more detail on the relationship between h_Δ and $h_{\Delta'}$ since this would lead us somewhat outside the scope of the book.

Discriminants of Class Number 1

The question of which discriminants have class number 1 has been much studied. This amounts to finding the discriminants for which all primitive forms are equivalent since if all primitive forms are equivalent, they are all equivalent to the principal form which has mirror symmetry so they are all properly equivalent to the principal form.

For elliptic forms the following nine fundamental discriminants have class number 1:

$$\Delta = -3,\ -4,\ -7,\ -8,\ -11,\ -19,\ -43,\ -67,\ -163$$

In addition there are four more which are not fundamental: -12, -16, -27, -28. It was conjectured by Gauss around 1800 that there are no other negative discriminants of class number 1. Over a century later in the 1930s it was shown that there is at most one more, and then in the 1950s and 1960s Gauss's conjecture was finally proved completely.

Another result from the 1930s is that for each number n there are only finitely many negative discriminants with class number n. Finding what these discriminants are is a difficult problem, however, and so far this has been done only in the range $n \leq 100$.

The situation for positive discriminants with class number 1 is not as well understood. Computations show that there are a large number of positive fundamental discriminants with class number 1, and it seems likely that there are in fact infinitely many. However, this has not been proved and remains one of the most basic unsolved problems about quadratic forms. If one allows nonfundamental discriminants then it is known that there are infinitely many with $h_\Delta = 1$, including for example the discriminants $\Delta = 2^{2k+1}$ for $k \geq 1$ and $\Delta = 5^{2k+1}$ for $k \geq 0$.

Returning to the nine negative fundamental discriminants of class number 1, it is easy to check in each case that all forms are equivalent. For example when $\Delta = -163$ and we apply the earlier algorithm to find all reduced forms we must have h odd with $h^2 \leq 163/3$ so the only possibilities are $h = 1, 3, 5, 7$. From the equation $4ac = 163 + h^2$ the corresponding values of ac are $41, 43, 47, 53$ which all happen to be prime, and since $a \leq c$ this forces a to be 1 in each case. But since $h \leq a$ this means h must be 1, and we obtain the single quadratic form $x^2 + xy + 41y^2$.

The corresponding polynomial $x^2 + x + 41$ has a curious property discovered by Euler: For each $x = 0, 1, 2, 3, \cdots, 39$ the value of $x^2 + x + 41$ is a prime number. Here are these forty primes:

41 43 47 53 61 71 83 97 113 131 151 173 197 223 251 281 313
347 383 421 461 503 547 593 641 691 743 797 853 911 971
1033 1097 1163 1231 1301 1373 1447 1523 1601

Notice that the successive differences between these primes are $2, 4, 6, 8, 10, \cdots, 78$ since $[(x+1)^2 + (x+1) + 41] - [x^2 + x + 41] = 2(x+1)$. The next number in the sequence after 1601 would be $1681 = 41^2$, not a prime. (Write $x^2 + x + 41$ as $x(x+1) + 41$ to see why $x = 40$ must give a nonprime.) A similar thing happens for the other negative fundamental discriminants of class number 1. The nontrivial cases are listed in the table below, where $D = -\Delta$.

D		
7	$x^2 + x + 2$	2
11	$x^2 + x + 3$	3 5
19	$x^2 + x + 5$	5 7 11 17
43	$x^2 + x + 11$	11 13 17 23 31 41 53 67 83 101
67	$x^2 + x + 17$	17 19 23 29 37 47 59 73 89 107 127 149 173 199 227 257

Satisfactory explanations are known for the occurrence of so many prime values of these quadratic polynomials but they involve fairly deep theory. It is curious that the lists of prime values account for all primes less than 100 except 79.

Suppose one asks about the next forty values of $x^2 + x + 41$ after the value 41^2 when $x = 40$. The next value, when $x = 41$, is $1763 = 41 \cdot 43$, also not a prime. After this the next two values are primes, then comes $2021 = 43 \cdot 47$, then four primes, then $2491 = 47 \cdot 53$, then six primes, then $3233 = 53 \cdot 61$, then eight primes, then $4331 = 61 \cdot 71$, then ten primes, then $5893 = 71 \cdot 83$. This last number was for $x = 76$, and the next four values are prime as well for $x = 77, 78, 79, 80$, completing the second 40 values. But then the pattern breaks down when $x = 81$ where one gets the value $6683 = 41 \cdot 163$. Thus, before the breakdown, not only were we getting sequences of 2, 4, 6, 8, 10 primes but the nonprime values were the products of two successive terms in the original sequence of prime values $41, 43, 47, 53, 61, \cdots$.

All this seems quite surprising, even if the nice patterns do not continue forever. A partial explanation can be found in the fact that the polynomial $P(x) = x^2 + x + 41$ satisfies the identity $P(40 + n^2) = P(n - 1)P(n)$ as one can easily check, so when $n = 1, 2, 3, \cdots$ we get $P(41) = P(0)P(1) = 41 \cdot 43$, $P(44) = 43 \cdot 47$, $P(49) = 47 \cdot 53$, $P(56) = 53 \cdot 61$, etc. However this does not explain why the intervening values of $P(x)$ should be prime. The polynomials in the preceding table exhibit similar behavior.

Exercises

1. Compute the class number for each of the following discriminants:
(a) -23 (b) -47 (c) -71 (d) -87 (e) -92 (f) 145 (g) 148.

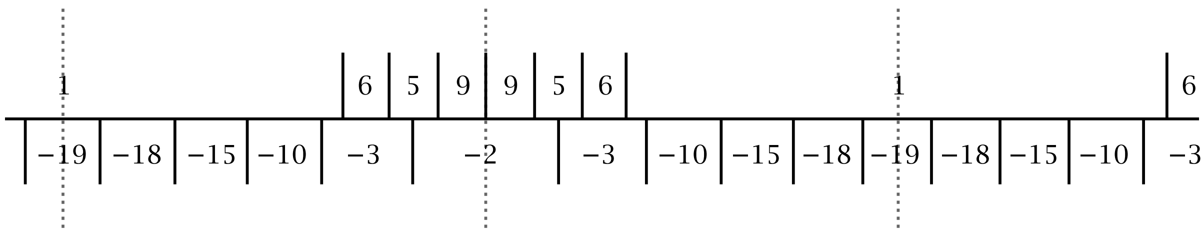
2. In this extended exercise the goal will be to show that the only negative even discriminants with class number 1 are -4, -8, -12, -16, and -28. (Of these only -4 and -8 are fundamental discriminants.) The strategy will be to exhibit an explicit reduced primitive form Q different from the principal form $x^2 + dy^2$ for each discriminant $-4d$ with $d > 4$ except $d = 7$. This will be done by breaking the problem into several cases, where in each case a form Q will be given and you are to show that this form has the desired properties, namely it is of discriminant $-4d$, primitive, reduced, and different from the principal form. You should also check that the cases considered cover all possibilities.
(a) Suppose d is not a prime power. Then it can be factored as $d = ac$ where $1 < a < c$ and a and c are coprime. In this case let Q be the form $ax^2 + cy^2$.
(b) The form $ax^2 + 2xy + cy^2$ will work provided that $d + 1$ factors as $d + 1 = ac$ where a and c are coprime and $1 < a < c$. If d is odd, for example a power of an odd prime, then $d + 1$ is even so it has such a factorization $d + 1 = ac$ unless $d + 1 = 2^n$.
(c) If $d = 2^n$ the cases we need to consider are $n \geq 3$ since $d > 4$. When $n = 3$ take Q to be $3x^2 + 2xy + 3y^2$ and when $n \geq 4$ take Q to be $4x^2 + 4xy + (2^{n-2} + 1)y^2$.
(d) When $d + 1 = 2^n$ the cases of interest are $n \geq 3$. When $n = 3$ we have $d = 7$ which is one of the allowed exceptions with class number 1. When $n = 4$ we have $d = 15$ and $3x^2 + 5y^2$ works as in part (a). When $n = 5$ we have $d = 31$ and we take the form $5x^2 + 4xy + 7y^2$. When $n \geq 6$ we use the form $8x^2 + 6xy + (2^{n-3} + 1)y^2$.

3. Show that the class number for discriminant $\Delta = q^2 > 1$ is $\varphi(q)$ where $\varphi(q)$ is the number of positive integers less than q and coprime to q.

5.4 Symmetries of Forms

We have observed that some topographs are symmetric in various ways. To give a precise meaning to this term, let us say that a ***symmetry*** of a form Q or its topograph is a transformation T in $LF(\mathbb{Z})$ that leaves all the values of Q unchanged, so $Q(T(x,y)) = Q(x,y)$ for all pairs (x,y). For example, every hyperbolic form has a periodic separator line, which means there is a symmetry that translates the separator line along itself. If T is the symmetry translating by one period in either direction, then all the positive and negative powers of T are also translational symmetries. Strictly speaking, the identity transformation is always a symmetry but we will sometimes ignore this trivial symmetry.

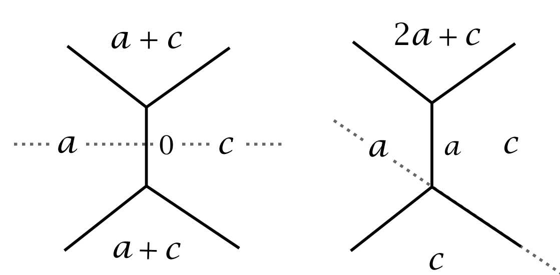
Some hyperbolic forms also have mirror symmetry, where the symmetry is reflection across a line perpendicular to the separator line. This reflector line could contain one of the edges leading off the separator line, or it could be halfway between two consecutive edges leading off the separator line on the same side. Both kinds of symmetry occur along the separator line of the form $x^2 - 19y^2$, for example:



Elliptic forms can have mirror symmetries as well, as we saw in the earlier example $\Delta = -260$ where two topographs had mirror symmetry across a line perpendicular to an edge and two had symmetry across a line containing an edge.

Proposition 5.6. *A number a appears on the reflector line of a mirror symmetry of the topograph of a form Q exactly when Q is equivalent to a form $ax^2 + cy^2$ or $ax^2 + axy + cy^2$. In both cases a divides the discriminant of Q.*

In particular the principal forms $x^2 - ky^2$ and $x^2 + xy - ky^2$ have mirror symmetry, so there is at least one form with mirror symmetry in each discriminant.



Proof: The figures at the right show the two types of mirror symmetries, where the reflector line is either the perpendicular bisector of an edge of the topograph or contains an edge of the topograph. Let a and c be the labels on the left and right regions as in the figures, so the reflector line passes through the a region. If the edge between the left and right regions is labeled h then the regions above and below this edge are labeled $a + c + h$ and $a + c - h$. In the first figure the mirror symmetry forces h to be 0 so the form is equivalent to the form $ax^2 + cy^2$. In the second figure the mirror symmetry forces the lower region to be labeled c and this forces h to equal a when the edge labeled h is oriented upward. The form is then equivalent to the form $ax^2 + axy + cy^2$.

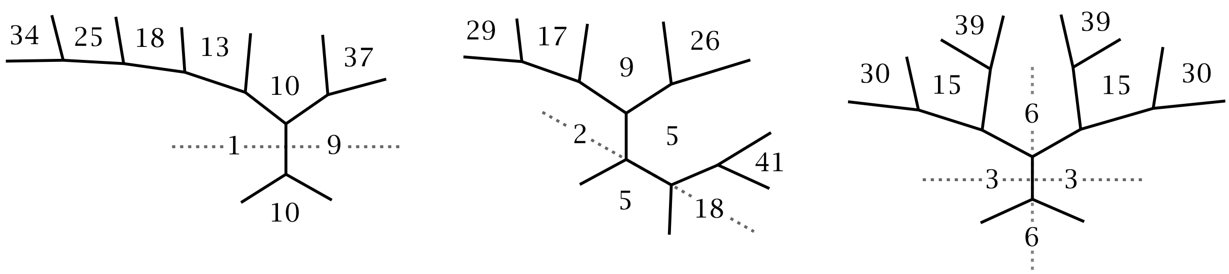
Conversely, the forms $ax^2 + cy^2$ and $ax^2 + axy + cy^2$ have topographs as shown in the figures, so these topographs have mirror symmetry with the reflector line passing through the a region. These two forms have discriminants $-4ac$ and $a^2 - 4ac$, both divisible by a. $\square$

As the proof showed, reflector lines crossing an edge in the topograph correspond to forms $ax^2 + cy^2$ and reflector lines containing an edge correspond to forms $ax^2 + axy + cy^2$. For example, a form $ax^2 + bxy + ay^2$ has mirror symmetry interchanging x and y, reflecting across the vertical axis of the circular Farey diagram which contains an edge of the topograph, so this form is equivalent to a form $Ax^2 + Axy + Cy^2$. The reflector line passes through regions of the topograph labeled $2a + b$ and $2a - b$ so A can be taken to be either $2a + b$ or $2a - b$, with $C = a$ since this is the value of the form at $x/y = 0/1$.

Proposition 5.7. *Let a be a divisor of the discriminant Δ that is either odd or twice an odd number. Then there exists a form $ax^2 + cy^2$ or $ax^2 + axy + cy^2$ of discriminant Δ having a in its topograph. If a is squarefree, a form of discriminant Δ with a in its topograph is unique up to equivalence, and a appears in the topograph only on a reflector line of a mirror symmetry.*

The conditions on the number a can be illuminated by looking at the case $\Delta = -36$ where there are three equivalence classes of forms:

$x^2 + 9y^2$ $\qquad$ $2x^2 + 2xy + 5y^2$ $\qquad$ $3x^2 + 3y^2$



The first two topographs have a single reflector line while the third has two reflector lines. The positive divisors of 36 are $1, 2, 3, 4, 6, 9, 12, 18$, and 36. The divisors that appear in the topographs are the ones that are odd or twice an odd number, so 4, 12, and 36 are excluded. Of the divisors that do appear, the ones that are not squarefree are 9 and 18, and these appear in more than one topograph, and off the reflector lines as well as on them.

Proof of Proposition 5.7: Suppose first that Δ is even. For the given divisor a of Δ let us first look for a form $ax^2 + cy^2$ since this has even discriminant. Thus we want an integer c such that $\Delta = -4ac$. Since Δ is even it is divisible by 4, so if a is odd and divides Δ then $4a$ divides Δ so the desired integer c exists in this case.

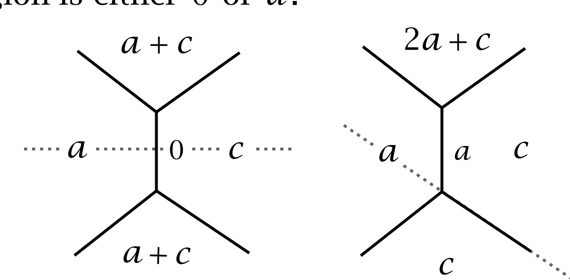
Since Δ is even it is either $8k$ or $8k + 4$ for some integer k. If $\Delta = 8k$ then $\Delta = -4ac$ can again be solved for c when a is twice an odd number.

When $\Delta = 8k+4$ and a is twice an odd number the equation $\Delta = -4ac$ will not have an integer solution c since $-4ac$ is divisible by 8, so we instead look for a form $ax^2+axy+cy^2$. This has $\Delta = a(a-4c)$ and we want to find an integer c such that $\Delta/a = a-4c$. This is equivalent to saying $\Delta/a \equiv a \bmod 4$. We have $a = 2(2m+1)$ so $a \equiv 2 \bmod 4$. For Δ/a, if we first divide Δ by 2 we get $4k+2$, then dividing by $2m+1$ can only change the congruence class mod 4 by a sign since odd numbers are $\pm 1 \bmod 4$. Thus $\Delta/a \equiv 2 \bmod 4$ so the congruence $\Delta/a \equiv a \bmod 4$ is satisfied. This finished the proof of the existence of a form ax^2+cy^2 or $ax^2+axy+cy^2$ when Δ is even.

Suppose now that Δ is odd, hence also its divisor a. Since Δ is odd, we are looking for a form $ax^2+axy+cy^2$. As above, the condition for having such a form is the congruence $\Delta/a \equiv a \bmod 4$. This is satisfied since $\Delta \equiv 1 \bmod 4$ and $a \equiv \pm 1 \bmod 4$.

Now we turn to the second statement in the proposition where we assume a is a squarefree divisor of Δ. Suppose that a appears in the topograph of a form of discriminant Δ. If b is one of the labels on an edge of the topograph bordering the region labeled a then we have $\Delta = b^2-4ac$ for c the label on the other region adjacent to the b edge. Since we assume a divides $\Delta = b^2-4ac$ it must also divide b^2, and if a is squarefree it will therefore divide b. Thus we have $b = ma$ for some integer m. The labels on the edges bordering the a region form an arithmetic progression with increment $2a$ so these are the numbers $b+2ka$ as k ranges over all integers. Since $b = ma$ we can factor $b+2ka$ as $(m+2k)a$. The numbers $m+2k$ for varying k form an arithmetic progression consisting of all even numbers if m is even and all odd numbers if m is odd. Thus we can choose k so that $m+2k$ is either 0 or 1, and hence the arithmetic progression $(m+2k)a$ contains either 0 or a. This means one of the edge labels on the border of the a region is either 0 or a.

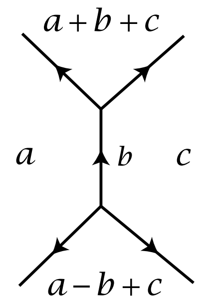
The topograph near this edge has the shape shown in one of the two figures at the right. From this we see that there is a reflector line passing through the a region and the form is equivalent to either ax^2+cy^2 or $ax^2+axy+cy^2$.



To finish the proof we only need to see that there cannot be both a form ax^2+cy^2 and a form $ax^2+axy+c'y^2$ with the same a and the same discriminant. Equating the discriminants of these two forms, we would have $-4ac = a^2-4ac'$ and therefore $a = 4(c'-c)$, but a would then be divisible by 4 and thus not squarefree. $\square$

Symmetries of Elliptic Forms

Let us consider now what sorts of symmetries are possible in general for the various types of forms, beginning with elliptic forms. For an elliptic form each symmetry must take the source vertex or edge to itself since this is where the smallest values of the form occur. In the case of a source edge, if a symmetry does not interchange the two ends of the source edge then the symmetry must be either the identity or a reflection across a line containing the source edge. If a symmetry does interchange the two ends of a source edge then it must either be a reflection across a line perpendicular to the edge or a 180 degree rotation of the topograph about the midpoint of the edge. Referring to the figure at the right, this rotation can only give a symmetry if $a = c$ and $a+b+c = a-b+c$ which is equivalent to having $b = 0$. Thus the form is $ax^2 + ay^2$ so if it is primitive it is just $x^2 + y^2$. Note that multiplying any form by a constant does not affect its symmetries so there is no harm in considering only primitive forms. For the form $x^2 + y^2$ note also that this form has both types of mirror symmetries, and the composition of these two mirror symmetries is the 180 degree rotational symmetry.



For a source vertex, a symmetry must take this vertex to itself. If a symmetry is orientation-preserving and not the identity then it must be a rotation about the source vertex by either one-third or two-thirds of a full turn. In either case this means that the three labels around the source vertex must be equal, so if the source vertex is the lower vertex in the figure above then the condition is $a = c = a - b + c$, which is equivalent to saying $a = b = c$. The form is then $ax^2 + axy + ay^2$ so if it is primitive it is $x^2 + xy + y^2$. The only other sort of symmetry for a source vertex is reflection across a line containing one of the three edges that meet at the source vertex. The only time there can be more than one such symmetry is when all three adjacent labels are equal so we are again in the situation of a form $ax^2 + axy + ay^2$.

For an elliptic form $ax^2 + bxy + cy^2$ that is reduced, so $0 \le b \le a \le c$, it is easy to recognize exactly when symmetries occur, namely when at least one of these three inequalities becomes an equality. Again using the figure above, when $b = 0$ one has a source edge with a mirror symmetry across the perpendicular line. When $b = a$ we have $a - b + c = c$ so there is a mirror symmetry across the lower right edge. And when $a = c$ one has mirror symmetry across the central edge. Since a and c are the two smallest labels on regions in the topograph, we see that reduced forms $ax^2 + bxy + ay^2$ occur when the smaller two of the three labels at the source vertex are equal, and reduced forms $ax^2 + axy + cy^2$ occur when the larger two labels are equal, at $0/1$ and $-1/1$.

Certain combinations of equalities in $0 \le b \le a \le c$ are also possible. If $b = 0$ and $a = c$ the form is $a(x^2 + y^2)$ with a source edge and both types of mirror symmetry as well as 180 degree rotational symmetry. Another possibility is that $b = a = c$ so the form is $a(x^2 + xy + y^2)$ with the symmetries described earlier. These are the only combinations of equalities that can occur since we must have $a > 0$ so $0 = b = a$ is impossible.

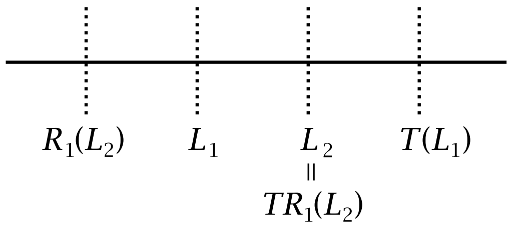
For reduced elliptic forms this exhausts all the possible symmetries since if we have strict inequalities $0 < b < a < c$ then the values of the form in the four regions shown in the figure above are all distinct. The first time this occurs is when the inequalities are $0 < 1 < 2 < 3$ so the form is $2x^2 + xy + 3y^2$ of discriminant -23.

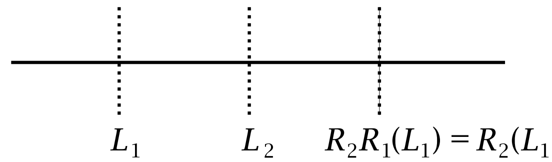
Symmetries of Hyperbolic Forms

Now consider hyperbolic forms. These all have periodic separator lines so they always have translational symmetries, and the question is what other sorts of symmetries are possible. For a hyperbolic form each symmetry must take the separator line to itself since this line consists of the edges that separate positive from negative values of the form. It is a simple geometric fact that a symmetry of a line L that is divided into a sequence of edges, say of length 1, extending to infinity in both directions, must be either a translation along L by some integer distance in either direction, or a reflection of L fixing either a vertex of L or the midpoint of an edge of L and interchanging the two halves of L on either side of the fixed point. This can be seen as follows. Symmetries of L are assumed to take vertices to vertices, so suppose the symmetry T sends a vertex v to the vertex $T(v)$. Then if T preserves the orientation of L it must be a translation along L by the distance from v to $T(v)$ as one can see by considering what T does to the two edges adjacent to v, then to the next two adjacent edges on either side, then the next two edges, and so on. If T reverses the orientation of L then either $T(v) = v$ or T fixes the midpoint of the segment from v to $T(v)$ since it sends this segment to a segment of the same length with one end at $T(v)$ but extending back toward v since T reverses orientation of L. Thus T fixes a point of L in either case, and it follows that T must reflect L across this fixed point, as one can again see by considering the edge or edges containing the fixed point, then the next two edges, and so on. If the distance from v to $T(v)$ is an even integer, the midpoint between v and $T(v)$ will be a vertex, and if it is odd, the midpoint will be a midpoint of an edge.

Returning to the situation of a symmetry T of the topograph of a hyperbolic form that takes the separator line L to itself, T must also take the side of L with positive labels to itself, so T preserves orientation of the plane exactly when it preserves orientation of L. Thus the only orientation-preserving symmetries of the topograph are translations along the separator line, and the only orientation-reversing symmetries are the two kinds of reflections across lines perpendicular to L.

If the separator line of a hyperbolic form has a mirror symmetry then because of periodicity there has to be at least one reflector line in each period, but in fact there are exactly two reflector lines in each period. To see this, let T be the translation by one period and let R_1 be a reflection across a reflector line L_1. Consider the composition TR_1, reflecting first by R_1 then translating by T, so TR_1 is an orientation-reversing symmetry. If L_2 is the line halfway between L_1 and $T(L_1)$ then $T(R_1(L_2)) = L_2$ as we can see in the first figure below:





Thus TR_1 is an orientation-reversing symmetry that takes L_2 to itself while preserving the positive and negative sides of the separator line, so TR_1 must be a reflection R_2 across L_2. This shows that there are at least two reflector lines in each period. There cannot be more than two since if R_1 and R_2 are the reflections across two adjacent reflector lines L_1 and L_2 as in the second figure then the composition R_2R_1, first reflecting by R_1 then by R_2, is orientation-preserving and sends L_1 to $R_2(R_1(L_1)) = R_2(L_1)$ so this composition is a symmetry translating the separator line by twice the distance between L_1 and L_2. The distance between L_1 and L_2 must then be half the length of the period, otherwise if the translation R_2R_1 were some power T^n of the basic periodicity translation T with $|n| > 1$, there would be fewer than two reflector lines in a period.

For completeness let us also describe the symmetries for the remaining two types of forms besides elliptic and hyperbolic forms. For a 0-hyperbolic form, if the two regions labeled 0 in the topograph have a border edge in common then a symmetry must take this edge to itself, and it cannot interchange the ends of the edge since positive values must go to positive values. The only possibility is then a reflection across this edge, which is always a symmetry of the topograph. If the two 0-regions do not have a common border edge they are joined by a finite separator line and a symmetry must take this line to itself without interchanging the positive and negative sides. The only possibility is a reflection across a line perpendicular to the separator line and passing through its midpoint. This reflection gives a symmetry only when the finite continued fraction associated to the form is palindromic.

A parabolic form has a single 0-region in its topograph, so the bordering line for this region must be taken to itself by any symmetry. Every symmetry of this bordering line gives a symmetry of the form, either a translation along the line or a reflection across a perpendicular line.

The preceding analysis shows in particular the following fact:

Proposition 5.8. *All orientation-reversing symmetries of the topograph of a form are mirror symmetries, reflecting across a line that is either perpendicular to or contains an edge of the topograph.*

Traditionally, a form whose topograph has an orientation-reversing symmetry is called *ambiguous* although there is really nothing about the form that is ambiguous in the usual sense of the word, unless perhaps it is the fact that such a form is indistinguishable from its mirror image.

The Symmetric Class Number

Let us define the ***symmetric class number*** h^s_Δ to be the number of equivalence classes of primitive forms of discriminant Δ with mirror symmetry. Recall that equivalence is the same as proper equivalence for forms with mirror symmetry. The ordinary class number h_Δ is thus h^s_Δ plus twice the number of equivalence classes of primitive forms without mirror symmetry. We have $h_\Delta \geq h^s_\Delta$, and in fact h_Δ is always a multiple of h^s_Δ as we will see in Proposition 7.16.

In contrast with h_Δ, the number h^s_Δ can be computed explicitly. Here is the result for elliptic and hyperbolic forms:

Theorem 5.9. *If Δ is a nonsquare discriminant and k is the number of distinct prime divisors of Δ then $h^s_\Delta = 2^{k-1}$ except in the following cases*:

(a) *If $\Delta = 4(4m+1)$ then $h^s_\Delta = 2^{k-2}$.*

(b) *If $\Delta = 32m$ then $h^s_\Delta = 2^k$.*

The exceptional cases (a) and (b) involve nonfundamental discriminants, so for fundamental discriminants we have $h^s_\Delta = 2^{k-1}$. For example, the discriminants $\Delta = 60 = 3\cdot 4\cdot 5$ and $\Delta = -260 = -4\cdot 5\cdot 13$ that we looked at in the previous section have three distinct prime divisors so the theorem says there are $2^2 = 4$ equivalence classes of mirror symmetric forms in these two cases. This agrees with what the topographs showed.

The proof of the theorem will involve considering a number of different cases. Fortunately most of the resulting complication disappears in the final answer.

Proof: By Proposition 5.6 every form with mirror symmetry is equivalent to a form $ax^2 + cy^2$ or $ax^2 + axy + cy^2$. The strategy will be to count how many of these special forms there are that are primitive with discriminant Δ, then determine which of these special forms are equivalent.

For counting the special forms ax^2+cy^2 and $ax^2+axy+cy^2$ we may assume $a > 0$ since a is the value of the form when $(x,y) = (1,0)$ and for elliptic forms we only consider those with positive values, while for hyperbolic forms we are free to change a form to its negative so it suffices to count only those with $a > 0$ and then double the result.

Case 1: Forms ax^2+cy^2. Then $\Delta = -4ac = 4\delta$ for $\delta = -ac$. Primitivity of the form is equivalent to a and c being coprime. The only way to have coprime factors a and c of $\delta = -ac$ is to take an arbitrary subset of the distinct primes dividing δ and let a be the product of these primes each raised to the same power as in δ (so $a = 1$ when we choose the empty subset). The number of such subsets is $2^{k'}$ where k' is the number of distinct prime divisors of δ, so there are $2^{k'}$ primitive forms ax^2+cy^2 with $a>0$.

Case 2: Forms $ax^2+axy+cy^2$ with Δ odd. We have $\Delta = a^2-4ac$ so Δ and a have the same parity. From $\Delta = a(a-4c)$ we see that a divides Δ. We claim that each divisor a of Δ gives rise to a form $ax^2+axy+cy^2$ of discriminant Δ. Solving $\Delta = a^2-4ac$ for c gives $c = (a^2-\Delta)/4a$. The numerator is divisible by 4 since a and Δ are odd and hence a^2 and Δ are both 1 mod 4, making the numerator 0 mod 4. The numerator is also divisible by a if a divides Δ. Since 4 and a are coprime when a is odd, it follows that $4a$ divides the numerator so c is an integer and we get a form $ax^2+axy+cy^2$ of discriminant Δ. This form is primitive exactly when a and c are coprime. This is equivalent to saying that the two factors of $\Delta = a(a-4c)$ are coprime since any divisor of a and c must divide the two factors, and conversely any divisor of the two factors must divide a and $4c$, hence also c since this divisor of the odd number a must be odd. As in Case 1, the only way to obtain a factorization $\Delta = a(a-4c)$ with the two factors coprime is to take an arbitrary subset of the distinct primes dividing Δ and let a be the product of these primes each raised to the same power as in Δ. The number of such subsets is 2^k so this is the number of primitive forms $ax^2+axy+cy^2$ with $a>0$ when Δ is odd.

There remain the forms $ax^2+axy+cy^2$ with $\Delta = 4\delta$. Again Δ and a have the same parity since $\Delta = a^2-4ac$, so a is even, say $a = 2d$. From $\Delta = a^2-4ac$ we then have $\delta = d^2-2dc = d(d-2c)$.

Case 3: Forms $ax^2+axy+cy^2$ with $\Delta = 4\delta$ and $a = 2d$ for odd d. By primitivity c must be odd. The two factors of $\delta = d(d-2c)$ are odd and must be distinct mod 4 since c is odd. Thus one factor is 1 mod 4 and the other is 3 mod 4, so $\delta \equiv 3 \bmod 4$, say $\delta = 4m+3$. We claim that when $\delta = 4m+3$, each divisor d of δ gives rise to a form $ax^2+axy+cy^2$ with $a = 2d$. To show this, note first that d must be odd since it divides δ which is odd. Solving $\delta = d(d-2c)$ for c gives $c = (d^2-\delta)/2d$. Since d and δ are odd, the numerator $d^2-\delta$ is even hence divisible by the 2 in the denominator. The numerator is also divisible by the d in the denominator if d divides δ. Since d is odd, this implies that $2d$ divides the numerator, so c is an integer for each divisor d of δ. In fact c is an odd integer since the numerator $d^2-\delta$ is 2 mod 4 and so $cd = (d^2-\delta)/2$ is odd, forcing c to be odd. For the form $ax^2+axy+cy^2$ to be primitive means that a and c are coprime. Since c is odd and $a = 2d$ this is equivalent to c and d being coprime. This in turn is equivalent to the two factors of $\delta = d(d-2c)$ being coprime since c and d are odd. Thus when $\delta = 4m+3$ we get a primitive form $ax^2+axy+cy^2$ for each choice of a subset of the distinct prime divisors of δ since this determines d as before, and d determines c and a. The number of primitive forms $ax^2+axy+cy^2$ is then $2^{k'}$ when Δ is even and $a = 2d$ with d odd, where k' is the number of distinct prime divisors of δ as in Case 1.

Case 4: Forms $ax^2 + axy + cy^2$ with Δ even and $a = 2d$ for even d, say $d = 2e$. Then $\delta = d(d - 2c) = 4e(e - c)$. Since c is odd by primitivity of the form, the two factors e and $e - c$ of $\delta = 4e(e - c)$ have opposite parity, hence δ must be divisible by 8, say $\delta = 8m$. We need to determine which choices of e and c yield primitive forms $ax^2 + axy + cy^2$. Let $\delta' = \delta/4 = 2m$ so the equation $\delta = 4e(e - c)$ becomes $\delta' = e(e - c)$. Thus e must divide δ'. We have $c = e - \delta'/e$ and this will be an integer if e divides δ'. From the equation $c = e - \delta'/e$ we see that any divisor of two of the three terms c, e, and δ'/e will divide the third. In particular, c and e will be coprime exactly when e and δ'/e are coprime. Since $\delta' = e \cdot \delta'/e$ this means we want to choose e by choosing some subset of the distinct prime divisors of δ' and letting e be the product of these primes raised to the same powers as in δ'. Then e and δ'/e will be coprime and of opposite parity since they are not both even and their product δ' is even. Their difference $c = e - \delta'/e$ will then be odd. Also, c and e will be coprime so c and $a = 4e$ will be coprime, making the form $ax^2 + axy + cy^2$ primitive. The number of distinct prime divisors of δ' is the same as for $\delta = 4\delta'$ since δ' is even. Thus in Case 4 the number of primitive forms $ax^2 + axy + cy^2$ with $a > 0$ is $2^{k'}$.

Note that $k' = k$ when δ is even and $k' = k - 1$ when δ is odd. By combining the four cases above and remembering to double the number of forms when $\Delta > 0$ to account for negative coefficients of x^2, we then obtain the following table for the number of forms of either of the types $ax^2 + cy^2$ or $ax^2 + axy + cy^2$:

Δ	odd	$4\delta,\ \delta = 4m+1$	$4\delta,\ \delta = 4m+3$
Cases	(2)	(1)	(1) and (3)
$\Delta < 0$	2^k	$2^{k'} = 2^{k-1}$	$2^{k'} + 2^{k'} = 2^{k'+1} = 2^k$
$\Delta > 0$	2^{k+1}	$2^{k'+1} = 2^k$	$2^{k'+1} + 2^{k'+1} = 2^{k'+2} = 2^{k+1}$

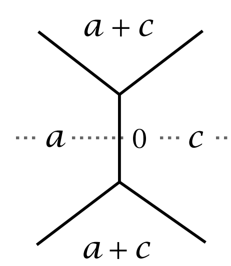
Δ	$4\delta,\ \delta = 8m$	$4\delta,\ \delta$ even, $\delta \neq 8m$
Cases	(1) and (4)	(1)
$\Delta < 0$	$2^{k'} + 2^{k'} = 2^{k'+1} = 2^{k+1}$	$2^{k'} = 2^k$
$\Delta > 0$	$2^{k'+1} + 2^{k'+1} = 2^{k'+2} = 2^{k+2}$	$2^{k'+1} = 2^{k+1}$

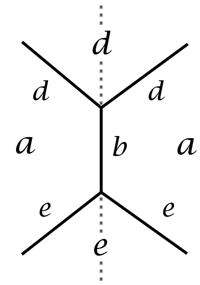
Comparing the results in the table with the statement of the theorem, we see that the proof will be finished when we show that under the relation of equivalence the special forms split up into pairs when $\Delta < 0$ and into groups of four when $\Delta > 0$.

Two easy cases that can be disposed of first are $\Delta = -3$ and $\Delta = -4$. Here all forms are equivalent and are primitive, and $k = 1$, so the theorem is true since the exceptional cases (a) and (b) in the statement of the theorem do not apply.

Our earlier analysis of symmetries of elliptic and hyperbolic forms shows that the only time that reflector lines can intersect is for elliptic forms equivalent to $ax^2 + ay^2$ or $ax^2 + axy + ay^2$, so when we restrict to primitive forms this means $\Delta = -3$ or $\Delta = -4$. Thus we may assume from now on that reflector lines do not intersect.

For a form $ax^2 + cy^2$ with a reflector line perpendicular to an edge of the topograph as in the first figure at the right we have $a \neq c$, otherwise there would be two intersecting reflector lines. Thus the reflector line corresponds to two distinct special forms, $ax^2 + cy^2$ and $cx^2 + ay^2$.



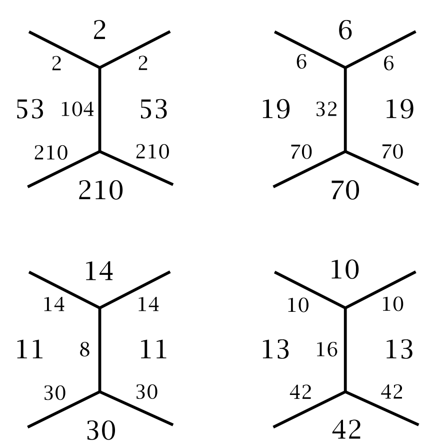


The second figure shows the case of a form with a reflector line containing an edge of the topograph. This edge corresponds to a form $ax^2 + bxy + ay^2$ and the adjacent edges correspond to two forms $dx^2 + dxy + ay^2$ and $ex^2 + exy + ay^2$ of the type $ax^2 + axy + cy^2$. These two forms are distinct since if $d = e$ there would be a second reflector line intersecting the first one. Thus the reflector line accounts for two special forms $ax^2 + axy + cy^2$.

Primitive elliptic forms with mirror symmetry and $\Delta \neq -3, -4$ have just one reflector line, so each equivalence class of such forms contains exactly two special forms. For hyperbolic forms with mirror symmetry there are two reflector lines in each period, with one pair of special forms for each reflector line. These two pairs give four distinct special forms, otherwise there would be a translational symmetry taking one reflector line to the other within a single period, which is impossible. Thus each equivalence class of mirror-symmetric hyperbolic forms contains exactly four special forms, and the proof is complete. □

We illustrate the theorem with an example, the first negative discriminant with four distinct prime divisors, $\Delta = -420 = -3 \cdot 4 \cdot 5 \cdot 7$. In this case $\Delta = 4(4m + 3)$ so the theorem says there are $2^3 = 8$ equivalence classes of symmetric primitive forms. If we compute all the reduced forms for $\Delta = -420$ by the method in Section 5.2 we get the following table, with the letter b replacing h so we are finding solutions of $b^2 + 420 = 4ac$ with $0 \le b \le a \le c$. The entries $[a, b, c]$ in the last column give the reduced forms $ax^2 + bxy + cy^2$.

b	ac	(a, c)	$[a, b, c]$
0	105	$(1, 105)$	$[1, 0, 105]$
		$(3, 35)$	$[3, 0, 35]$
		$(5, 21)$	$[5, 0, 21]$
		$(7, 15)$	$[7, 0, 15]$
2	106	$(2, 53)$	$[2, 2, 53]$
4	109	—	
6	114	$(6, 19)$	$[6, 6, 19]$
8	121	$(11, 11)$	$[11, 8, 11]$
10	130	$(10, 13)$	$[10, 10, 13]$



Thus all forms of discriminant -420 are symmetric. The first four have $b = 0$ so these arise in Case 1 in the proof of the theorem where we set $\Delta = 4\delta$, so $\delta = -3 \cdot 5 \cdot 7$ and we get a form $[a, 0, c]$ for each positive divisor a of δ, the eight numbers

$1, 3, 5, 7, 15, 21, 35$, and 105. These forms $[a, 0, c]$ are the first four entries in the last column of the table along with the equivalent forms obtained by reversing a and c. The remaining four forms in the last column have b nonzero and are instances of forms $[a, a, c]$ and $[a, b, a]$. The relevant parts of the topographs of these four forms are shown in the figure to the right of the table. Each edge in the figure gives a form $[a, b, a]$, $[a, a, c]$, or $[a, c, c]$. For example the third figure gives the forms $[11, 8, 11]$, $[11, 14, 14]$, $[14, 14, 11]$, $[11, 30, 30]$, and $[30, 30, 11]$. In the proof of the theorem we were only counting the forms $[a, a, c]$, not $[a, b, a]$ or $[a, c, c]$. According to Case 3 in the proof of the theorem the numbers a in the forms $[a, a, c]$ should be twice the numbers a in the forms $[a, 0, c]$, and they are: $2 = 2\cdot 1$, $6 = 2\cdot 3$, $10 = 2\cdot 5$, $14 = 2\cdot 7$, $30 = 2\cdot 15$, $42 = 2\cdot 21$, $70 = 2\cdot 35$, and $210 = 2\cdot 105$.

Corollary 5.10. *The nonsquare discriminants Δ with $h^s_\Delta = 1$ are $\Delta = -4, \pm 8, -16, \pm p^{2k+1}$, and $\pm 4p^{2k+1}$ for odd primes p with $p \equiv 1$ mod 4 when $\Delta > 0$ and $p \equiv 3$ mod 4 when $\Delta < 0$. In particular, the only fundamental discriminants with $h^s_\Delta = 1$ are $\Delta = -4, \pm 8$, and $\pm p$ for odd primes p, with $p \equiv 1$ mod 4 when $\Delta > 0$ and $p \equiv 3$ mod 4 when $\Delta < 0$.*

Proof: Consider first the case $\Delta > 0$. If we are not in one of the exceptional cases (a) and (b) in Theorem 5.9 then Δ must have just one distinct prime divisor so it must be a power of a prime, in fact an odd power if it is not a square. Thus for p odd we have $\Delta = p^{2k+1}$ and we must have $p \equiv 1$ mod 4 in order to have $\Delta \equiv 1$ mod 4. For odd powers of $p = 2$ the only possibility is $\Delta = 8$ since Δ cannot be 2 and odd powers beyond 8 are of the form $\Delta = 32m$, the exceptional case (b) where $h^s_\Delta \ge 2$ so this is ruled out as well. In the exceptional case (a) we have $\Delta = 4(4m + 1)$ with $4m + 1$ a prime power p^{2k+1} with $p \equiv 1$ mod 4 since $\Delta = 4p^{2k}$ is a square.

When $\Delta < 0$ the reasoning is similar, the main difference being that $-p^{2k}$ and $-4p^{2k}$ are ruled out, not because squares are excluded, but because p^{2k} is always 1 mod 4 when p is odd, so $-p^{2k}$ is 3 mod 4. This rules out $-p^{2k}$ as a discriminant, and it rules out $-4p^{2k}$ being an exceptional case $\Delta = 4(4m + 1)$.

Requiring Δ to be a fundamental discriminant eliminates the cases $\Delta = -16$ and $\pm 4p^{2k+1}$ and restricts the exponent in $\pm p^{2k+1}$ to be 1. $\square$

We have mentioned the fact that h_Δ is always a multiple of h^s_Δ, which will be proved in Proposition 7.17. This tells us nothing about h_Δ when $h^s_\Delta = 1$, but we will also prove that $h^s_\Delta = 1$ exactly when h_Δ is odd. Thus the preceding corollary gives a way to determine whether h_Δ is even or odd. In the examples we have looked at so far h_Δ has been either 1 or even, but odd numbers greater than 1 can also occur as class numbers. The table on the next page gives some examples for negative discriminants, so we are determining the reduced forms $ax^2 + bxy + cy^2$ by finding the solutions of $b^2 + |\Delta| = 4ac$ with $0 \le b \le a \le c$. The forms other than the principal form in each discriminant lack mirror symmetry so they count twice in the class number, making the class number odd. The discriminants in the table are all fundamental discriminants, and in each case they are the first negative discriminant with the given class number.

Δ	b	ac	(a,c)	h_Δ
-23	1	6	$(1,6),(2,3)$	3
-47	1	12	$(1,12),(2,6),(3,4)$	5
	3	14	—	
-71	1	18	$(1,18),(2,9),(3,6)$	7
	3	20	$(4,5)$	
-199	1	50	$(1,50),(2,25),(5,10)$	9
	3	52	$(4,13)$	
	5	56	$(7,8)$	
	7	62	—	
-167	1	42	$(1,42),(2,21),(3,14),(6,7)$	11
	3	44	$(4,11)$	
	5	48	$(6,8)$	
	7	54	—	
-191	1	48	$(1,48),(2,24),(3,16),(4,12),(6,8)$	13
	3	50	$(5,10)$	
	5	54	$(6,9)$	
	7	60	—	
-239	1	60	$(1,60),(2,30),(3,20),(4,15),(5,12),(6,10)$	15
	3	62	—	
	5	66	$(6,11)$	
	7	72	$(8,9)$	

Besides the cases when $h^s_\Delta = 1$, another nice situation is when $h_\Delta = h^s_\Delta$ so all primitive forms of discriminant Δ have mirror symmetry. We call such discriminants ***fully symmetric***. As we will see in later chapters, forms with fully symmetric discriminants have very special properties. A table at the end of the book lists the 101 known negative discriminants that are fully symmetric, ranging from -3 to -7392.

Of the 101 known fully symmetric negative discriminants, 65 are fundamental discriminants, the largest being -5460. Since 5460 factors as $3\cdot 4\cdot 5\cdot 7\cdot 13$ with five distinct prime factors, Theorem 5.9 says that $h^s_\Delta = 2^4 = 16$. This is in fact the largest value of h^s_Δ among the 101 discriminants in the list. Computer calculations have extended to much larger negative discriminants without finding any more that are fully symmetric. It has not yet been proved that no more exist, although it is known that there are at most two more. For positive discriminants there are probably infinitely many that are fully symmetric since it is likely that there are already infinitely many with $h_\Delta = 1$.

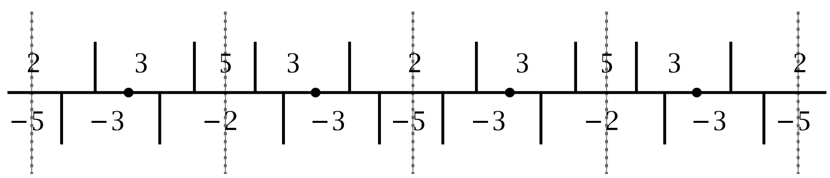
Skew Symmetries

Among the examples of hyperbolic forms we have considered there were some whose topograph had a “symmetry” which was a glide reflection along the separator line that had the effect of changing each value to its negative rather than preserving the values. These are not actual symmetries according to the definition we have given, so let us call such a transformation that takes each value of a form to its negative a ***skew symmetry***. (Compare this with skew-symmetric matrices in linear algebra which equal the negative of their transpose.)

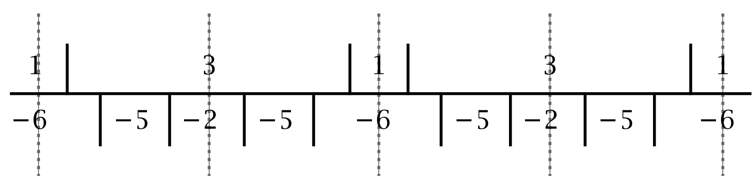
A skew symmetry must take the separator line to itself while interchanging the two sides of the separator line, so it either translates the separator line along itself and hence is a glide reflection, or it reflects the separator line, interchanging its two ends as well as the two sides of the separator line, making it a 180 degree rotation about a point of the separator line. Examples of forms with this sort of skew symmetry occurred in Chapter 4, the forms $x^2 - 13y^2$ and $10x^2 - 29y^2$.

The figures below show forms whose separator lines have all the possible combinations of symmetries and skew symmetries.

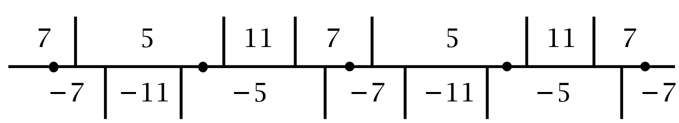
$2x^2 - 5y^2$



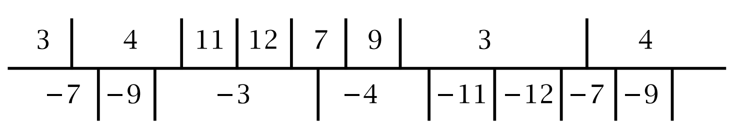
$x^2 - 6y^2$



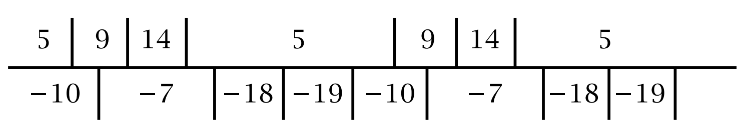
$7x^2 + 5xy - 7y^2$



$3x^2 + 8xy - 7y^2$



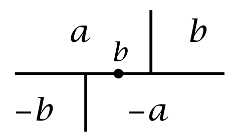
$5x^2 + 14xy - 10y^2$



The first form has all four types: translations, mirror symmetries, glide reflections, and rotations. The next three forms have only one type of symmetry or skew symmetry besides translations, while the last form has only translational symmetries and no mirror symmetries or skew symmetries. It is not possible to have two of the three types of nontranslational symmetries and skew symmetries without having the third since the composition of two of these three types gives the third type. One can see this by considering the effect of a symmetry or skew symmetry on the orientation of the plane and the orientation of the separator line. The four possible combinations distinguish the four types of transformations according to the following chart, where a plus sign means orientation-preserving and a minus sign means orientation-reversing.

	plane orientation	line orientation
translation	+	+
rotation	+	−
glide reflection	−	+
reflection	−	−

A rotational skew symmetry is a rotation about the midpoint of an edge of the separator line where the two adjacent regions have labels a and $-a$. If the edge separating these two regions has label b then the form associated to this edge is $ax^2 + bxy - ay^2$. Conversely, any form $ax^2 + bxy - ay^2$ whose discriminant $\Delta = b^2 + 4a^2$ is not a square (although it is the sum of two squares) will be a hyperbolic form having a rotational skew symmetry, as one can see in the figure at the right. Note that the form $ax^2 + bxy - ay^2$ will be one of the reduced forms in the equivalence class of the given form since the two edges leading off the separator line at the ends of the edge labeled b do so on opposite sides of the separator line. Thus rotational skew symmetries can be detected by looking just at the reduced forms. The same is true for mirror symmetries and glide reflection skew symmetries, but for these one must look at the arrangement of the whole cycle of reduced forms rather than just the individual reduced forms.



For rotational skew symmetries there are two rotation points along the separator line in each period, just as reflector lines occur in pairs in each period.

Exercises

1. Show that the number of symmetries of an elliptic form, including the identity transformation, is 1, 2, 4, or 6.

2. Show that the number of equivalence classes of forms of discriminant 45 with mirror symmetry is not a power of 2 if nonprimitive as well as primitive forms are allowed. (Compare this with Theorem 5.9.)

3. In the text an example was given of a hyperbolic form having only translational symmetries and no skew symmetries, the form $5x^2 + 14xy - 10y^2$. Find another example of the same sort which is not equivalent to this form or a constant times it. *Hint*: First find a separator line with the desired properties, without any labels along the line, then find a form realizing that separator line.

4. Show that a positive nonsquare number is the discriminant of some hyperbolic form whose topograph has a rotational skew symmetry if and only if the number is the sum of two squares at least one of which is even.

5. Verify that the following discriminants are fully symmetric, so all primitive forms of that discriminant have mirror symmetry:

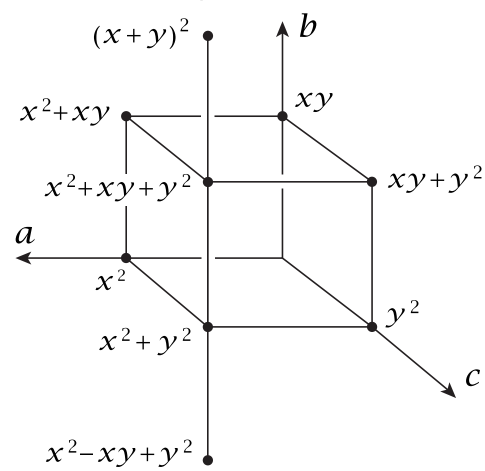
(a) -195 (b) -660 (c) 195

6. Show that the topograph of a primitive 0-hyperbolic form $qxy - py^2$ has mirror symmetry exactly when $p^2 \equiv 1 \bmod q$, and has rotational skew symmetry exactly when $p^2 \equiv -1 \bmod q$. (See the discussion at the end of Secion 2.1 about the relation between the continued fraction for p/q and the continued fraction obtained by reversing the order of the terms.)

5.5 Charting All Forms

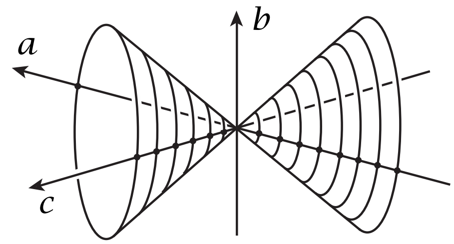
We have used the Farey diagram to study individual quadratic forms through their topographs, and in this section we will see that the Farey diagram also appears in another way when one creates a global picture mapping out all forms simultaneously. This viewpoint will not play an essential role in later chapters, however, so this section can be regarded as something of a digression from the main line of the book.

Quadratic forms are defined by formulas $ax^2+bxy+cy^2$, and our point of view will be to regard the coefficients a, b, and c as parameters that vary over all integers independently. It is natural to consider the triples (a,b,c) as points in 3-dimensional Euclidean space $\mathbb{R}^3$, and more specifically as points in the integer lattice $\mathbb{Z}^3$ consisting of points (a,b,c) whose coordinates are integers. We will exclude the origin $(0,0,0)$ since this corresponds to the trivial form that is identically zero. Instead of using the usual (x,y,z) as coordinates for $\mathbb{R}^3$ we will use (a,b,c), but since a and c play a symmetric role as the coefficients of the squared terms x^2 and y^2 we will position the a-axis and the c-axis in a horizontal plane, with the b-axis vertical, perpendicular to the ac-plane.



Along a ray starting at $(0,0,0)$ and passing through another lattice point (a,b,c) there are infinitely many lattice points (ka,kb,kc) for positive integers k. If a, b, and c have a greatest common divisor larger than 1 we can cancel this common divisor to get a primitive triple (a,b,c) corresponding to a primitive form $ax^2+bxy+cy^2$. Then all the other lattice points on the ray through (a,b,c) are the positive integer multiples (ka,kb,kc), corresponding to the nonprimitive forms $kax^2+kbxy+kcy^2$. Thus primitive forms correspond exactly to rays from the origin passing through lattice points. These are the same as rays passing through points (a,b,c) with rational coordinates since denominators can always be eliminated by multiplying a, b, and c by a common denominator.

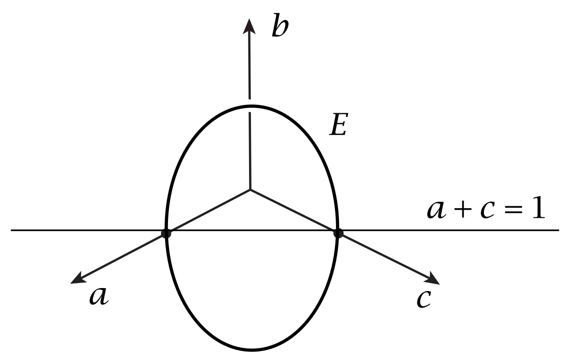
Since the discriminant $\Delta = b^2 - 4ac$ plays such an important role in the classification of forms, let us see how this fits into the picture in (a, b, c) coordinates. When $b^2 - 4ac$ is zero we have the special class of parabolic forms, and the points in $\mathbb{R}^3$ satisfying the equation $b^2 - 4ac = 0$ form a double cone with the common vertex of the two cones at the origin. The double cone intersects the ac-plane in the a-axis and the c-axis. The central axis of the double cone is the line $a = c$ in the ac-plane. Parabolic forms are the lattice points on these cones.



Elliptic and Parabolic Forms

Points (a, b, c) inside either cone have $b^2 - 4ac < 0$ so the lattice points inside the cones correspond to elliptic forms. Positive elliptic forms have $a > 0$ and $c > 0$ so they lie inside the cone projecting to the first quadrant of the ac-plane. We call this the *positive cone*. Inside the other cone are the negative elliptic forms, those with $a < 0$ and $c < 0$. Outside the cones is a single region consisting of points with $b^2 - 4ac > 0$ so the lattice points here correspond to hyperbolic forms and 0-hyperbolic forms.

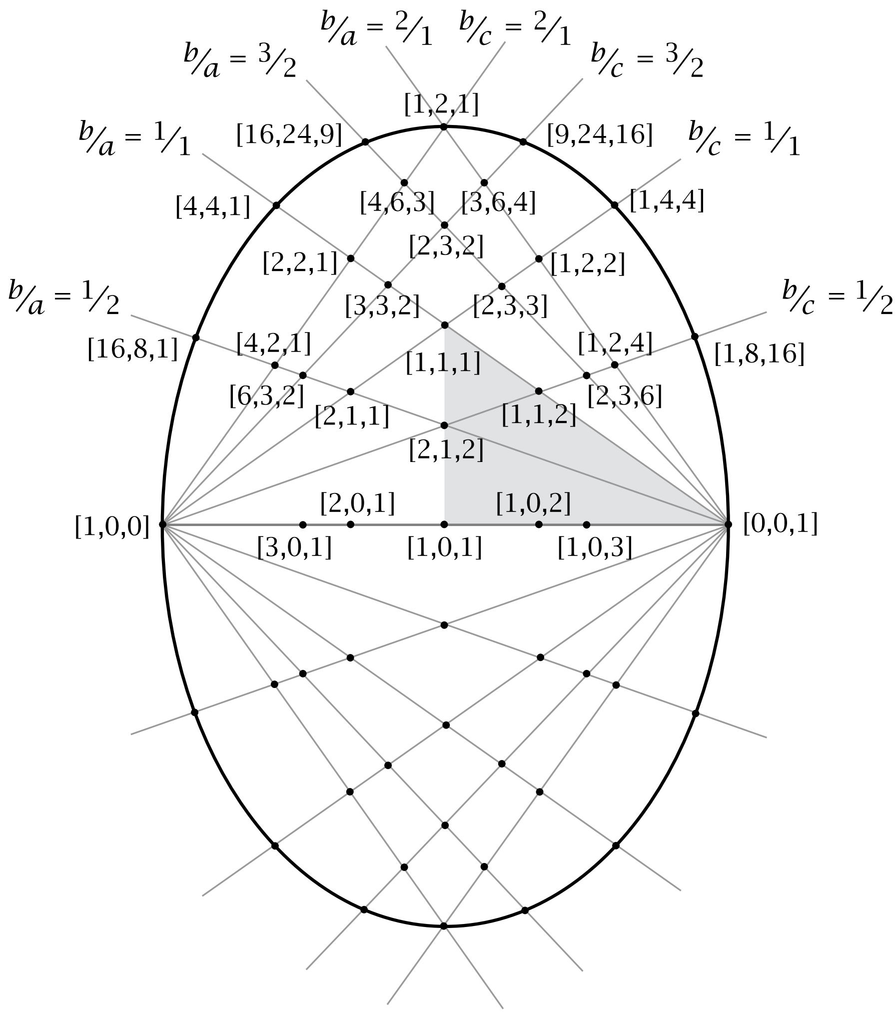
If one slices the positive cone via the vertical plane $a + c = 1$ perpendicular to the axis of the cone then the intersection of the cone with this plane is an ellipse which we denote E. The top and bottom points of E are $(a, b, c) = (1/2, \pm 1, 1/2)$ so its height is 2. The left and right points of E are $(1, 0, 0)$ and $(0, 0, 1)$ so its width is $\sqrt{2}$. Thus E is somewhat elongated vertically. If we wanted, we could compress the vertical coordinate to make E a circle, but there is no special advantage to doing this.



If we take a lattice point (a, b, c) corresponding to a primitive positive elliptic form and project this lattice point along the ray to the origin passing through (a, b, c), this ray intersects the plane $a + c = 1$ in the point $(a/_{a+c}, b/_{a+c}, c/_{a+c})$ since this is the rescaling of (a, b, c) for which the sum of the first and third coordinates is 1. This point lies inside the ellipse E and has rational coordinates. Conversely, every point inside E with rational coordinates is the radial projection of a unique primitive positive elliptic form, obtained by multiplying the coordinates of the point by the least common multiple of their denominators. Thus the rational points inside E parametrize primitive positive elliptic forms. We will use the notation $[a, b, c]$ to denote both the form $ax^2 + bxy + cy^2$ and the corresponding rational point $(a/_{a+c}, b/_{a+c}, c/_{a+c})$

inside E.

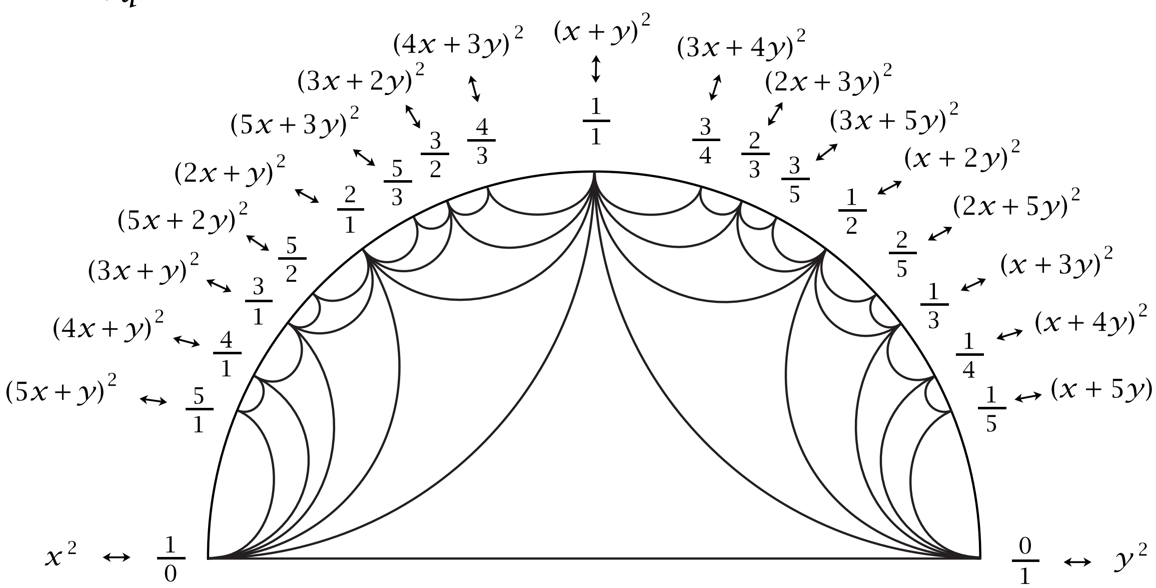
The figure below shows some examples, including a few parabolic forms on E itself. The lines radiating out from the points $[1,0,0]$ and $[0,0,1]$ consist of the points $[a,b,c]$ with a fixed ratio b/c or b/a respectively. The ratios a/c are fixed along vertical lines. For most points inside E any two out of these three ratios determine the third since $b/a \cdot a/c = b/c$. The exceptions are the points on the segment between $[1,0,0]$ and $[0,0,1]$ where b/a and b/c are both 0 but a/c can be anything.



Of special interest are the reduced primitive elliptic forms $[a,b,c]$, which are the ones satisfying $0 \le b \le a \le c$ where a, b, and c have no common divisor. These correspond to the rational points in the shaded triangle in the figure with vertices $[1,1,1]$, $[1,0,1]$, and $[0,0,1]$. The edges of the triangle correspond to one of the three inequalities $0 \le b \le a \le c$ becoming an equality, so $b = 0$ for the lower edge, $a = c$ for the vertical edge, and $a = b$ for the hypotenuse. Thus the three edges correspond to the reduced forms with mirror symmetry, the forms $[a,0,c]$ for the bottom edge, $[a,b,a]$ for the left edge, and $[a,a,c]$ for the diagonal edge. The vertices $[1,0,1]$ and $[1,1,1]$ correspond to the reduced elliptic forms with more than one mirror symmetry, and hence also rotational symmetry. Points in the interior of the triangle correspond to forms with no symmetry.

Just as rational points inside the ellipse E correspond to primitive positive elliptic forms, the rational points on E itself correspond to primitive positive parabolic forms. As we saw in Section 5.2, every parabolic form is equivalent to a form ax^2 for some nonzero integer a. For this to be primitive means that $a = \pm 1$, so every positive primitive parabolic form is equivalent to x^2. Equivalent forms can be obtained from each other by a change of variables, replacing (x, y) by $(px + qy, rx + sy)$ for integers p, q, r, s satisfying $ps - qr = \pm 1$. For the form x^2 this means that the primitive positive parabolic forms are the forms $(px+qy)^2 = p^2x^2 + 2pqxy + q^2y^2$ for coprime integers p and q. In $[a, b, c]$ notation this is $[p^2, 2pq, q^2]$, defining a point on the ellipse E.

More concisely, we could label the rational point on E corresponding to the form $(px + qy)^2$ just by the fraction p/q. Thus at the left and right sides of E we have the fractions $1/0$ and $0/1$ corresponding to the forms x^2 and y^2, while at the top and bottom of E we have $1/1$ and $-1/1$ corresponding to $(x+y)^2$ and $(x-y)^2 = (-x+y)^2$. Changing the signs of both p and q does not change the form $(px + qy)^2$ or the fraction p/q.



In the first quadrant of the ellipse the fractions p/q increase monotonically from $0/1$ to $1/1$ since the ratio b/c equals $2p/q$ and b is increasing while c is decreasing so $2p/q$ is increasing, and hence so is p/q. Similarly in the second quadrant the values of p/q increase from $1/1$ to $1/0$ since we have $b/a = 2q/p$ which decreases as b decreases and a increases. In the lower half of the ellipse we have just the negatives of the values in the upper half since the sign of b has changed from plus to minus.

This labeling of the rational points of E by fractions p/q seems very similar to the labeling of vertices in the circular Farey diagram. As we saw in Section 1.1, if the Farey diagram is drawn with $1/0$ at the top of the unit circle in the xy-plane, then the point on the unit circle labeled p/q has coordinates $(x, y) = (2pq/p^2+q^2, p^2-q^2/p^2+q^2)$. After rotating the circle to put $1/0$ on the left side by replacing (x, y) by $(-y, x)$

this becomes $({}^{q^2-p^2}\!/\!_{p^2+q^2}, {}^{2pq}\!/\!_{p^2+q^2})$. Here the y-coordinate ${}^{2pq}\!/\!_{p^2+q^2}$ is the same as the b-coordinate of the point of E labeled ${}^{p}\!/\!_{q}$, which is the point $(a,b,c) = ({}^{p^2}\!/\!_{p^2+q^2}, {}^{2pq}\!/\!_{p^2+q^2}, {}^{q^2}\!/\!_{p^2+q^2})$. Since the vertical coordinates of points in either the left or right half of the unit circle or the ellipse E determine the horizontal coordinates uniquely, this means that the labeling of points of E by fractions ${}^{p}\!/\!_{q}$ is really the same as in the circular Farey diagram.

Change of Variables

Let us return now to the general picture of how forms $ax^2 + bxy + cy^2$ are represented by points (a,b,c) in $\mathbb{R}^3$. As we know, a change of variables by a linear transformation T sends (x,y) to $T(x,y) = (px+qy, rx+sy)$, where p,q,r,s are integers with $ps - qr = \pm 1$. This change of variables transforms each form into an equivalent form. To see the effect of this change of variables on the coefficients (a,b,c) of a form $Q(x,y) = ax^2 + bxy + cy^2$ we do a simple calculation:

$$\begin{aligned} Q(px+qy, rx+sy) &= a(px+qy)^2 + b(px+qy)(rx+sy) + c(rx+sy)^2 \\ &= (ap^2 + bpr + cr^2)x^2 + (2apq + bps + bqr + 2crs)xy \\ &\qquad + (aq^2 + bqs + cs^2)y^2 \end{aligned}$$

This means that the (a,b,c) coordinates of points in $\mathbb{R}^3$ are transformed according to the following formula:

$$T^*(a,b,c) = (p^2a + prb + r^2c, 2pqa + (ps+qr)b + 2rsc, q^2a + qsb + s^2c)$$

For fixed values of p,q,r,s this T^* is a linear transformation of the variables a,b,c. Its matrix is:

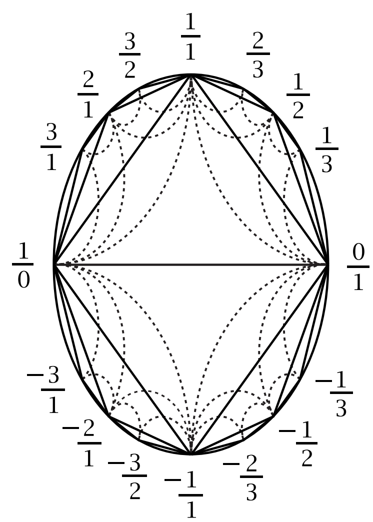
$$\begin{pmatrix} p^2 & pr & r^2 \\ 2pq & ps+qr & 2rs \\ q^2 & qs & s^2 \end{pmatrix}$$

Since T^* is a linear transformation, it takes lines to lines and planes to planes, but T^* also has another special geometric property. Since equivalent forms have the same discriminant, this means that each surface defined by an equation $b^2 - 4ac = k$ for k a constant is taken to itself by T^*. In particular, the double cone $b^2 - 4ac = 0$ is taken to itself, and in fact each of the two cones separately is taken to itself since one cone consists of positive parabolic forms and the other cone of negative parabolic forms (as one can see just by looking at the coefficients a and c), and positive parabolic forms are never equivalent to negative parabolic forms. When $k > 0$ the surface $b^2 - 4ac = k$ is a hyperboloid of one sheet and when $k < 0$ it is a hyperboloid of two sheets. In the case of two sheets the lattice points on one sheet give positive elliptic forms and the lattice points on the other sheet give negative elliptic forms.

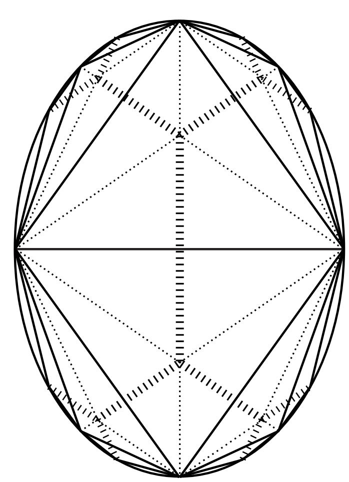
Since T^* takes lines through the origin to lines through the origin and the double cone $b^2 - 4ac = 0$ to itself, this means that T^* gives a transformation of the ellipse E

to itself, taking rational points to rational points since rational points on E correspond to lattice points on the cones. Regarding E as the boundary circle of the Farey diagram, we know that linear fractional transformations give symmetries of the Farey diagram, also taking rational points on the boundary circle to rational boundary points. And in fact, the transformation of this circle defined by T^* is exactly one of these linear fractional transformations. This is because T^* takes the parabolic form $(dx+ey)^2$ to the form $(d(px+qy)+e(rx+sy))^2 = ((dp+er)x+(dq+es)y)^2$ so in the fractional labeling of points of E this says $T^*({}^{d}/_{e}) = {}^{pd+re}/_{qd+se}$ which is a linear fractional transformation. If we write this using the variables x and y instead of d and e it would be $T^*({}^{x}/_{y}) = {}^{px+ry}/_{qx+sy}$. This is not quite the same as the linear fractional transformation $T({}^{x}/_{y}) = {}^{px+qy}/_{rx+sy}$ defined by the original change of variables $T(x,y) = (px+qy, rx+sy)$, but rather T^* is obtained from T by transposing the matrix of T, interchanging the off-diagonal terms q and r.

Via radial projection, the transformation T^* determines a transformation not just of E but also of the interior of E in the plane $a+c=1$. This transformation, which we still call T^* for simplicity, takes lines inside E to lines inside E since T^* takes planes through the origin to planes through the origin. This leads us to consider a linear version of the Farey diagram in which each circular arc of the original Farey diagram is replaced by a straight line segment joining the two endpoints of the circular arc. These line segments divide the interior of E into triangles, just as the original Farey diagram divides the disk into curvilinear triangles. The transformation T^* takes each of these triangles onto another triangle, analogous to the way that linear fractional transformations provide symmetries of the original Farey diagram.



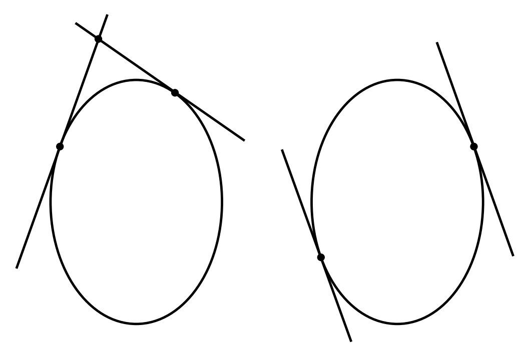
Suppose we divide each triangle of the linear Farey diagram into six smaller triangles as in the figure at the right, by adding diagonals to each quadrilateral formed by two adjacent triangles of the Farey diagram. The transformation T^* takes each of these small triangles onto another small triangle since it takes lines to lines. One of these small triangles is the triangle defined by the inequalities $0 \le b \le a \le c$ that we considered earlier. The fact that every positive primitive elliptic form is equivalent to exactly one reduced form, corresponding to a rational point in this special triangle, is now visible geometrically as the fact that there is always exactly one transformation T^* taking a given small triangle to this one special small triangle.



Elliptic forms whose topograph contains a source edge are equivalent to forms $ax^2 + cy^2$ so these are the forms corresponding to rational points on the edges of the original linear Farey diagram, before the subdivision into smaller triangles. These are the forms whose topograph has a symmetry reflecting across a line perpendicular to the source edge. (This line is just the edge in the Farey diagram containing the given form.) The other type of reflectional symmetry in the topograph of an elliptic form is reflection across an edge of the topograph. Forms with this sort of symmetry correspond to rational points in the dotted edges in the preceding figure, the edges we added to subdivide the Farey diagram into the smaller triangles. The dotted edges are of two types according to whether the two equal values of the form in the three regions surrounding the source vertex occur for the smallest value of the form (wide dotted edges) or the next-to-smallest value of the form (narrow dotted edges). The wide dotted edges form the dual tree of the Farey diagram.

Hyperbolic and 0-Hyperbolic Forms

Hyperbolic and 0-hyperbolic forms correspond to integer lattice points that lie outside the two cones. For a point (a, b, c) outside the double cone there are exactly two planes in $\mathbb{R}^3$ that are tangent to the double cone and pass through (a, b, c). Each of these planes is tangent to the double cone along a line through the origin. The two tangent planes through (a, b, c) are determined by their intersection with the plane $a + c = 1$, which consists of two lines tangent to the ellipse E. These two lines can either intersect or be parallel. The latter possibility occurs when the point (a, b, c) lies in the plane $a + c = 0$, so the two tangent planes intersect in a line in this plane. For example, if the point (a, b, c) we start with happens to lie on the b-axis, then the tangent planes are the ab-plane and the bc-plane. These intersect the plane $a + c = 1$ in the two vertical tangent lines to the ellipse E.



Our goal will be to show the following:

Proposition 5.11. *Let $Q(x, y) = ax^2 + bxy + cy^2$ be a form of positive discriminant, either hyperbolic or 0-hyperbolic. Then the two points where the tangent lines to E determined by (a, b, c) touch E are the points diametrically opposite the two points that are the endpoints of the separator line in the topograph of Q in the case that Q is hyperbolic, or the two points labeling the regions in the topograph of Q where Q takes the value zero in the case that Q is 0-hyperbolic.*

Proof: We begin with a few preliminary remarks that will allow us to treat the hyperbolic and 0-hyperbolic cases in the same way. A form $Q(x, y) = ax^2 + bxy + cy^2$

of positive discriminant can always be factored as $(px+qy)(rx+sy)$ with p,q,r,s real numbers since if $a=0$ we have the factorization $y(bx+cy)$ and if $a\neq 0$ then the associated quadratic equation $ax^2+bx+c=0$ has positive discriminant so it has two distinct real roots α and β. This leads to the factorization $ax^2+bxy+cy^2 = a(x-\alpha y)(x-\beta y)$ which can be rewritten as $(px+qy)(rx+sy)$ by incorporating a into either factor. If Q is hyperbolic then the discriminant is not a square and hence the factorization $(px+qy)(rx+sy)$ will involve coefficients that are quadratic irrationals. If Q is 0-hyperbolic then the discriminant is a square so the roots α and β are rational and we obtain a factorization of Q as $(px+qy)(rx+sy)$ with rational coefficients. In fact we can take p,q,r,s to be integers in this case since we know every 0-hyperbolic form is equivalent to a form $y(bx+cy)$ so we can obtain the given form Q from $y(bx+cy)$ by replacing x and y by certain linear combinations $dx+ey$ and $fx+gy$ with integer coefficients d,e,f,g.

The points where the tangent planes touch the double cone correspond to forms of discriminant zero, with coefficients that may not be integers or even rational. A simple way to construct two such forms from a given form $Q=(px+qy)(rx+sy)$ is just to take the squares of the two linear factors, so we obtain the forms $(px+qy)^2$ and $(rx+sy)^2$, each of discriminant zero. We will show that each of these two forms lies on the line of tangency for one of the two tangent planes determined by Q.

To do this for the case of $(px+qy)^2$ we consider the line L in $\mathbb{R}^3$ passing through the two points corresponding to the forms $(px+qy)(rx+sy)$ and $(px+qy)^2$. We claim that L consists of the forms

$$Q_t(x,y) = (px+qy)\Big[(1-t)(rx+sy)+t(px+qy)\Big]$$

as t varies over all real numbers. When $t=0$ or $t=1$ we obtain the two forms $Q_0=(px+qy)(rx+sy)$ and $Q_1=(px+qy)^2$ so these forms lie on L. Also, we can see that the forms Q_t do form a straight line in $\mathbb{R}^3$ by rewriting the formula for $Q_t(x,y)$ as $ax^2+bxy+cy^2$ with the coefficients a,b,c given by:

$$(a,b,c) = (pr(1-t)+p^2t,(ps+qr)(1-t)+2pqt,qs(1-t)+q^2t)$$

This defines a line since p,q,r,s are constants, so each coordinate is a linear function of t. Since the forms Q_t factor as the product of two linear factors, they have nonnegative discriminant for all t. This means that the line L does not go into the interior of either cone. It also does not pass through the origin since if it did, it would have to be a subset of the double cone since it contains the form Q_1 which lies in the double cone. From these facts we deduce that L must be a tangent line to the double cone. Hence the plane containing L and the origin must be tangent to the double cone along the line containing the origin and Q_1. The same reasoning shows that the other tangent plane that passes through $(px+qy)(rx+sy)$ intersects the double cone along the line containing the origin and $(rx+sy)^2$.

The labels of the points of E corresponding to the two forms $(px+qy)^2$ and $(rx+sy)^2$ are p/q and r/s according to the convention we have adopted. On the other hand, when the form $(px+qy)(rx+sy)$ is hyperbolic the ends of the separator line in its topograph are at the two points where this form is zero, which occur when x/y is $-q/p$ and $-s/r$. These are the negative reciprocals of the previous two points p/q and r/s so they are the diametrically opposite points in E. Similarly, when the form $(px+qy)(rx+sy)$ is 0-hyperbolic the vertices of the Farey diagram where it is zero are at $-q/p$ and $-s/r$, again diametrically opposite p/q and r/s. □

It might have been nicer if the statement of the previous proposition did not involve passing to diametrically opposite points, but to achieve this we would have had to use a different rule for labeling the points of E, with the label p/q corresponding to the form $(qx-py)^2$ instead of $(px+qy)^2$. This 180 degree rotation of the labels would put the negative labels in the upper half of E rather than the lower half, which does not seem like a good idea.

Next let us investigate how hyperbolic and 0-hyperbolic forms are distributed over the lattice points outside the double cone $b^2-4ac=0$. This is easier to visualize if we project such points radially into the plane $a+c=1$. This only works for forms $ax^2+bxy+cy^2$ with $a+c>0$, but the forms with $a+c<0$ are just the negatives of these so they give nothing essentially new. The forms with $a+c=0$ will be covered after we deal with those with $a+c>0$.

Forms with $a+c>0$ that are hyperbolic or 0-hyperbolic correspond via radial projection to points in the plane $a+c=1$ outside the ellipse E. As we have seen, each such point determines a pair of tangent lines to E intersecting at the given point.

For a 0-hyperbolic form $(px+qy)(rx+sy)$ the points of tangency in E have rational labels p/q and r/s. We know that every 0-hyperbolic form is equivalent to a form $y(rx+sy)$ with $a=0$, so $p/q=0/1$ and one line of tangency is the vertical line tangent to E on the right side. The form $y(rx+sy)$ corresponds to the point $(0,r,s)$ in the plane $a=0$ tangent to the double cone. Projecting radially into the vertical tangent line to E, we obtain the points $(0,r/s,1)$, where r/s is an arbitrary rational number. Thus 0-hyperbolic forms are dense in this vertical tangent line to E. Choosing any rational number r/s, the other tangent line for the form $y(rx+sy)$ is tangent to E at the point labeled r/s.

An arbitrary 0-hyperbolic form $(px+qy)(rx+sy)$ is obtained from one with $p/q=0/1$ by applying a linear fractional transformation T taking $0/1$ to p/q, so the vertical tangent line to E at $0/1$ is taken to the tangent line at p/q, and the dense set of 0-hyperbolic forms in the vertical tangent line is taken to a dense set of 0-hyperbolic forms in the tangent line at p/q. Thus we see that the 0-hyperbolic forms in the plane $a+c=1$ consist of all the rational points on all the tangent lines to E at rational points p/q of E.

In the case of a hyperbolic form $ax^2+bxy+cy^2$ with $a+c>0$ the two tangent lines intersect E at a pair of conjugate quadratic irrationals, the negative reciprocals of the roots α and $\overline{\alpha}$ of the equation $ax^2+bx+c=0$. Since α determines $\overline{\alpha}$ uniquely, one tangent line determines the other uniquely, unlike the situation for 0-hyperbolic forms whose rational tangency points p/q and r/s can be varied independently. A consequence of this uniqueness for hyperbolic forms is that each of the two tangent lines contains only one rational point, the intersection point of the two lines. This is because any other rational point would correspond to another form having one of its tangent lines the same as for $ax^2+bxy+cy^2$ and the other tangent line different, contradicting the previous observation that each tangent line for a hyperbolic form determines the other. (The hypothetical second form would also be hyperbolic since the common tangency point for the two forms is not a rational point on E.)

The points in the plane $a+c=1$ that correspond to 0-hyperbolic forms are dense in the region of this plane outside E since for an arbitrary point in this region we can first take the two tangent lines to E through this point and then take a pair of nearby lines that are tangent at rational points of E since points in E with rational labels are dense in E. It is also true that points in the plane $a+c=1$ that correspond to hyperbolic forms are dense in the region outside E. To see this we can proceed in two steps. First consider the case of a point in this region whose two tangent lines to E are tangent at irrational points of E. These two irrational points are the endpoints of an infinite strip in the Farey diagram that need not be periodic. However we can approximate this strip by a periodic strip by taking a long finite segment of the infinite strip and then repeating this periodically at each end. This means that the given point in the region outside E lies arbitrarily close to points corresponding to hyperbolic forms. Finally, a completely arbitrary point in the region outside E can be approximated by points whose tangent lines to E touch E at irrational points since irrational numbers are dense in real numbers.

It remains to consider hyperbolic and 0-hyperbolic forms $(px+qy)(rx+sy)$ corresponding to points (a,b,c) in the plane $a+c=0$. Such a form determines a line through the origin in this plane, and the tangent planes to the double cone that intersect in this line intersect the plane $a+c=1$ in two parallel lines tangent to E at two diametrically opposite points p/q and $-q/p$. This means that the form is $(px+qy)(qx-py)$, up to a constant multiple. If p/q is rational this is a 0-hyperbolic form. Examples are:

- xy with vertical tangents to E at $1/0$ and $0/1$.
- $x^2-y^2=(x+y)(x-y)$ with horizontal tangents to E at $1/1$ and $-1/1$.
- $2x^2-3xy-2y^2=(2x+y)(x-2y)$ with parallel tangents at $2/1$ and $-1/2$.

If p/q and $-q/p$ are conjugate quadratic irrationals then we have a hyperbolic form $ax^2+bxy+cy^2=a(x-\alpha)(x-\overline{\alpha})$ where $\alpha\overline{\alpha}=-1$ since $c=-a$ when $a+c=0$. Thus α and $\overline{\alpha}$ are negative reciprocals of each other that are interchanged by 180 degree rotation of E. As examples we have:

$$x^2 + xy - y^2 = \left(x - \frac{-1+\sqrt{5}}{2}y\right)\left(x - \frac{-1-\sqrt{5}}{2}y\right)$$
$$2x^2 + xy - 2y^2 = 2\left(x - \frac{-1+\sqrt{17}}{4}y\right)\left(x - \frac{-1-\sqrt{17}}{4}y\right)$$

One can consider a pair of parallel tangent lines to E as the limit of a pair of intersecting tangents where the point of intersection moves farther and farther away from E in a certain direction which becomes the direction of the pair of parallel tangents.

6 Representations by Quadratic Forms

With the various things we have learned about quadratic forms so far, let us return to the basic representation problem of determining what values a given form $Q(x,y) = ax^2 + bxy + cy^2$ can take on when x and y are integers, or in other words, which numbers can be represented as $ax^2 + bxy + cy^2$ for some choice of integers x and y. Remember that it suffices to restrict attention to the values of Q appearing in the topograph since these are the values for primitive pairs (x,y), and to get all other values one just multiplies the values in the topograph by arbitrary squares. With this in mind we will adopt the following convention in the rest of the book:

When we say that a form Q represents a number n we mean that $n = Q(x,y)$ for some **primitive** *pair of integers $(x,y) \neq (0,0)$.*

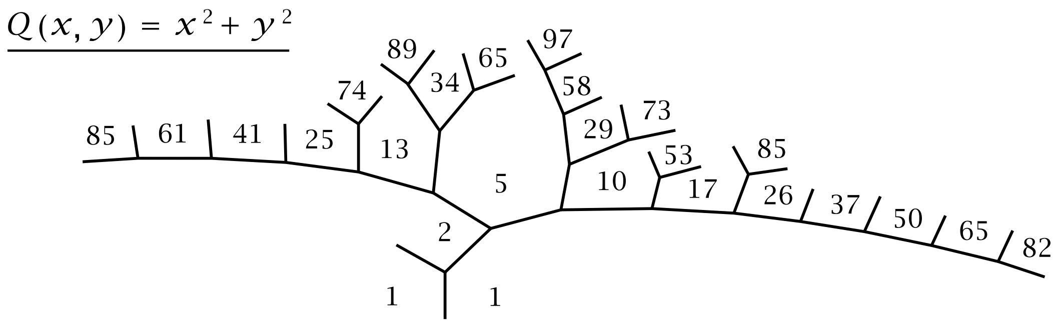
This differs from the traditional terminology in which any solution of $n = Q(x,y)$ is called a representation of n, without requiring (x,y) to be a primitive pair, and when (x,y) is primitive it is called a proper or primitive representation of n. However, since we will rarely consider the case that (x,y) is not a primitive pair, it will save many words not to have to insert the extra modifier for every representation.

We will focus on forms that are either elliptic or hyperbolic, as these are the most interesting cases.

6.1 Three Levels of Complexity

In this section we will look at a series of examples to try to narrow down what sort of answer one could hope to obtain for the representation problem. The end result will be a reasonable guess that will be verified in the rest of this chapter and the next one, at least for fundamental discriminants. For nonfundamental discriminants there is sometimes a small extra wrinkle that seems to be rather subtle and more difficult to analyze.

As a first example let us try to find a general pattern in the values of the form x^2+y^2. In view of the symmetry of the topograph for this form it suffices to look just in the first quadrant of the topograph. Part of this quadrant is shown in the figure below, somewhat distorted to fit more numbers into the picture. What is shown is all the numbers in the topograph that are less than 100.



At first glance it may be hard to detect any patterns here. Both even and odd numbers occur, but none of the even numbers are divisible by 4 so they are all twice an odd number, and in fact an odd number that appears in the topograph. Considering the odd numbers, one notices they are all congruent to 1 mod 4 and not 3 mod 4, which is the other possibility for odd numbers. On the other hand, not all odd numbers congruent to 1 mod 4 appear in the topograph. Up to 100, the ones that are missing are 9, 21, 33, 45, 49, 57, 69, 77, 81, and 93. Each of these has at least one prime factor congruent to 3 mod 4, while all the odd numbers that do appear have all their prime factors congruent to 1 mod 4. Conversely, all products of primes congruent to 1 mod 4 are in the topograph.

This leads us to guess that the following might be true:

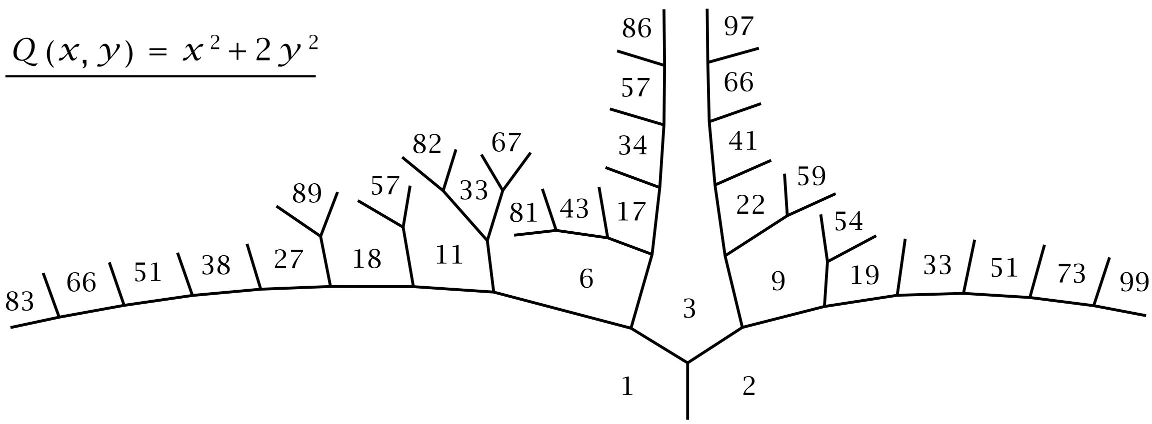
Conjecture. *The numbers that appear in the topograph of* $x^2 + y^2$ *are precisely the numbers* $n = 2^a p_1 p_2 \cdots p_k$ *where* $a \le 1$ *and each* p_i *is a prime congruent to* 1 *mod* 4*. Consequently, the values of the quadratic form* $Q(x, y) = x^2 + y^2$ *as* x *and* y *range over all integers (not just the primitive pairs) are exactly the numbers* $n = m^2 p_1 p_2 \cdots p_k$ *where* m *is an arbitrary integer and each* p_i *is either* 2 *or a prime congruent to* 1 *mod* 4*.*

In both statements the index k denoting the number of prime factors p_i is allowed to be zero as well as any positive integer. The restriction $a \le 1$ in the first statement disappears in the second statement since higher powers of 2 can occur when we multiply by arbitrary squares. We will prove the conjecture later in the chapter.

A weaker form of the conjecture can be proved just by considering congruences mod 4 as follows. An even number squared is congruent to 0 mod 4 and an odd number squared is congruent to 1 mod 4, so $x^2 + y^2$ must be congruent to 0, 1, or 2 mod 4. Moreover, the only way that $x^2 + y^2$ can be 0 mod 4 is for both x and y to be even, which cannot happen for primitive pairs. Thus all numbers in the topograph must be congruent to 1 or 2 mod 4. This says that the odd numbers in the topograph are congruent to 1 mod 4 and the even numbers are each twice an odd number.

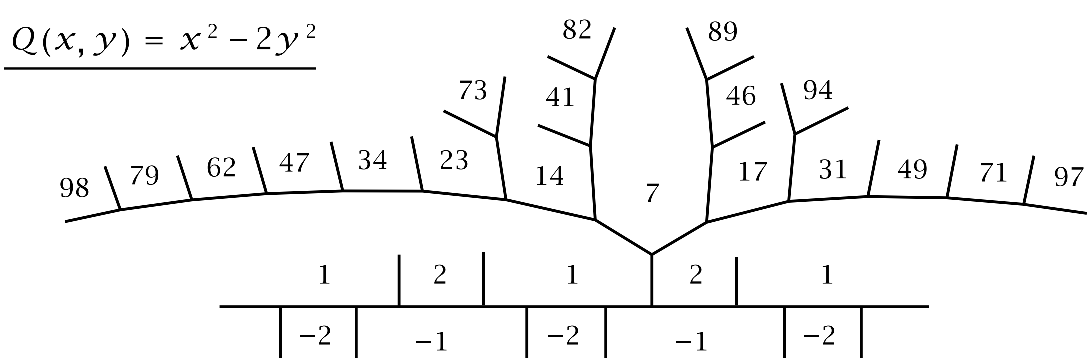
However, these simple observations say nothing about the role played by primes and prime factorizations, nor do they include any positive assertions about which numbers actually are represented by $x^2 + y^2$. It definitely takes more work to show for example that every prime $p = 4k+1$ can be represented as the sum of two squares.

Let us look at a second example to see whether the same sorts of patterns occur, this time for the form $Q(x,y) = x^2+2y^2$. Here is a portion of its topograph showing all values less than 100, with the lower half of the topograph omitted since it is just the mirror image of the upper half:



Again the even values are just the doubles of the odd values. The odd prime values are $3, 11, 17, 19, 41, 43, 59, 67, 73, 83, 89, 97$ and the other odd values are all the products of these primes. The odd prime values are not determined by their values mod 4 in this case, but instead by their values mod 8 since the primes we just listed are exactly the primes less than 100 that are congruent to 1 or 3 mod 8. Apart from this change, the answer to the representation problem for $x^2 + 2y^2$ is completely analogous to the answer for $x^2 + y^2$. Namely, the numbers represented by $x^2 + 2y^2$ are the numbers $n = 2^a p_1 p_2 \cdots p_k$ with $a \le 1$ and each p_i a prime congruent to 1 or 3 mod 8. Using congruences mod 8 we could easily prove the weaker statement that all numbers represented by $x^2 + 2y^2$ must be congruent to $1, 2, 3$, or 6 mod 8, so all odd numbers in the topograph must be congruent to 1 or 3 mod 8 and all even numbers must be twice an odd number.

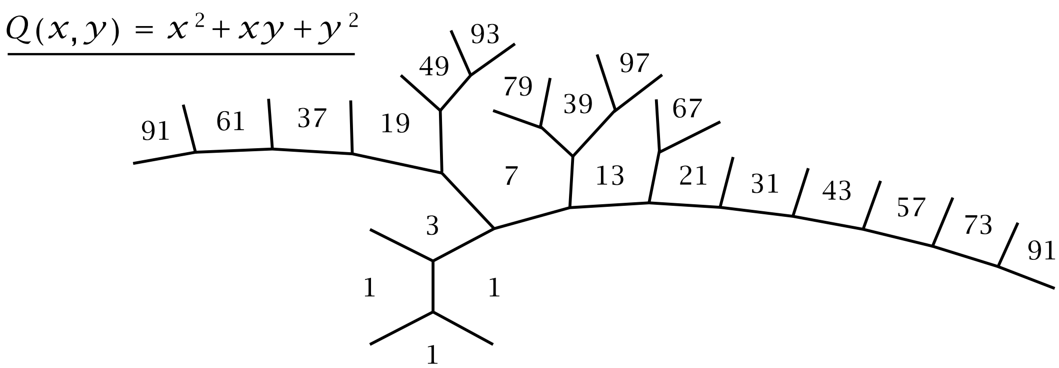
These two examples were elliptic forms, but the same sort of behavior can occur for hyperbolic forms as we see in the next example, the form $x^2 - 2y^2$. The negative values of this form happen to be just the negatives of the positive values, so we need only show the positive values in the topograph:



Here the primes that occur are 2 and primes congruent to ± 1 mod 8. The nonprime values that occur are the products of primes congruent to ± 1 mod 8 and twice these products. Again there is a weaker statement that can be proved using just congruences mod 8.

In these three examples the guiding principle was to look at prime factorizations and at primes modulo certain numbers, the numbers 4, 8, and 8 in the three cases. Notice that these numbers are just the absolute values of the discriminants -4, -8, and 8. Looking at primes mod $|\Delta|$ turns out to be a key idea for all quadratic forms.

Another example of the same sort is the form $x^2 + xy + y^2$ of discriminant -3. This time it is the prime 3 that plays a special role rather than 2.



We only have to draw one-sixth of the topograph because of all the symmetries. Notice that all the values are odd, so the prime 2 plays no role here. Since the discriminant is -3 we are led to consider congruences mod 3. The primes in the topograph are 3 and the primes congruent to 1 mod 3 (which in particular excludes the prime 2), namely the primes $7, 13, 19, 31, 37, 43, 61, 67, 73, 79, 97$. The nonprime values are the products of these primes with the restriction that the prime 3 never has an exponent greater than 1. This is analogous to the prime 2 never having an exponent greater than 1 in the preceding examples. In all four examples the "special" primes whose exponents are restricted are just the prime divisors of the discriminant. This is a general phenomenon, that primes dividing the discriminant behave differently from primes that do not divide the discriminant.

A special feature of the discriminants -4, -8, 8, and -3 is that in each case all forms of that discriminant are equivalent. We will see that the representation problem always has the same type of answer for discriminants with a single equivalence class of forms.

Before going on to the next level of complexity let us digress to describe a nice property that forms of the first level of complexity have. As we know, if an equation $Q(x,y) = n$ has an integer solution (x,y) then so does $Q(x,y) = m^2 n$ for every integer m. The converse is not always true however. For example the equation $2x^2 + 7y^2 = 9$ has the solution $(x,y) = (1,1)$ but $2x^2 + 7y^2 = 1$ obviously has no solution with x and y integers. Nevertheless, this converse property does hold for forms such as those in the preceding four examples where the numbers n for which $Q(x,y) = n$ has an integer solution are exactly the numbers that can be factored as $n = m^2 p_1 p_2 \cdots p_k$ for primes p_i satisfying certain conditions and m an arbitrary integer. This is because if a number n has a factorization of this type then we can cancel any square factor of n and the result still has a factorization of the same type.

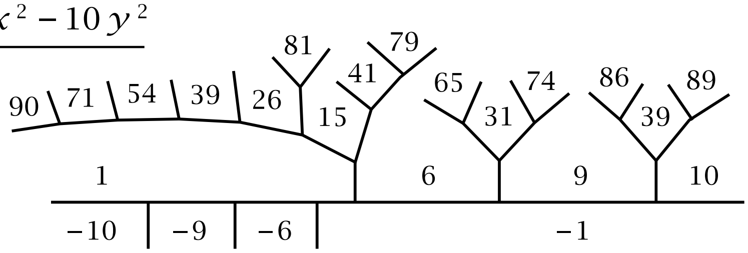
Let us apply this "square-cancellation" property in the case of the form x^2+y^2 to determine the numbers n such that the circle $x^2+y^2 = n$ contains a rational point, and hence, as in Chapter 0, an infinite dense set of rational points. Suppose first that the circle $x^2+y^2 = n$ contains a rational point, so after putting the two coordinates over a common denominator the point is $(x,y) = (a/c, b/c)$. The equation $x^2+y^2 = n$ then becomes $a^2+b^2 = c^2 n$. This means that the equation $x^2+y^2 = c^2 n$ has an integer solution. Then the square-cancellation property implies that the original equation $x^2+y^2 = n$ has an integer solution. Thus we see that if there are rational points on the circle $x^2+y^2 = n$ then there are integer points on it. This is not something that is true for all quadratic curves, as shown by the example of the ellipse $2x^2+7y^2 = 1$ which has rational points such as $(1/3, 1/3)$ but no integer points.

From the solution to the representation problem for x^2+y^2 we deduce that the circle $x^2+y^2 = n$ contains rational points exactly when $n = m^2 p_1 p_2 \cdots p_k$ where m is an arbitrary integer and each p_i is either 2 or a prime congruent to 1 mod 4. The first few values of n satisfying this condition are $1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, \cdots$.

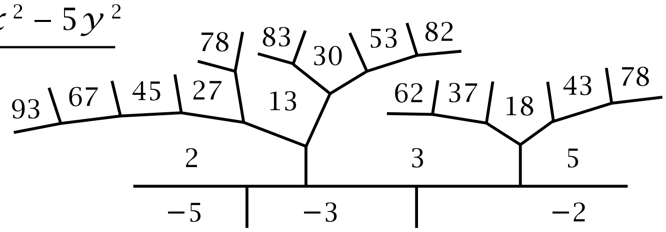
The Second Level of Complexity

For an example with slightly greater complexity consider discriminant 40 where the class number is 2 and two nonequivalent forms are $x^2 - 10y^2$ and $2x^2 - 5y^2$. The topographs below show the positive values less than 100.

$Q_1(x,y) = x^2 - 10y^2$



$Q_2(x,y) = 2x^2 - 5y^2$



The topographs are periodic and also have mirror symmetry so it suffices to show half of one period. There is no need to show any more of the negative values since these will just be the negatives of the positive values.

For the form $x^2 - 10y^2$ the prime values less than 100 are $31, 41, 71, 79, 89$. These are the primes congruent to ± 1 or ± 9 mod 40, the discriminant. However, in contrast to what happened in the previous examples, there are many nonprime values of this form that are not products of these prime values. The prime factors of these nonprime values are $2, 3, 5, 13, 37, 43$, none of which occur in the topograph of the first form. Rather miraculously, these prime values are realized instead by the second form $2x^2 - 5y^2$. The prime values this form takes on are 2 and 5, which are the prime divisors of the discriminant 40, along with primes congruent to ± 3 and ± 13 mod 40, namely $3, 13, 37, 43, 53, 67$, and 83.

Apart from the primes 2 and 5 that divide the discriminant, the possible values of primes mod 40 are $\pm 1, \pm 3, \pm 7, \pm 9, \pm 11, \pm 13, \pm 17, \pm 19$ since even numbers and multiples of 5 are excluded. There are sixteen different congruence classes here, and exactly half of them, eight, are realized by one or the other of the two forms $x^2 - 10y^2$ and $2x^2 - 5y^2$, with four classes realized by each form. The other eight congruence classes are not realized by any form of discriminant 40 since every form of discriminant 40 is equivalent to one of the two forms $x^2 - 10y^2$ or $2x^2 - 5y^2$, as is easily checked by the methods from the previous chapter.

This turns out to be a general phenomenon valid for elliptic and hyperbolic forms of any discriminant Δ: If one excludes the primes that divide Δ, then the prime values of quadratic forms of discriminant Δ are exactly the primes in half of the congruence classes mod Δ of numbers coprime to Δ. This will be proved in Proposition 6.23. Also, each form represents primes in the same number of congruence classes. For $\Delta = 40$ this is four congruence classes for each form.

The primes 2 and 5 that divide the discriminant occur in the topographs only to the first power, and in fact no numbers in the topographs are divisible by 2^2 or 5^2. This is similar to what happened in the earlier examples where there was only one prime dividing the discriminant. Apart from this restriction it appears that each product of primes represented by Q_1 or Q_2 is also represented by Q_1 or Q_2. The problem is to decide which form represents which products. For numbers in the topographs not divisible by 2 or 5 it seems that these numbers are subject to the same congruence conditions as for primes, so they are congruent to ± 1 or ± 9 for Q_1 and to ± 3 or ± 13 for Q_2.

If one includes numbers divisible by 2 or 5 the following statements seem to be true, provided that numbers divisible by 2^2 or 5^2 are excluded:

- The product of two numbers represented by Q_1 is again represented by Q_1.
- The product of two numbers represented by Q_2 is represented by Q_1.
- The product of a number represented by Q_1 with a number represented by Q_2 is represented by Q_2.

To illustrate the first statement, the numbers 6, 9, and 10 appear in the topograph of Q_1 hence so do $6 \cdot 9$, $9 \cdot 9$, and $9 \cdot 10$, but not $6 \cdot 10$ since this is divisible by 2^2. For the second statement, the numbers 2, 3, and 5 are in the topograph of Q_2 so $2 \cdot 3$, $3 \cdot 3$, $2 \cdot 5$, and $3 \cdot 5$ are in the topograph of Q_1 but not $2 \cdot 2$ or $5 \cdot 5$. The product $2 \cdot 3 \cdot 5$ is then in the topograph of Q_2 by the third statement.

An abbreviated way of writing the three rules is by the formulas $Q_1 Q_1 = Q_1$, $Q_2 Q_2 = Q_1$, and $Q_1 Q_2 = Q_2$. One can see that these are formally the same as the rules for addition of integers mod 2: $0 + 0 = 0$, $1 + 1 = 0$, and $0 + 1 = 1$. The two formulas $Q_1 Q_1 = Q_1$ and $Q_1 Q_2 = Q_2$ say that Q_1 serves as an identity element for this multiplication operation, and then the formula $Q_2 Q_2 = Q_1$ can be interpreted as saying that Q_2 is equal to its own inverse, so $Q_2 = Q_2^{-1}$.

This way of "multiplying" forms is more than just shorthand notation, and in Chapter 7 we will develop a general method for forming products of primitive forms of a fixed discriminant that will be a key ingredient in reducing the representation problem to the special case of representing primes.

The various observations we have made so far about the two forms of discriminant 40 lead to the following:

Conjecture. *The positive numbers represented by either Q_1 or Q_2 are exactly the products $2^a 5^b p_1 p_2 \cdots p_k$ where $a, b \le 1$ and each p_i is a prime congruent to ± 1, ± 3, ± 9, or ± 13 mod 40. The form Q_1 represents the primes $p_i \equiv \pm 1$ and ± 9 while Q_2 represents 2, 5, and the primes $p_i \equiv \pm 3$ and ± 13. One can determine which form will represent a product $2^a 5^b p_1 p_2 \cdots p_k$ by the rule that if the number of terms in the product that are represented by Q_2 is even then the product is represented by Q_1 and if it is odd then the product is represented by Q_2.*

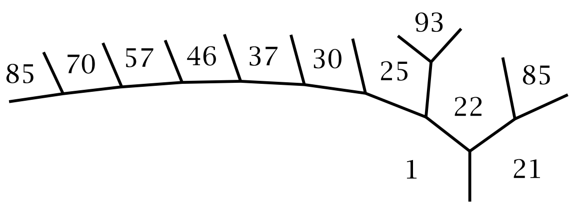
For example, the topograph of Q_1 contains the even powers of 3 while the topograph of Q_2 contains the odd powers. Another consequence is that the even values in one topograph are just the doubles of the odd values in the other topograph.

This characterization of numbers represented by these two forms also implies that no number is represented by both Q_1 and Q_2. However, for some discriminants it is possible for two nonequivalent forms of that discriminant to represent the same nonzero number, as we will see.

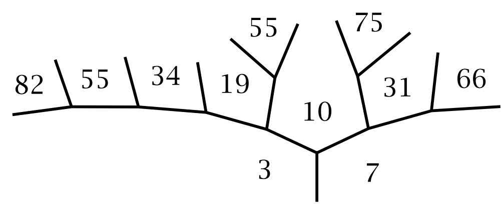
The Conjecture will be proved piece by piece as we gradually develop the necessary general theory. The first statement will be an application of Theorem 6.8 together with later facts in Section 6.2. The second statement will be an application of Proposition 6.19 and the rest of the Conjecture will use results from Chapter 7, particularly Theorem 7.7.

Let us look at another example where the representation problem has an answer that is qualitatively similar to the preceding example but just a little more complicated, the case of discriminant -84. Here there are twice as many equivalence classes of forms, four instead of two, with topographs shown below.

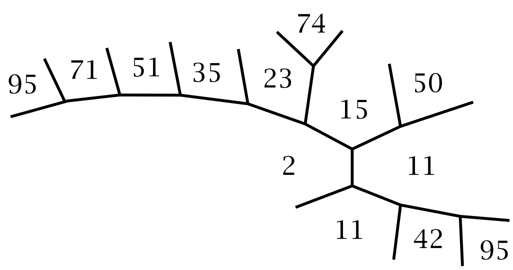
$Q_1(x,y) = x^2 + 21y^2$



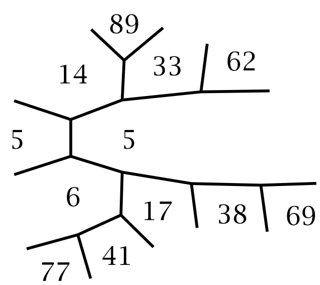
$Q_2(x,y) = 3x^2 + 7y^2$



$Q_3(x,y) = 2x^2 + 2xy + 11y^2$



$Q_4(x,y) = 5x^2 + 4xy + 5y^2$



The primes dividing the discriminant -84 are 2, 3, and 7, and these primes are each represented by one of the forms. In fact the divisors of the discriminant that appear in the topographs are $1, 2, 3, 6, 7, 14, 21$, and 42 which are precisely the squarefree divisors of the discriminant. These squarefree divisors of Δ are exactly the numbers appearing on reflector lines of mirror symmetries of the topographs. This was the case also in the previous examples, as one can check, and is a general phenomenon for fundamental discriminants as we saw in Propositions 5.6 and 5.7.

For the primes not dividing the discriminant, we will show in Section 6.3 that the primes represented by each form are as follows:

- For Q_1 the primes $p \equiv 1, 25, 37 \bmod 84$.
- For Q_2 the primes $p \equiv 19, 31, 55 \bmod 84$.
- For Q_3 the primes $p \equiv 11, 23, 71 \bmod 84$.
- For Q_4 the primes $p \equiv 5, 17, 41 \bmod 84$.

This agrees with what is shown in the four topographs above, and one could expand the topographs to get further evidence that these are the right answers. Passing from primes to arbitrary numbers appearing in at least one of the topographs, these appear to be exactly the products $2^a 3^b 7^c p_1 \cdots p_k$ with $a, b, c \le 1$ and each p_i one of the other primes represented by Q_1, Q_2, Q_3, or Q_4.

One can work out hypothetical rules for multiplying the forms by considering how products of two primes are represented. For example, 3 is represented by Q_2 and 11 is represented by Q_3, while their product $3 \cdot 11 = 33$ is represented by Q_4, so we might guess that $Q_2Q_3 = Q_4$. Some other products that give the same conclusion are $3\cdot 2 = 6$, $3\cdot 23 = 69$, $7\cdot 2 = 14$, $7\cdot 11 = 77$, and $31\cdot 2 = 62$. In the same way one can determine tentative rules for all the products Q_iQ_j, with the following results:

- The principal form Q_1 acts as the identity, so $Q_1Q_i = Q_i$ for each i.
- $Q_iQ_i = Q_1$ for each i so each Q_i equals its own inverse.
- The product of any two out of Q_2, Q_3, Q_4 is equal to the third.

These multiplication rules are formally identical to how one would add pairs (m,n) of integers mod 2 by adding their two coordinates separately. The form Q_1 corresponds to $(0,0)$ and the first of the three rules above becomes $(0,0)+(m,n)=(m,n)$. The forms Q_2, Q_3, and Q_4 correspond to $(1,0)$, $(0,1)$, and $(1,1)$ in any order, and the second rule above becomes $(m,n)+(m,n) = (0,0)$ which is valid for addition mod 2, while the third rule becomes the fact that the sum of any two of $(1,0)$, $(0,1)$, and $(1,1)$ is equal to the third if we do addition mod 2.

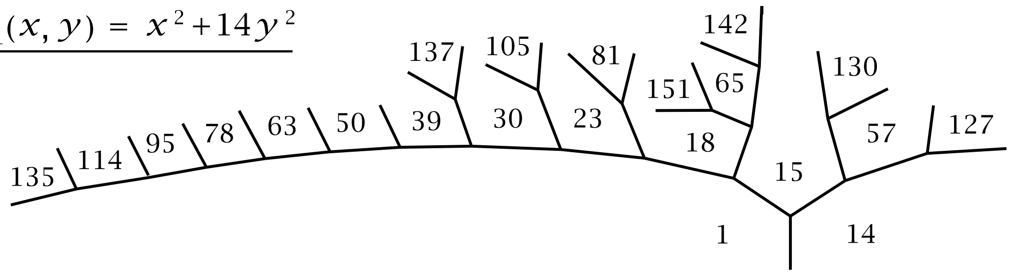
The multiplication rules determine which form represents a given number n by replacing each prime in the prime factorization of n by the form Q_i that represents it, then multiplying out the resulting product using the three multiplication rules, keeping in mind that 2, 3, and 7 can never occur with an exponent greater than 1. For example, for $n = 70 = 2\cdot 5\cdot 7$ we get the product $Q_3Q_4Q_2$ which equals Q_1 and so 70 is represented by Q_1, as the topograph shows. For $n = 66 = 2\cdot 3\cdot 11$ we get $Q_3Q_2Q_3 = Q_2$ and 66 is represented by Q_2. In general, for a number $n = 2^a3^b7^cp_1\cdots p_k$ we can determine which form represents n by the following steps. First compute the number q_i of prime factors of n represented by Q_i. Next compute the sum $q_1(0,0)+q_2(1,0)+q_3(0,1)+q_4(1,1) = (q_2+q_4, q_3+q_4)$ where $(0,0),(1,0),(0,1),(1,1)$ correspond to Q_1,Q_2,Q_3,Q_4 respectively. The resulting sum (r,s) mod 2 then tells which form represents n.

An interesting feature of all the forms at the first or second level of complexity that we have examined so far is that their topographs have mirror symmetry. This is actually a general phenomenon: Whenever all the forms of a given discriminant have mirror symmetry, then one can determine which primes are represented by each form just in terms of congruence conditions modulo the discriminant. And in fact this is the only time when congruences modulo the discriminant determine how primes are represented, at least if one restricts attention just to primitive forms. This will be shown in Corollary 6.29. In Chapter 5 we called discriminants for which all primitive forms have mirror symmetry *fully symmetric* discriminants, and we observed that they are unfortunately rather rare, with only 101 negative discriminants known to have this property, and probably no more.

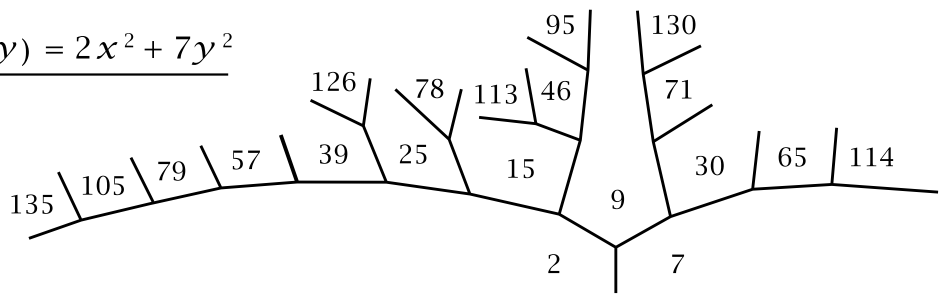
The Third Level of Complexity

A deeper degree of complexity is illustrated by the case $\Delta = -56$ where there are three equivalence classes of forms, with topographs shown below. The first two topographs have mirror symmetry but the third topograph does not, so the third form counts twice when determining the class number for discriminant -56, which is therefore 4 rather than 3.

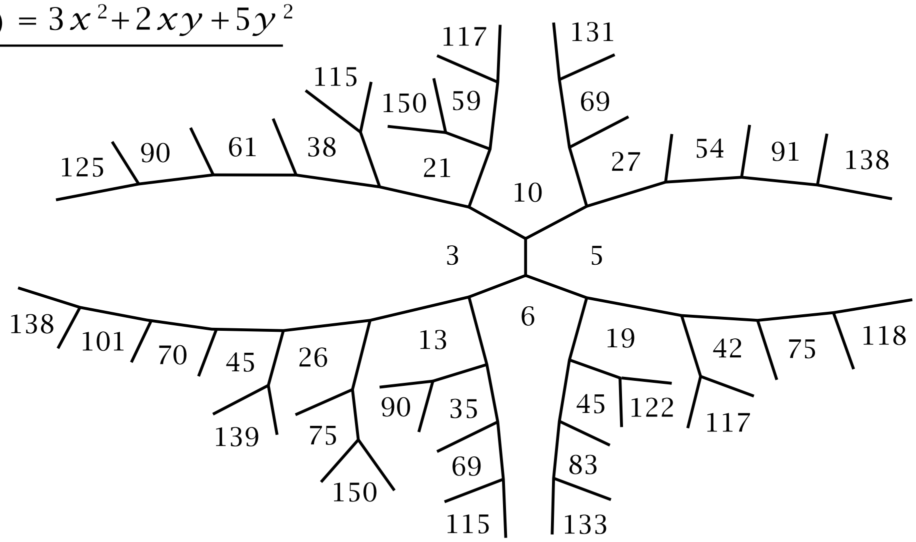
$Q_1(x, y) = x^2 + 14y^2$



$Q_2(x, y) = 2x^2 + 7y^2$



$Q_3(x, y) = 3x^2 + 2xy + 5y^2$



The behavior of divisors of the discriminant is the same as in the previous examples. Only the squarefree divisors appear, 1, 2, 7, and 14, and these are the numbers appearing on the reflector lines.

In the examples at the first two levels of complexity it was possible to determine which numbers are represented by a given form by looking at primes and which congruence classes they fall into modulo the discriminant. The primes represented by a given form were exactly the primes in certain congruence classes modulo the discriminant. This is no longer true for discriminant -56 however. For example the primes 23 and 79 are congruent mod 56, and yet 23 is represented by $Q_1 = x^2 + 14y^2$ since

$Q_1(3,1) = 23$, while 79 is represented by $Q_2 = 2x^2 + 7y^2$ since $Q_2(6,1) = 79$.

Another nice property that held in the previous examples was that no number appeared in more than one topograph for the given discriminant, but this too fails for discriminant -56 since there are many nonprimes that occur in the topographs of both Q_1 and Q_2 starting with $15, 30, 39, 57, 65, 78, 95, 105, 114, 130$, and 135.

Apart from the primes 2 and 7 that divide the discriminant -56, all other primes belong to the following 24 congruence classes mod 56, corresponding to odd numbers less than 56 not divisible by 7:

$$\underline{1}\ \overline{3}\ \overline{5}\ \underline{9}\ 11\ \overline{13}\ \underline{15}\ 17\ \overline{19}\ \underline{23}\ \underline{25}\ \overline{27}\ 29\ 31\ 33\ 37\ \underline{39}\ 41\ 43\ \overline{45}\ 47\ 51\ 53\ 55$$

The six congruence classes whose prime elements are represented by Q_1 or Q_2 are indicated by underlines, and the six congruence classes whose prime elements are represented by Q_3 are indicated by overlines. Primes not represented by any of the three forms are in the remaining twelve congruence classes.

The new thing that happens in this example is that one cannot tell whether a prime is represented by Q_1 or Q_2 just by considering congruence classes mod the discriminant. We saw this for the pair of primes 23 and 79, and another such pair visible in the topographs is 71 and 127. By extending the topographs we could find many more such pairs. One might try using congruences modulo some other number besides 56, but it is known that this does not help.

Congruences mod 56 suffice to tell which primes are represented by Q_3, but there is a different sort of novel behavior involving Q_3 when we look at representing products of primes. To illustrate this, observe that the primes 3 and 5 are represented by Q_3 but their product 15 is represented by both Q_1 and Q_2. This means there is some ambiguity about whether the product Q_3Q_3 should be Q_1 or Q_2. The same thing happens in fact for any pair of coprime numbers represented by Q_3, for example 5 and 6 whose product is represented by both Q_1 and Q_2.

For other products Q_iQ_j there seems to be no ambiguity. The principal form Q_1 acts as the identity for multiplication, while $Q_2Q_2 = Q_1$ and $Q_2Q_3 = Q_3$, although this last formula is somewhat odd since it seems to imply that Q_3 does not have a multiplicative inverse since if it did, we could multiply the equation $Q_2Q_3 = Q_3$ by this inverse to get $Q_2 = Q_1$.

There is a way out of these difficulties, discovered by Gauss. The troublesome form Q_3 is different from the other forms in this example and in the preceding examples in that it does not have mirror symmetry. Thus the equivalence class of Q_3 splits into two proper equivalence classes, with Q_3 having a mirror image form $Q_4 = 3x^2 - 2xy + 5y^2$ obtained from Q_3 by changing the sign of either x or y and hence changing the coefficient of xy to its negative. Using Q_4 we can then resolve the ambiguous product Q_3Q_3 by setting $Q_3Q_3 = Q_2 = Q_4Q_4$ and $Q_3Q_4 = Q_1$ so that Q_4 is the inverse of Q_3. This means that each Q_i has its inverse given by the mirror image topograph since Q_1 and Q_2 have mirror symmetry and equal their own inverses.

The rigorous justification for the formulas $Q_3Q_3 = Q_2 = Q_4Q_4$ and $Q_3Q_4 = Q_1$ will come in Chapter 7, but for the moment one can check that these formulas are at least consistent with the topographs.

Since $Q_3^2 = Q_2$ we have $Q_3^4 = Q_2^2 = Q_1$. Multiplying the equation $Q_3^4 = Q_1$ by Q_4, the inverse of Q_3, gives $Q_3^3 = Q_4$. Thus all four proper equivalence classes of forms are powers of the single form Q_3 since $Q_3^2 = Q_2$, $Q_3^3 = Q_4$, and $Q_3^4 = Q_1$. This is corroborated by the representations of powers of 3 since 3 is represented by Q_3, 3^2 by $Q_3^2 = Q_2$, 3^3 by $Q_3^3 = Q_4$, and 3^4 by $Q_3^4 = Q_1$. Products of powers Q_3^i are computed by adding exponents mod 4 since Q_3^4 is the identity. Thus multiplication of the four forms is formally identical with addition of integers mod 4. The earlier doubtful formula $Q_2Q_3 = Q_3$ is resolved into the two formulas $Q_2Q_3 = Q_4$ and $Q_2Q_4 = Q_3$, which become $Q_3^2Q_3 = Q_3^3$ and $Q_3^2Q_3^3 = Q_3^5 = Q_3$.

The appearance of the same number in two different topographs is easy to explain now that we have two forms Q_3 and Q_4 representing exactly the same numbers. For example, to find all appearances of the number $15 = 3 \cdot 5$ in the topographs we observe that its prime factors 3 and 5 appear in the topographs of both Q_3 and Q_4 so 15 will appear in the topographs of $Q_3Q_3 = Q_2$, $Q_3Q_4 = Q_1$, and $Q_4Q_4 = Q_2$, although this last formula gives no new representations.

The procedure for finding which forms represent a number $n = 2^a7^bp_1 \cdots p_k$ with $a, b \leq 1$ and primes p_i different from 2 or 7 is to replace each prime factor in this product by a form Q_j that represents it, then multiply out the resulting product of forms Q_j. There is also an extra condition that will be justified in Chapter 7: Whenever a prime p_i appears more than once in the prime factorization of n, we should replace all of its appearances by the same Q_j. For example, the forms representing $18 = 2 \cdot 3^2$ are just the products $Q_2Q_3^2 = Q_1$ and $Q_2Q_4^2 = Q_1$ and not $Q_2Q_3Q_4 = Q_2$, as one can see in the topographs. Similarly, $9 = 3 \cdot 3$ is represented only by $Q_3^2 = Q_2 = Q_4^2$ and not by $Q_3Q_4 = Q_1$.

We will show in Chapter 7 that the set of proper equivalence classes of primitive forms of fixed discriminant always has a multiplication operation compatible with multiplying values of forms of that discriminant in the way illustrated by the preceding examples. This multiplication operation gives this set the structure of a group, that is, a set with an associative multiplication operation for which there is an element of the set that functions as an identity for the multiplication, and such that each element of the set has a multiplicative inverse in the set whose product with the given element is the identity element. The set of proper equivalence classes of primitive forms with this group structure is called the *class group* for the given discriminant. The identity element is the class of the principal form, and the inverse of a class is obtained by taking the mirror image topograph.

The class group has the additional property that the multiplication is commutative. This makes its algebraic structure much simpler than the typical noncommutative group. An example of a noncommutative group that we have seen is the group $LF(\mathbb{Z})$ of linear fractional transformations, where the multiplication comes from multiplication of 2×2 matrices, or equivalently, composition of the transformations.

For a given discriminant, if the numbers represented by two primitive forms cannot be distinguished by congruences modulo the discriminant, then these two forms are said to belong to the same ***genus***. Thus in the preceding example of discriminant -56 the two forms Q_1 and Q_2 are of the same genus while Q_3 is of a different genus from Q_1 and Q_2, so there are two different genera ("genera" is the plural of "genus").

Equivalent forms always belong to the same genus since their topographs contain exactly the same numbers. The first two of the three levels of complexity we have described correspond to the discriminants where there is only one equivalence class in each genus. As we stated earlier, this desirable situation is also characterized by the condition that all primitive forms of the given discriminant have mirror symmetry. For larger discriminants there can be large numbers of genera and large numbers of equivalence classes within a genus. However, for a fixed discriminant there are always the same number of proper equivalence classes within each genus, as we will show in Corollary 7.27. This is illustrated by the case $\Delta = -56$ where one genus consists of Q_1 and Q_2 and the other genus consists of Q_3 and Q_4.

Dirichlet's Theorem on Primes in Arithmetic Progressions

The examples in this section show the significance of primes in certain congruence classes for solving the representation problem. In the examples there seems to be no shortage of primes in each of the relevant congruence classes. For example, for the form $x^2 + y^2$ the primes represented, apart from 2, seem to be the primes congruent to 1 mod 4, the primes of the form $4k + 1$ starting with $5, 13, 17, 29, 37, 41, 53, \cdots$. The other possibility for odd primes is the sequence $3, 7, 11, 19, 23, 31, 43, 47, \cdots$, primes of the form $4k + 3$, or equivalently $4k - 1$.

Such sequences form arithmetic progressions $an + b$ for fixed positive integers a and b and varying $n = 0, 1, 2, 3, \cdots$. It is natural to ask whether there are infinitely many primes in each arithmetic progression $an + b$. For this to be true an obvious restriction is that a and b should be coprime since any common divisor of a and b will divide every number $an + b$, so there could be at most one prime in the progression.

A famous theorem of Dirichlet from 1837 asserts that every arithmetic progression $an + b$ with a and b coprime contains an infinite number of primes. This can be rephrased as saying that within each congruence class of numbers $x \equiv b$ mod a there are infinitely many primes whenever a and b are coprime. Dirichlet's theorem actually says more, that primes are approximately equally distributed among the various congruence classes mod a for a fixed a. For example, there are approximately as many primes $p = 4n + 1$ as there are primes $p = 4n - 1$.

Dirichlet's Theorem is not easy to prove, and a proof would require methods quite different from anything else in this book so we will not be giving a proof. However a few special cases of Dirichlet's Theorem can be proved by elementary arguments. The simplest case is the arithmetic progression $3, 7, 11, \cdots$ of numbers $n = 4n - 1$, using a variant of Euclid's proof that there are infinitely many primes. First let us recall how Euclid's argument goes: Suppose that $p_1, \cdots, p_k$ is a finite list of primes, and consider the number $N = p_1 \cdots p_k + 1$. This must be divisible by some prime p, but p cannot be any of the primes p_i on the list since dividing p_i into N gives a remainder of 1. Thus no finite list of primes can be complete and hence there must be infinitely many primes.

To adapt this argument to primes of the form $4n-1$, suppose that $p_1, \cdots, p_k$ is a finite list of such primes, and consider the number $N = 4p_1 \cdots p_k - 1$. The prime divisors of N must be odd since N is odd. If all these prime divisors were of the form $4n + 1$ then N would be a product of numbers of the form $4n + 1$ hence N itself would have this form, contradicting the fact that N has the form $4n - 1$. Hence N must have a prime factor $p = 4n - 1$. This p cannot be any of the primes p_i since dividing p_i into N gives a remainder of -1. Thus no finite list of primes $4n - 1$ can be a complete list.

This argument does not work for primes $p = 4n + 1$ since a number $N = 4p_1 \cdots p_k + 1$ can be a product of primes of the form $4n-1$, for example $21 = 3 \cdot 7$, so one could not deduce that N had a prime factor $p = 4n + 1$.

However, the quadratic form $x^2 + y^2$ can be used to show there are infinitely many primes $p = 4n+1$. In Proposition 6.18 we will show that for each discriminant Δ there are infinitely many primes represented by forms of discriminant Δ. In the case $\Delta = -4$ all forms are equivalent to the form x^2+y^2, so this form must represent infinitely many primes. None of these primes can be of the form $4n-1$ since all values of $x^2 + y^2$ are congruent to 0, 1, or 2 mod 4, as squares are always 0 or 1 mod 4. Thus there must be infinitely many primes $p = 4n + 1$.

The same arguments work also for primes $p = 3n + 1$ and $p = 3n - 1$. For $p = 3n - 1$ one argues just as for $4n - 1$, using numbers $N = 3p_1 \cdots p_k - 1$. For $p = 3n + 1$ one uses the form $x^2 + xy + y^2$ of discriminant -3. Here again all forms of this discriminant are equivalent so Proposition 6.18 says that $x^2 + xy + y^2$ represents infinitely many primes. All values of x^2+xy+y^2 are congruent to 0 or 1 mod 3 as one can easily check by listing the various possibilities for x and y mod 3. Thus there are infinitely many primes $p = 3n + 1$.

We can try these arguments for arithmetic progressions $5n \pm 1$ and $5n \pm 2$ but there are problems. The Euclidean argument we have given fails in each case for much the same reason that it failed for primes $p = 4n + 1$. For the approach via quadratic forms we would use the form $x^2 + xy - y^2$ of discriminant 5. This is the only form of this discriminant, up to equivalence, so Proposition 6.18 implies that it represents infinitely many primes. The methods in the next section will show that the primes represented by this form are the primes $p = 5n \pm 1$, so there are infinitely many primes $p = 5n+1$ or $p = 5n-1$ but we cannot be more specific than this. Dirichlet's Theorem says there are infinitely primes of each type, and in fact there are fancier forms of the Euclidean argument that prove this, but these Euclidean arguments do not work for the other cases $p = 5n \pm 2$.

We have just seen three quadratic forms that represent infinitely many primes, for discriminants -4, -3, and 5, and Proposition 6.18 provides other examples for each discriminant with class number 1. (Nonprimitive forms obviously cannot represent infinitely many primes, so these forms can be ignored.) For discriminants with larger class numbers Proposition 6.18 only implies that there is at least one form representing infinitely many primes. However there is another hard theorem of Dirichlet which does say that each primitive form of nonsquare discriminant represents infinitely many primes.

Exercises

1. For the form $Q(x,y) = x^2 + xy - y^2$ do the following things:

(a) Draw enough of the topograph to show all the values less than 100 that occur in the topograph. This form is hyperbolic and it takes the same negative values as positive values, so you need not draw all the negative values.

(b) Make a list of the primes less than 100 that occur in the topograph, and a list of the primes less than 100 that do not occur.

(c) Characterize the primes in the two lists in part (b) in terms of congruence classes mod $|\Delta|$ where Δ is the discriminant of Q.

(d) Characterize the nonprime values in the topograph in terms of their factorizations into primes in the lists in part (b).

(e) Summarize the previous parts by giving a simple criterion for determining the numbers n such that $Q(x,y) = n$ has an integer solution (x,y), primitive or not. The criterion should say something like $Q(x,y) = n$ is solvable if and only if $n = m^2 p_1 \cdots p_k$ where each p_i is a prime such that ...

(e) Check that all forms having the same discriminant as Q are equivalent to Q.

2. Do the same things for the form $x^2 + xy + 2y^2$, except that this time you only need to consider values less than 50 instead of 100.

3. For discriminant $\Delta = -24$ do the following:

(a) Verify that the class number is 2 and find two quadratic forms Q_1 and Q_2 of discriminant -24 that are not equivalent.

(b) Draw topographs for Q_1 and Q_2 showing all values less than 100. (You do not have to repeat parts of the topographs that are symmetric.)

(c) Divide the primes less than 100 into three lists: those represented by Q_1, those represented by Q_2, and those represented by neither Q_1 nor Q_2. (No primes are represented by both Q_1 and Q_2.)

(d) Characterize the primes in the three lists in part (c) in terms of congruence classes mod $|\Delta| = 24$.

(e) Characterize the nonprime values in the topograph of Q_1 in terms of their factorizations into primes in the lists in part (c), and then do the same thing for Q_2. Your answers should be in terms of whether there are an even or an odd number of prime factors from certain of the lists.

(f) Summarize the previous parts by giving a criterion for which numbers n the equation $Q_1(x,y) = n$ has an integer solution and likewise for the equation $Q_2(x,y) = n$.

4. This problem will show how things can be more complicated than in the previous problems.

(a) Show that the number of equivalence classes of forms of discriminant -23 is 2 while the number of proper equivalence classes is 3, and find reduced forms Q_1 and Q_2 of discriminant -23 that are not equivalent.

(b) Draw the topographs of Q_1 and Q_2 up to the value 70. (Again you do not have to repeat symmetric parts.)

(c) Find a number n that occurs in both topographs, and find the x and y values that give $Q_1(x_1,y_1) = n = Q_2(x_2,y_2)$. (This sort of thing never happens in the previous problems.)

(d) Find a prime p_1 in the topograph of Q_1 and a different prime p_2 in the topograph of Q_2 such that p_1 and p_2 are congruent mod $|\Delta| = 23$. (This sort of thing also never happens in the previous problems.)

5. Show there are infinitely many primes of the form $6m-1$ by an argument similar to the one used for $4m-1$.

6. Consider a discriminant $\Delta = q^2$, $q > 0$, corresponding to 0-hyperbolic forms. Using the description of the topographs of such forms obtained in the previous chapter, show:

(a) Every number is represented by at least one form of discriminant Δ, so in particular all primes are represented.

(b) The primes represented by a given form of discriminant Δ are exactly the primes in certain congruence classes mod q (and hence also mod Δ).

(c) For $q = 1, 2, 7$, and 15 determine the class number for discriminant $\Delta = q^2$ and find which primes are represented by the forms in each equivalence class.

6.2 Representations in a Fixed Discriminant

The problem of determining the numbers represented by a given form is difficult in general, so in this section we will consider the somewhat easier question of determining which numbers n are represented by at least one form of a given discriminant Δ, without specifying which form this will be. We refer to this as *representing n in discriminant Δ*.

On several occasions we will make use of the following fact: A form Q represents a number a if and only if Q is equivalent to a form $ax^2+bxy+cy^2$ with leading coefficient a. To see this, note first that the form $ax^2+bxy+cy^2$ obviously represents a when $(x,y)=(1,0)$, hence any form equivalent to $ax^2+bxy+cy^2$ also represents a. Conversely, if a form Q represents a then a appears in the topograph of Q, and by applying a suitable linear fractional transformation we can bring the region where a appears to the $1/0$ region, changing Q to an equivalent form $ax^2+bxy+cy^2$ where c is the new label on the $0/1$ region and b is the new label on the edge between the $1/0$ and $0/1$ regions.

Here is our first use of this principle:

Proposition 6.1. *If a number n is represented in discriminant Δ then so is every divisor of n.*

Thus for representations in a given discriminant, if we find which primes are represented and then which products of these primes are represented, we will have found all numbers that are represented.

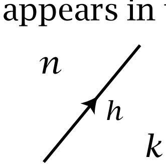
Proof: If n is represented in discriminant Δ then there is a form $nx^2+bxy+cy^2$ of discriminant Δ. If n factors as $n=n_1n_2$ then n_1 is represented by the form $n_1x^2+bxy+n_2cy^2$ which has the same discriminant as $nx^2+bxy+cy^2$. □

There is a simple congruence criterion for when a number is represented in a given discriminant:

Proposition 6.2. *There exists a form of discriminant Δ that represents n if and only if Δ is congruent to a square mod $4n$.*

Note that if n is negative then "mod $4n$" means the same thing as "mod $4|n|$" since being divisible by a number d is equivalent to being divisible by $-d$ when we are considering both positive and negative numbers.

Proof: Suppose n is represented by a form Q of discriminant Δ, so n appears in the topograph of Q. If we look at an edge of the topograph bordering a region labeled n then we obtain an equation $\Delta = h^2 - 4nk$ where h is the label on the edge and k is the label on the region on the opposite



side of this edge. The equation $\Delta = h^2 - 4nk$ implies the congruence $\Delta \equiv h^2 \bmod 4n$ so Δ is a square mod $4n$.

Conversely, suppose that Δ is the square of some integer h mod $4n$. This means that $h^2 - \Delta$ is an integer times $4n$, or in other words $h^2 - \Delta = 4nk$ for some k. This equation can be rewritten as $\Delta = h^2 - 4nk$, so the form $nx^2 + hxy + ky^2$ has discriminant Δ, and this form represents n when $(x,y) = (1,0)$. □

Let us see what this proposition implies about representing small numbers n. For $n = 1$ it says that there is a form of discriminant Δ representing 1 if and only if Δ is a square mod 4. The squares mod 4 are 0 and 1, and we already know that discriminants of forms are always congruent to 0 or 1 mod 4. So we conclude that for every possible value of the discriminant there exists a form that represents 1. This is not new information, however, since the principal forms x^2+dy^2 and $x^2+xy+dy^2$ represent 1 and there is a principal form in each discriminant.

In the next case $n = 2$ the possible values of the discriminant mod $4n = 8$ are $0,1,4,5$, and the squares mod 8 are $0,1,4$ since $0^2 = 0$, $(\pm 1)^2 = 1$, $(\pm 2)^2 = 4$, $(\pm 3)^2 \equiv 1$, and $(\pm 4)^2 \equiv 0$. Thus 2 is not represented by any form of discriminant Δ when $\Delta \equiv 5 \bmod 8$, but for all other discriminants there is a form representing 2. Explicit forms representing 2 are $2x^2 - ky^2$ for $\Delta = 8k$, $2x^2 + xy - ky^2$ for $\Delta = 8k+1$, and $2x^2 + 2xy - ky^2$ for $\Delta = 8k + 4$.

Moving on to the next case $n = 3$, the discriminants mod 12 are $0,1,4,5,8,9$ and the squares mod 12 are $0,1,4,9$ since $0^2 = 0$, $(\pm 1)^2 = 1$, $(\pm 2)^2 = 4$, $(\pm 3)^2 = 9$, $(\pm 4)^2 \equiv 4$, $(\pm 5)^2 \equiv 1$, and $(\pm 6)^2 \equiv 0$. The excluded discriminants are thus those congruent to 5 or 8 mod 12. Again explicit forms are easily given, the forms $3x^2 + hxy - ky^2$ with $\Delta = 12k + h^2$ for $h = 0,1,2,3$.

We could continue in this direction, exploring which discriminants have forms that represent a given number, but this is not really the question we want to answer, which is to start with a given discriminant and decide which numbers are represented in this discriminant. The sort of answer we are looking for, based on the various examples we looked at earlier, is also a different sort of congruence condition, with congruence modulo the discriminant rather than congruence mod $4n$. So there is more work to be done before we would have the sort of answer we want. Nevertheless, the representability criterion in Proposition 6.2 is the starting point.

Our approach will be to reduce the representation problem in discriminant Δ first to the case of representing prime powers and then to representing primes themselves. Here is the first step.

Proposition 6.3. *If two coprime numbers m and n are both represented in discriminant Δ then so is their product mn.*

Applying this repeatedly, we see that if a number n has the prime factorization $n = p_1^{e_1} \cdots p_k^{e_k}$ for distinct primes p_i, and if $p_i^{e_i}$ is represented in discriminant Δ for each i, then n is represented in discriminant Δ.

The main ingredient in the proof of the proposition will be the following:

Lemma 6.4. *If a number x is a square mod m_1 and also a square mod m_2 where m_1 and m_2 are coprime, then x is a square mod $m_1 m_2$.*

For example, the number 2 is a square mod 7 (since $3^2 \equiv 2 \bmod 7$) and also mod 17 (since $6^2 \equiv 2 \bmod 17$) so 2 must also be a square mod $7 \cdot 17 = 119$. And in fact $2 \equiv 11^2 \bmod 119$.

Proof: This will be a consequence of the Chinese Remainder Theorem. If x is a square mod m_1 and also a square mod m_2 then there are numbers a_1 and a_2 such that $x \equiv a_1^2 \bmod m_1$ and $x \equiv a_2^2 \bmod m_2$. If m_1 and m_2 are coprime then by the Chinese Remainder Theorem there is a number a that is congruent to $a_1 \bmod m_1$ and to $a_2 \bmod m_2$, hence $a^2 \equiv a_1^2 \bmod m_1$ and $a^2 \equiv a_2^2 \bmod m_2$. Thus $x \equiv a^2 \bmod m_1$ and mod m_2. This implies $x \equiv a^2 \bmod m_1 m_2$ since the difference $x - a^2$ is divisible by both m_1 and m_2 and hence by their product $m_1 m_2$ since m_1 and m_2 are coprime. This shows that x is a square mod $m_1 m_2$. □

Proof of Proposition 6.3: Let m and n be coprime. At least one of them must be odd, say n is odd. If m and n are represented in discriminant Δ then Δ is a square mod $4m$ and mod $4n$, hence also mod n. Since $4m$ and n are coprime, the lemma then says that Δ is a square mod $4mn$, so mn is represented in discriminant Δ. □

Next we try to reduce further from prime powers to primes themselves. This is possible for most primes by the following more technical result:

Lemma 6.5. *If a number x is a square mod p for an odd prime p not dividing x, then x is also a square mod p^r for each $r > 1$. The corresponding statement for the prime $p = 2$ is that if an odd number x is a square mod 8 then x is also a square mod 2^r for each $r > 3$.*

For example, 2 is a square mod 7 since $2 \equiv 3^2 \bmod 7$, so 2 is also a square mod 7^2, namely $2 \equiv 10^2 \bmod 49$. It is also a square mod $7^3 = 343$ since $2 \equiv 108^2 \bmod 343$. Likewise it must be a square mod 7^4, mod 7^5, etc. The proof of the lemma will give a method for refining the initial congruence $2 \equiv 3^2 \bmod 7$ to each subsequent congruence $2 \equiv 10^2 \bmod 49$, $2 \equiv 108^2 \bmod 343$, etc.

For the prime $p = 2$ we have to begin with squares mod 8 since 3 is a square mod 2 but not mod 4, while 5 is a square mod 4 but not mod 8.

Proof of Lemma 6.5: We will show that if x is a square mod p^r then it is also a square mod p^{r+1}, assuming $r \geq 1$ in the case that p is odd and $r \geq 3$ in the case $p = 2$. By induction this will prove the lemma.

We begin by assuming that x is a square mod p^r, so there is a number y such that $x \equiv y^2 \bmod p^r$ or in other words p^r divides $x - y^2$, say $x - y^2 = p^r l$ for some integer l. We would like to find a number z such that $x \equiv z^2 \bmod p^{r+1}$, so it is reasonable to look for a z with $z \equiv y \bmod p^r$, or in other words $z = y + kp^r$ for some k. Thus we want to choose k so that $x \equiv (y + kp^r)^2 \bmod p^{r+1}$. In other words we want p^{r+1} to divide $x - (y + kp^r)^2$. This can be rewritten as:

$$\begin{aligned} x - (y + kp^r)^2 &= x - (y^2 + 2kp^r y + k^2 p^{2r}) \\ &= x - y^2 - 2kp^r y - k^2 p^{2r} \\ &= p^r l - 2kp^r y - k^2 p^{2r} \quad \text{since } x - y^2 = p^r l \\ &= p^r (l - 2ky - k^2 p^r) \end{aligned}$$

For this to be divisible by p^{r+1} means that p should divide $l - 2ky - k^2 p^r$. Since we assume $r \geq 1$ this is equivalent to p dividing $l - 2ky$, or in other words, $l - 2ky = pq$ for some integer q. Rewriting this as $l = 2yk + pq$, we see that this linear Diophantine equation with unknowns k and q always has a solution when p is odd since $2y$ and p are coprime if p is odd, in view of the fact that p does not divide y since $x \equiv y^2 \bmod p^r$ and we assume x is not divisible by p. This finishes the induction step in the case that p is odd.

When $p = 2$ this argument breaks down at the last step since the equation $l = 2yk + pq$ becomes $l = 2yk + 2q$ and this will not have a solution when l is odd. To modify the proof so that it works for $p = 2$ we would like to get rid of the factor 2 in the equation $l = 2yk + pq$ which arose when we squared $y + kp^r$. To do this, suppose that instead of trying $z = y + k \cdot 2^r$ we try $z = y + k \cdot 2^{r-1}$. Then we would want 2^{r+1} to divide $x - (y + k \cdot 2^{r-1})^2$. Again this can be rewritten:

$$\begin{aligned} x - (y + k \cdot 2^{r-1})^2 &= x - y^2 - k \cdot 2^r y - k^2 2^{2r-2} \\ &= 2^r l - k \cdot 2^r y - k^2 2^{2r-2} \quad \text{since } x - y^2 = 2^r l \\ &= 2^r (l - ky - k^2 2^{r-2}) \end{aligned}$$

Assuming $r \geq 3$, this means 2 should divide $l - ky$, or in other words $l = yk + 2q$ for some integer q. The number y is odd since $y^2 \equiv x \bmod 2^r$ and x is odd by assumption. This implies the equation $l = yk + 2q$ has a solution (k, q). □

Proposition 6.6. *If a prime p not dividing the discriminant Δ is represented by a form of discriminant Δ then every power of p is also represented by a form of discriminant Δ.*

Proof: First we consider odd primes p. If p is represented in discriminant Δ then Δ is a square mod $4p$ and hence mod p. The preceding lemma then says that Δ is a square mod each power p^r. From this it follows by Lemma 6.4 that Δ is also a square mod $4p^r$ since Δ is always a square mod 4. Thus by Proposition 6.2 all powers of p are represented in discriminant Δ.

For $p = 2$ the argument is almost the same. In this case the representability of 2 implies that Δ is a square mod $4 \cdot 2 = 8$ so the lemma implies that Δ is also a square mod $4 \cdot 2^r$ for all $r \geq 1$ so all powers of 2 are represented. □

In the examples for the representation problem that we looked at in the preceding section we saw that primes that divide the discriminant behave differently from primes that do not, and the differences begin at this point:

Proposition 6.7. *Each prime dividing the discriminant Δ is represented in discriminant Δ. If a prime p divides Δ but not the conductor of Δ then no form of discriminant Δ represents p^2 or any higher power of p.*

Recall that the conductor for discriminant Δ is the largest positive number d such that $\Delta = d^2\Delta'$ for some discriminant Δ'. This Δ' is then a fundamental discriminant. Fundamental discriminants are those with conductor 1.

Proof: The representability of primes dividing Δ follows from Proposition 5.7, but it can also be deduced from the congruence criterion of Proposition 6.2 as follows. For a prime p dividing Δ we have $\Delta \equiv 0 \bmod p$ so Δ is a square mod p, namely 0^2. When p is odd it follows that Δ is also a square mod $4p$ since Δ is always a square mod 4. Hence p is represented in discriminant Δ in this case. If p is 2 and divides Δ then $\Delta \equiv 0 \bmod 4$ so $\Delta = 8k$ or $8k+4$. Thus $\Delta \equiv 0$ or $4 \bmod 8$ and so Δ is a square mod 8, which means that 2 is represented in discriminant Δ.

Suppose now that p is a prime dividing Δ and some form of discriminant Δ represents p^2. This form is equivalent to a form $p^2x^2 + bxy + cy^2$ with p dividing $\Delta = b^2 - 4p^2c$ so p must divide b^2. Since p is prime it must then divide b, so in fact p^2 divides b^2. Therefore p^2 divides $\Delta = b^2 - 4p^2c$ and we have $\Delta = p^2\Delta'$ for some integer Δ'.

Consider first the case that p is odd. Then $p^2 \equiv 1 \bmod 4$ so $\Delta \equiv \Delta' \bmod 4$. This means that Δ' is also a discriminant, so by the definition of the conductor, p divides the conductor. Thus if p divides Δ but not the conductor then p^2 cannot be represented by any form of discriminant Δ.

In the case that $p = 2$ the assumption that p divides Δ means that Δ is even and hence so is b. The discriminant equation $\Delta = b^2 - 4p^2c$ is now $\Delta = b^2 - 4\cdot 2^2c$ so $\Delta \equiv b^2 \bmod 16$. The only squares of even numbers mod 16 are 0 and 4, as one sees by checking 0^2, $(\pm 2)^2$, $(\pm 4)^2$, $(\pm 6)^2$, and $(\pm 8)^2$, so Δ is either $16k = 4(4k)$ or $16k+4 = 4(4k+1)$. In both cases Δ is 4 times a discriminant so 2 divides the conductor.

Once we know that p^2 is not represented in discriminant Δ then neither is any multiple of p^2 by Proposition 6.1, and in particular higher powers of p are not represented. □

Here is a summary of what we have shown so far in the case of fundamental discriminants:

Theorem 6.8. *If Δ is a fundamental discriminant then the numbers $n > 1$ that are represented by at least one form of discriminant Δ are exactly the numbers* *that factor as a product* $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ *of powers of distinct primes* p_i *each of which is represented by some form of discriminant* Δ*, with the restriction that* $e_i \le 1$ *for primes* p_i *dividing* Δ*.*

The situation for nonfundamental discriminants is more complicated and will be described later in Theorem 6.11.

Quadratic Reciprocity

For the problem of determining which primes are represented in a given discriminant we already know when 2 is represented and we know that primes dividing the discriminant are always represented. After these special cases what remains are the odd primes not dividing the discriminant, which can be regarded as the generic case.

An odd prime p will be represented in discriminant Δ exactly when Δ is a square mod p. Let us introduce some convenient notation for this condition. For p an odd prime and a an integer not divisible by p, define the ***Legendre symbol*** $\left(\frac{a}{p}\right)$ by

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & \text{if } a \text{ is a square mod } p \\ -1 & \text{if } a \text{ is not a square mod } p \end{cases}$$

Using this notation we can say:

- *An odd prime* p *that does not divide* Δ *is represented in discriminant* Δ *if and only if* $\left(\frac{\Delta}{p}\right) = +1$*.*

It will therefore be useful to know how to compute $\left(\frac{a}{p}\right)$. The following four basic properties of the Legendre symbol make this a feasible task:

(1) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

(2) $\left(\frac{-1}{p}\right) = +1$ if $p \equiv 1$ mod 4 and $\left(\frac{-1}{p}\right) = -1$ if $p \equiv 3$ mod 4.

(3) $\left(\frac{2}{p}\right) = +1$ if $p \equiv \pm 1$ mod 8 and $\left(\frac{2}{p}\right) = -1$ if $p \equiv \pm 3$ mod 8.

(4) If p and q are distinct odd primes then $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ unless p and q are both congruent to 3 mod 4, in which case $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$.

Property (1), applied repeatedly, reduces the calculation of $\left(\frac{a}{p}\right)$ to the calculation of $\left(\frac{q}{p}\right)$ for the various prime factors q of a, along with $\left(\frac{-1}{p}\right)$ when a is negative. Note that $\left(\frac{q^2}{p}\right) = +1$ so we can immediately reduce to the case that $|a|$ is a product of distinct primes. Property (2) will be used when dealing with negative discriminants, and property (3) will be used for certain even discriminants.

Property (4) is called ***quadratic reciprocity***. This is by far the most subtle of the four properties, and proving it is considerably more difficult than for the other three properties. We will give a proof in Section 6.4, obtaining proofs of the first three properties along the way.

For a quick illustration of the usefulness of these properties let us see how they can be used to compute the values of Legendre symbols. Suppose for example that one wanted to know whether 78 was a square mod 89. The naive approach would be to list the squares of all the numbers $\pm 1, \cdots, \pm 44$ and see whether any of these was congruent to 78 mod 89, but this would be rather tedious. Since 89 is prime we can instead evaluate $\left(\frac{78}{89}\right)$ using the basic properties of Legendre symbols. First we factor 78 to get $\left(\frac{78}{89}\right) = \left(\frac{2}{89}\right)\left(\frac{3}{89}\right)\left(\frac{13}{89}\right)$. By property (3) we have $\left(\frac{2}{89}\right) = +1$ since $89 \equiv 1$ mod 8. Next, reciprocity gives $\left(\frac{3}{89}\right) = \left(\frac{89}{3}\right)$ and $\left(\frac{13}{89}\right) = \left(\frac{89}{13}\right)$ since $89 \equiv 1$ mod 4. After this we use the fact that $\left(\frac{a}{p}\right)$ depends only on the value of a mod p to reduce $\left(\frac{89}{3}\right)$ to $\left(\frac{2}{3}\right)$ and $\left(\frac{89}{13}\right)$ to $\left(\frac{11}{13}\right)$. Using property (3) again, we have $\left(\frac{2}{3}\right) = -1$, confirming the obvious fact that 2 is not a square mod 3. For $\left(\frac{11}{13}\right)$, reciprocity says this equals $\left(\frac{13}{11}\right)$. This reduces to $\left(\frac{2}{11}\right) = -1$. Summarizing, we have:

$$\left(\frac{78}{89}\right) = \left(\frac{2}{89}\right)\left(\frac{3}{89}\right)\left(\frac{13}{89}\right) = (+1)(-1)(-1) = +1$$

Thus we see that 78 is a square mod 89, even though we have not found an actual number x such that $x^2 \equiv 78$ mod 89.

In this example we used the fact that the modulus 89 was prime, but we have already seen how to reduce to the case of prime moduli. For example, if we wanted to determine whether 78 is a square mod 88 we know this is the case exactly when it is a square mod 8 and mod 11. The squares mod 8 are 0, 1, and 4 whereas $78 \equiv 6$ mod 8 so 78 is not a square mod 8 and therefore not mod 88 either, even though $78 \equiv 1$ mod 11 so 78 is a square mod 11.

Returning now to quadratic forms, let us see what the basic properties of Legendre symbols tell us about which primes are represented by some of the forms discussed at the beginning of the chapter. In the first four cases the class number is 1 so we will be determining which primes are represented by the given form, and Theorem 6.8 will then say exactly which numbers are represented by this form, confirming the conjectures made when we looked at the topographs.

Example: $x^2 + y^2$ with $\Delta = -4$. This form obviously represents 2, the only prime dividing Δ, and it represents an odd prime p exactly when $\left(\frac{-4}{p}\right) = +1$. Using the first of the four properties we have $\left(\frac{-4}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{2}{p}\right) = \left(\frac{-1}{p}\right)$, and the second property says this is $+1$ exactly for primes $p = 4k + 1$. Thus we see the primes represented by $x^2 + y^2$ are 2 and the primes $p = 4k + 1$.

Example: $x^2 + 2y^2$ with $\Delta = -8$. Again the only prime dividing Δ is 2, and it is represented. For odd primes p we have $\left(\frac{-8}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)^3 = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)$. In the four cases $p \equiv 1, 3, 5, 7$ mod 8 this is, respectively, $(+1)(+1)$, $(-1)(-1)$, $(+1)(-1)$, and $(-1)(+1)$. We conclude that the primes represented by the form $x^2 + 2y^2$ are 2 and primes congruent to 1 or 3 mod 8.

Example: $x^2 - 2y^2$ with $\Delta = 8$. The only prime dividing Δ is 2 which is represented when $(x, y) = (2, 1)$. For odd primes p we have $\left(\frac{8}{p}\right) = \left(\frac{2}{p}\right)^3 = \left(\frac{2}{p}\right)$ so property (3) implies that the primes represented by $x^2 - 2y^2$ are 2 and $p \equiv \pm 1$ mod 8.

Example: x^2+xy+y^2 with $\Delta=-3$. The only prime dividing the discriminant is 3 and it is represented. The prime 2 is not represented since $\Delta\equiv 5 \bmod 8$. For primes $p>3$ we can evaluate $\left(\frac{-3}{p}\right)$ using quadratic reciprocity:

$$\left(\frac{-3}{p}\right)=\left(\frac{-1}{p}\right)\left(\frac{3}{p}\right)=\begin{cases}(+1)\left(\frac{p}{3}\right) & \text{if } p=4k+1\\ (-1)(-\left(\frac{p}{3}\right)) & \text{if } p=4k+3\end{cases}$$

So we get $\left(\frac{p}{3}\right)$ in both cases. Since $\left(\frac{p}{3}\right)$ only depends on p mod 3, we have $\left(\frac{p}{3}\right)=+1$ if $p\equiv 1 \bmod 3$ and $\left(\frac{p}{3}\right)=-1$ if $p\equiv 2 \bmod 3$. (Since $p\neq 3$ we do not need to consider the possibility $p\equiv 0 \bmod 3$.) The conclusion is that the primes represented by x^2+xy+y^2 are 3 and the primes $p\equiv 1 \bmod 3$.

Example: $\Delta=40$. Here all forms are equivalent to either x^2-10y^2 or $2x^2-5y^2$. The primes dividing 40 are 2 and 5 so these are represented by one form or the other, and in fact both are represented by $2x^2-5y^2$ as the topographs showed. For other primes p we have $\left(\frac{40}{p}\right)=\left(\frac{2}{p}\right)^3\left(\frac{5}{p}\right)=\left(\frac{2}{p}\right)\left(\frac{p}{5}\right)$. The factor $\left(\frac{2}{p}\right)$ depends only on p mod 8 and $\left(\frac{p}{5}\right)$ depends only on p mod 5, so their product depends only on p mod 40. The following table lists all the possibilities for congruence classes mod 40 not divisible by 2 or 5:

	1	3	7	9	11	13	17	19	21	23	27	29	31	33	37	39
$\left(\frac{2}{p}\right)$	+1	−1	+1	+1	−1	−1	+1	−1	−1	+1	−1	−1	+1	+1	−1	+1
$\left(\frac{p}{5}\right)$	+1	−1	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1

The product $\left(\frac{2}{p}\right)\left(\frac{p}{5}\right)$ is $+1$ in exactly the eight cases $p\equiv 1,3,9,13,27,31,37,39$ mod 40. We conclude that these are the eight congruence classes containing primes (other than 2 and 5) represented by one of the two forms x^2-10y^2 and $2x^2-5y^2$. This agrees with our earlier observations based on the topographs. However, we have yet to verify our earlier guesses as to which congruence classes are represented by which form. We will see how to do this in the next section.

In the examples above we were able to express $\left(\frac{\Delta}{p}\right)$ in terms of Legendre symbols $\left(\frac{-1}{p}\right)$, $\left(\frac{2}{p}\right)$, and $\left(\frac{p}{p_i}\right)$ for odd primes p_i dividing Δ. The following result shows that this can be done for all Δ:

Proposition 6.9. *Let the nonzero integer Δ be factored as $\Delta=\varepsilon 2^s p_1\cdots p_k$ for $\varepsilon=\pm1$, $s\geq 0$, and each p_i an odd prime. (We allow $k=0$ when $\Delta=\varepsilon 2^s$.) Then for odd primes p not dividing Δ the Legendre symbol $\left(\frac{\Delta}{p}\right)$ has the value given in the following table*:

Δ	$\left(\frac{\Delta}{p}\right)$
$2^{2l}(4m+1)$	$\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$
$2^{2l}(4m+3)$	$\left(\frac{-1}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$
$2^{2l+1}(4m+1)$	$\left(\frac{2}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$
$2^{2l+1}(4m+3)$	$\left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$

Proof: For $\Delta = \varepsilon 2^s p_1 \cdots p_k$ quadratic reciprocity gives

$$\left(\frac{\Delta}{p}\right) = \left(\frac{\varepsilon}{p}\right)\left(\frac{2}{p}\right)^s\left(\frac{p_1}{p}\right)\cdots\left(\frac{p_k}{p}\right) = \left(\frac{\varepsilon}{p}\right)\left(\frac{2}{p}\right)^s\left(\frac{\omega}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$$

where ω is $+1$ or -1 according to whether there are an even or an odd number of factors $p_i \equiv 3 \bmod 4$. The exponent s in this formula can be replaced by 0 or 1 according to whether s is even or odd. In the first and third rows of the table the odd part of Δ is $4m+1$ so we have $\varepsilon = \omega$ and therefore $\left(\frac{\varepsilon}{p}\right)\left(\frac{\omega}{p}\right) = 1$. In the second and fourth rows the factor $4m+1$ is replaced by $4m+3$ and we have $\varepsilon = -\omega$, hence $\left(\frac{\varepsilon}{p}\right)\left(\frac{\omega}{p}\right) = \left(\frac{-1}{p}\right)$. □

Corollary 6.10. *The representability of an odd prime p in discriminant Δ depends only on the congruence class of p mod Δ.*

Proof: The class of p mod Δ determines its class mod p_i for each i and this determines $\left(\frac{p}{p_i}\right)$. For the terms $\left(\frac{-1}{p}\right)$ and $\left(\frac{2}{p}\right)$ in the last three rows of the table, note first that l must be at least 1 in these rows since Δ is a discriminant. In the second row the class of p mod Δ determines its class mod 4 so it determines $\left(\frac{-1}{p}\right)$. In the third and fourth rows the class of p mod Δ determines its class mod 8 so both $\left(\frac{-1}{p}\right)$ and $\left(\frac{2}{p}\right)$ are determined. Thus in all cases the factors of $\left(\frac{\Delta}{p}\right)$ are determined by the class of p mod Δ so $\left(\frac{\Delta}{p}\right)$ is determined. □

Complications for Nonfundamental Discriminants

Our next result generalizes Theorem 6.8 to cover all discriminants. As one can see, the general statement is considerably more complicated than for fundamental discriminants.

Theorem 6.11. *A number $n > 1$ is represented by at least one form of discriminant Δ exactly when n factors as a product $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ of powers of distinct primes p_i each of which is represented by some form of discriminant Δ, where $e_i \le 1$ for primes p_i dividing Δ but not the conductor, while for primes $p = p_i$ dividing the conductor the allowed exponents $e = e_i$ are given by the following rules. First write $\Delta = p^s q$ with p^s the highest power of p dividing Δ. Then if p is odd the allowable exponents e are those for which either*

(a) $e \le s$ *or*

(b) $e > s$, s *is even, and* $\left(\frac{q}{p}\right) = +1$.

If $p = 2$ then the allowable exponents e are those for which either

(a) $e \le s - 2$ *or*

(b) s *is even and* e *is as in the following table:*

q mod 8	1	3	5	7
e	all	$\le s-1$	$\le s$	$\le s-1$

Examples will be given following the proof. The main part of the proof is contained in a lemma:

Lemma 6.12. *Suppose that a number x divisible by a prime p factors as $p^s q$ where p does not divide q, so p^s is the largest power of p dividing x. Then*:

(a) *x is a square mod p^r for each $r \le s$.*

(b) *If $r > s$ and s is odd then x is not a square mod p^r.*

(c) *If $r > s$ and s is even then x is a square mod p^r if and only if q is a square mod p^{r-s}.*

Proof: Part (a) is easy since x is 0 mod p^s hence also mod p^r if $r \le s$, and 0 is always a square mod anything.

For (b) we assume $r > s$ and s is odd. Suppose $p^s q$ is a square mod p^r, so $p^s q = y^2 + lp^r$ for some integers y and l. Then p^s divides $y^2 + lp^r$ and it divides lp^r (since $r > s$) so p^s divides y^2. Since s is assumed to be odd and the exponent of p in y^2 must be even, this implies p^{s+1} divides y^2. It also divides lp^r since $s+1 \le r$, so from the equation $p^s q = y^2 + lp^r$ we conclude that p divides q, contrary to the definition of q. This contradiction shows that $p^s q$ is not a square mod p^r when $r > s$ and s is odd, so statement (b) is proved.

For (c) we assume $r > s$ and s is even. As in part (b), if $p^s q$ is a square mod p^r we have an equation $p^s q = y^2 + lp^r$ and this implies that p^s divides y^2. Since s is now even, this means $y^2 = p^s z^2$ for some number z. Canceling p^s from $p^s q = y^2 + lp^r$ yields an equation $q = z^2 + lp^{r-s}$, which says that q is a square mod p^{r-s}. Conversely, if q is a square mod p^{r-s} we have an equation $q = z^2 + lp^{r-s}$ and hence $p^s q = p^s z^2 + lp^r$. Since s is even, this says that $p^s q$ is a square mod p^r. □

Proof of Theorem 6.11: As in the proof of Theorem 6.8 the question reduces to representing powers of primes. We know from Proposition 6.6 that all powers of a prime not dividing the discriminant Δ are represented if the prime itself is represented. By Proposition 6.7 we also know that primes p dividing Δ are represented, and their powers p^e with $e > 1$ cannot be represented unless p divides the conductor. For the remaining case of primes dividing the conductor we will apply the preceding lemma with $x = \Delta$.

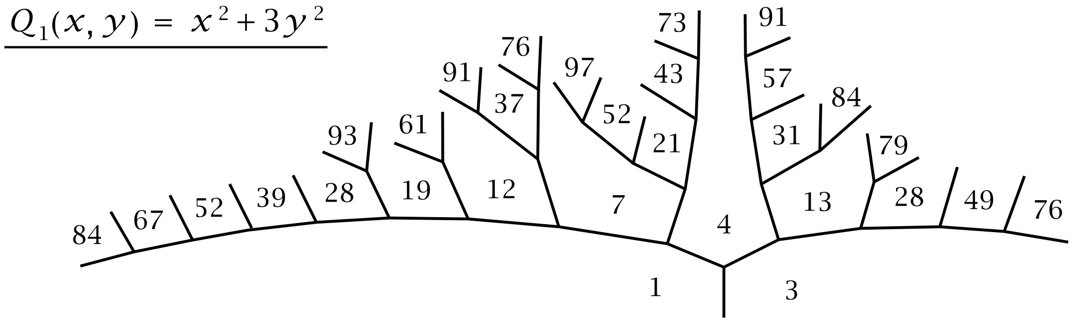
For odd p dividing Δ we need to determine when Δ is a square mod p^e. By the lemma the times this happens are when $e \le s$, or when $e > s$ and s is even and q is a square mod p^{e-s}. When $e > s$ this last condition amounts just to q being a square mod p by Lemma 6.5, or in other words $\left(\frac{q}{p}\right) = +1$.

When $p = 2$ we need to determine when Δ is a square mod $4 \cdot 2^e = 2^{e+2}$. By the lemma this happens only when $e \le s - 2$ or when s is even and q (which is odd) is a square mod 2^{e+2-s}. If $e = s - 1$ then $e + 2 - s = 1$ and every q is a square mod $2^{e+2-s} = 2$. If $e = s$ then $e + 2 - s = 2$ and q is a square mod $2^{e+2-s} = 4$ only when

$q = 4k + 1$. And if $e \geq s + 1$ then $e + 2 - s \geq 3$ and q is a square mod 2^{e+2-s} only when it is a square mod 8, which means $q = 8k + 1$. ☐

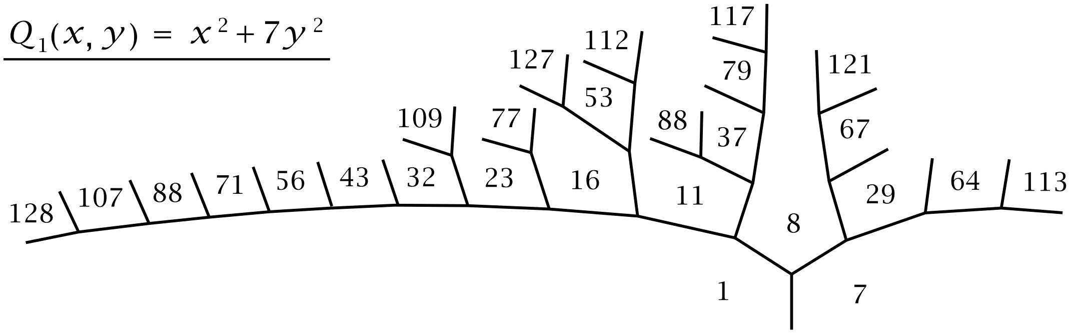
Let us look at two examples illustrating some of the more subtle possibilities in the preceding theorem. The examples involve the rather simple forms $x^2 + ny^2$ whose discriminant $-4n$ is sometimes not a fundamental discriminant such as when n is congruent to 3 mod 4. The examples will be the cases $n = 3, 7$.

Example: $\Delta = -12$ with conductor 2. The two forms here are $Q_1 = x^2 + 3y^2$ and the nonprimitive form $Q_2 = 2x^2 + 2xy + 2y^2$.



The primes represented in discriminant -12 are 2, 3, and primes p with $\left(\frac{-12}{p}\right) = \left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right) = +1$, so these are the primes $p \equiv 1$ mod 3. By Theorem 6.11 the numbers represented in discriminant -12 are the numbers $n = 2^a 3^b p_1 \cdots p_k$ with $a \leq 2$, $b \leq 1$, and each p_i a prime congruent to 1 mod 3. (When we apply the theorem for $p_i = 2$ we have $s = 2$ and $q = -3$.) We can in fact determine which of Q_1 and Q_2 is giving these representations. The form Q_2 is twice $x^2 + xy + y^2$ and we have already determined which numbers the latter form represents, namely the products $3^b p_1 \cdots p_k$ with $b \leq 1$ and each prime $p_i \equiv 1$ mod 3. Thus, of the numbers represented by Q_1 or Q_2, the numbers represented by Q_2 are those with $a = 1$. None of these numbers with $a = 1$ are represented by Q_1 since $x^2 + 3y^2$ is never 2 mod 4, as x^2 and y^2 must be 0 or 1 mod 4.

Example: $\Delta = -28$ with conductor 2 again. Here the only two forms up to equivalence are $Q_1 = x^2 + 7y^2$ and $Q_2 = 2x^2 + 2xy + 4y^2$ which is not primitive.



The primes represented in discriminant -28 are 2, 7, and odd primes p with $\left(\frac{-28}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{7}{p}\right) = \left(\frac{p}{7}\right) = +1$ so $p \equiv 1, 2, 4$ mod 7. According to Theorem 6.11 the numbers represented by Q_1 or Q_2 are the numbers $n = 2^a 7^b p_1 \cdots p_k$ with $b \le 1$ and each p_i an odd prime congruent to 1, 2, or 4 mod 7. There is no restriction on a since when we apply the theorem with $p_i = 2$ we have $s = 2$ and $q = -7 = 8l + 1$.

We can say exactly which numbers are represented by Q_2 since it is twice the form $x^2 + xy + 2y^2$ of discriminant -7, which is a fundamental discriminant of class number 1 so Theorem 6.8 tells us which numbers this form represents. These are the numbers $7^b p_1 \cdots p_k$ with $b \le 1$ and primes $p_i \equiv 1,2,4$ mod 7, including now the possibility $p_i = 2$. Thus Q_2 represents exactly the numbers $2^a 7^b p_1 \cdots p_k$ with $a \ge 1$, $b \le 1$ and odd primes $p_i \equiv 1,2,4$ mod 7. Hence Q_1 must represent at least the numbers $2^a 7^b p_1 \cdots p_k$ with $a = 0$, $b \le 1$, and odd primes $p_i \equiv 1,2,4$ mod 7. These numbers are all odd since $a = 0$, but Q_1 also represents some even numbers since $x^2 + 7y^2$ is even whenever both x and y are odd.

From the topograph we might conjecture that Q_1 represents exactly the numbers $2^a 7^b p_1 \cdots p_k$ with $a \ne 1,2$ and the same conditions on b and the primes p_i as before. For example one can see that 8, 16, 32, 64, and 128 are represented. It is not difficult to exclude $a = 1$ and $a = 2$ by considering the values of $x^2 + 7y^2$ mod 4 and mod 8. To see that Q_1 represents all the predicted numbers with $a \ge 3$ we use the following result.

Proposition 6.13. *For a prime p, if a product $p^k q$ with $k > 0$ is represented by a primitive form of discriminant Δ then $p^{k+2} q$ is represented by a primitive form of discriminant $p^2\Delta$.*

Applying this to the case at hand with $p = 2$, the form $x^2 + xy + 2y^2$ represents all the products $2^a 7^b p_1 \cdots p_k$ as above with $a \ge 1$, so $x^2 + 7y^2$ represents all these products with $a \ge 3$.

Proof: Suppose we have a primitive form of discriminant Δ representing $p^k q$, so the topograph of this form has a region labeled $p^k q$. If $k > 0$ then at least one of the regions adjacent to this region must have a label not divisible by p, otherwise a vertex in the boundary of this region would have all three adjacent labels divisible by p so the form would be p times another form, making it nonprimitive. Thus the given form is equivalent to a form $p^k q x^2 + bxy + cy^2$ with c not divisible by p. The form $p^{k+2} q x^2 + pbxy + cy^2$ has discriminant $p^2\Delta$ and is primitive since its coefficients are not all divisible by p, nor are they divisible by any other prime since such a prime would have to divide q, b, and c making the previous form $p^k q x^2 + bxy + cy^2$ nonprimitive. $\square$

For nonfundamental discriminants Theorem 6.11 says nothing about whether the representing forms are primitive. As we will see in Theorem 7.7, determining the numbers represented by primitive forms of a given discriminant also reduces to the special case of representing prime powers by primitive forms. Namely, a product of powers $p_i^{k_i}$ of distinct primes p_i is represented by a primitive form exactly when each of the prime powers $p_i^{k_i}$ is represented by a primitive form. Most prime powers are represented only by primitive forms, according to the following easy result:

Proposition 6.14. *A form of discriminant Δ representing a power p^k of a prime p not dividing the conductor of Δ is primitive.*

Proof: If a form Q representing p^k is not primitive it is a multiple of another form by some integer $d > 1$. This number d divides every number represented by Q so in particular d divides p^k and hence p divides d. Since d divides the conductor, this means that p divides the conductor. Thus if p does not divide the conductor then Q must be primitive. □

For primes dividing the conductor one can get some idea of the complications that can occur from the table on the next page. This lists all the equivalence classes of forms, both primitive and nonprimitive, for nonfundamental negative discriminants up to -99, along with the prime powers p^k represented by these forms for primes p dividing the conductor d. To save space the table uses the abbreviated notation $[a,b,c]$ for the form $ax^2+bxy+cy^2$.

Some information in the table can be deduced from the earlier Proposition 6.13, such as the fact that if nonprimitive forms of a given discriminant represent all powers p^k with $k \geq 1$ then primitive forms of that discriminant represent all powers p^k with $k \geq 3$. This statement is optimal for some discriminants such as -28 and -60 but not for others such as -72 and -99 where p^2 is also represented by a primitive form.

In the table one can see that primitive forms represent powers of primes dividing the conductor but not these primes themselves. As we will show in Proposition 6.15, a prime can only be represented by a single equivalence class of forms of a given discriminant, and a prime p dividing the conductor for discriminant Δ is represented by p times the principal form of discriminant Δ/p^2, so p is represented by a nonprimitive form and hence cannot also be represented by a primitive form. The uniqueness of forms representing primes holds also for powers of primes that do not divide the conductor, but we see from the table that this uniqueness may not hold for primes that do divide the conductor, even if we restrict attention just to primitive forms, as for example in the case $\Delta = -32$ where 2^3 is represented by two nonequivalent primitive forms, or discriminants -72 and -99 where there are infinitely many different powers p^k represented by different primitive forms.

The entries in the table where Theorem 6.11 says that only finitely many powers p^k are represented can be checked just by drawing topographs, but in the other cases one must use general theory. We already explained the first case $\Delta = -28$ in the earlier analysis of the form $x^2 + 7y^2$. For the next case $\Delta = -60$ the methods in the next section will suffice. A technique for handling the last few cases in the table will be explained at the end of Chapter 8.

Δ	d	Q prim.	p^k	Q nonprim.	p^k
-12	2	$[1,0,3]$	2^2	$2[1,1,1]$	2^1
-16	2	$[1,0,4]$	$2^2, 2^3$	$2[1,0,1]$	$2^1, 2^2$
-27	3	$[1,1,7]$	$3^2, 3^3$	$3[1,1,1]$	$3^1, 3^2$
-28	2	$[1,0,7]$	$2^3, 2^4, 2^5, \cdots$	$2[1,1,2]$	$2^1, 2^2, 2^3, \cdots$
-32	2	$[1,0,8]$	2^3	$2[1,0,2]$	$2^1, 2^2$
		$[3,2,3]$	$2^2, 2^3$		
-36	3	$[1,0,9]$	3^2	$3[1,0,1]$	3^1
		$[2,2,5]$	3^2		
-44	2	$[1,0,11]$	—	$2[1,1,3]$	2^1
		$[3,2,4]$	2^2		
-48	4	$[1,0,12]$	2^4	$2[1,0,3]$	$2^1, 2^3$
		$[3,0,4]$	$2^2, 2^4$	$4[1,1,1]$	2^2
-60	2	$[1,0,15]$	$2^4, 2^6, 2^8, 2^{10}, \cdots$	$2[1,1,4]$	$2^1, 2^3, 2^5, 2^7, \cdots$
		$[3,0,5]$	$2^3, 2^5, 2^7, 2^9, \cdots$	$2[2,1,2]$	$2^2, 2^4, 2^6, 2^8, \cdots$
-63	3	$[1,1,16]$	—	$3[1,1,2]$	3^1
		$[2,1,8]$	3^2		
		$[4,1,4]$	3^2		
-64	4	$[1,0,16]$	$2^4, 2^5$	$2[1,0,4]$	$2^1, 2^3, 2^4$
		$[4,4,5]$	$2^2, 2^4, 2^5$	$4[1,0,1]$	$2^2, 2^3$
-72	3	$[1,0,18]$	$3^3, 3^4, 3^5, 3^6, \cdots$	$3[1,0,2]$	$3^1, 3^2, 3^3, 3^4, \cdots$
		$[2,0,9]$	$3^2, 3^3, 3^4, 3^5, \cdots$		
-75	5	$[1,1,19]$	5^2	$5[1,1,1]$	5^1
		$[3,3,7]$	5^2		
-76	2	$[1,0,19]$	—	$2[1,1,5]$	2^1
		$[4,2,5]$	2^2		
-80	2	$[1,0,20]$	—	$2[1,0,5]$	2^1
		$[4,0,5]$	2^2	$2[2,2,3]$	2^2
		$[3,2,7]$	2^3		
-92	2	$[1,0,23]$	$2^5, 2^8, 2^{11}, 2^{14}, \cdots$	$2[1,1,6]$	$2^1, 2^4, 2^7, 2^{10}, \cdots$
		$[3,2,8]$	$2^3, 2^4, 2^6, 2^7, \cdots$	$2[2,1,3]$	$2^2, 2^3, 2^5, 2^6, 2^8, 2^9, \cdots$
-96	2	$[1,0,24]$	—	$2[1,0,6]$	2^1
		$[3,0,8]$	2^3	$2[2,0,3]$	2^2
		$[5,2,5]$	2^3		
		$[4,4,7]$	2^2		
-99	3	$[1,1,25]$	$3^3, 3^4, 3^5, 3^6, \cdots$	$3[1,1,3]$	$3^1, 3^2, 3^3, 3^4, \cdots$
		$[5,1,5]$	$3^2, 3^3, 3^4, 3^5, \cdots$		

Unique Representability for Primes and Prime Powers

In Section 6.1 we saw examples where two nonequivalent forms of the same discriminant both represent the same number. However, this does not happen for representations of 1 or primes or powers of most primes:

Proposition 6.15. *If Q_1 and Q_2 are two forms of the same discriminant that both represent the same prime p or both represent 1, then Q_1 and Q_2 are equivalent. The same conclusion holds when Q_1 and Q_2 both represent the same power p^k of an odd prime p that does not divide the discriminant.*

The last statement is also true for $p = 2$ but the proof is more difficult so we will wait until the next chapter to deduce this from a more general result, Theorem 7.7. Examples showing that powers of primes dividing the discriminant can be represented by nonequivalent forms of the same discriminant can be found in the table on the previous page. In these examples the prime in question divides the conductor, not just the discriminant, but this has to be the case since for primes p dividing the discriminant but not the conductor the only power p^k represented by a form of the given discriminant is p itself, by Proposition 6.7.

Proof: Suppose that Q is a form representing a number p that is either 1 or a prime. The topograph of Q then has a region labeled p, and we have seen that the h-labels on the edges adjacent to this p-region form an arithmetic progression with increment $2p$ when these edges are all oriented in the same direction. We have the discriminant formula $\Delta = h^2 - 4pq$ where h is the label on one of these edges and q is the value of Q for the region on the other side of this edge. Since p is nonzero the equation $\Delta = h^2 - 4pq$ determines q in terms of Δ and h. This implies that Δ and the arithmetic progression determine the form Q up to equivalence since the progression determines p, and any h-value in the progression then determines the q-value corresponding to this h-value, so Q is equivalent to $px^2 + hxy + qy^2$.

In the case that $p = 1$ the increment in the arithmetic progressions is 2 so the two possible progressions of h-values adjacent to the p-region are the even numbers and the odd numbers. We know that h has the same parity as Δ, so Δ determines which of the two progressions we have. As we saw in the preceding paragraph, this implies that the form is determined by Δ, up to equivalence.

Now we consider the case that p is prime. Let Q_1 and Q_2 be two forms of the same discriminant Δ both representing p. For Q_1 choose an edge in its topograph adjacent to the p-region, with h-label h_1 and q-label q_1. For the form Q_2 we similarly choose an edge with associated labels h_2 and q_2. Both h_1 and h_2 have the same parity as Δ. We have $\Delta = h_1^2 - 4pq_1 = h_2^2 - 4pq_2$ and hence $h_1^2 \equiv h_2^2 \bmod 4p$. This implies $h_1^2 \equiv h_2^2 \bmod p$, so p divides $h_1^2 - h_2^2 = (h_1 + h_2)(h_1 - h_2)$. Since p is prime, it must divide one of the two factors and hence we must have $h_1 \equiv \pm h_2 \bmod p$. By changing the orientations of the edges in the topograph for Q_1 or Q_2 if necessary, we can assume that $h_1 \equiv h_2 \bmod p$.

If p is odd we can improve this congruence to $h_1 \equiv h_2 \bmod 2p$ since we know that $h_1 - h_2$ is divisible by both p and 2 (since h_1 and h_2 have the same parity), hence $h_1 - h_2$ is divisible by $2p$. The congruence $h_1 \equiv h_2 \bmod 2p$ implies that the arithmetic progression of h-values adjacent to the p-region for Q_1 is the same as for Q_2 since $2p$ is the increment for both progressions. By what we showed earlier, this implies that Q_1 and Q_2 are equivalent.

When $p = 2$ this argument needs to be modified slightly. We still have $h_1^2 \equiv h_2^2 \bmod 4p$ so when $p = 2$ this becomes $h_1^2 \equiv h_2^2 \bmod 8$. Since $2p = 4$ the four possible arithmetic progressions of h-values are $h \equiv 0, 1, 2$, or $3 \bmod 4$. We can interchange the possibilities 1 and 3 just by reorienting the edges, leaving only the possibilities $h \equiv 0, 1$, or $2 \bmod 4$. These are distinguished from each other by the congruence $h_1^2 \equiv h_2^2 \bmod 8$ since $(4k)^2 \equiv 0 \bmod 8$, $(4k+1)^2 \equiv 1 \bmod 8$, and $(4k+2)^2 \equiv 4 \bmod 8$.

Finally we have the case that Q_1 and Q_2 both represent the power p^k of an odd prime p not dividing Δ, with $k > 1$. Following the line of proof above we see that p^k divides $h_1^2 - h_2^2 = (h_1 + h_2)(h_1 - h_2)$. If p^k divides either factor we can proceed exactly as before to show that Q_1 and Q_2 are equivalent since we assume p is odd, hence also p^k. If p^k does not divide either factor then both factors are divisible by p, hence p divides their sum $2h_1$. Since p is odd this implies that p divides h_1, and so p divides $\Delta = h_1^2 - 4p^k q_1$. Thus if p does not divide Δ then the case that p^k divides neither $h_1 + h_2$ nor $h_1 - h_2$ does not arise. □

The same argument shows another interesting fact:

Proposition 6.16. *If the topograph of a form has two regions with the same label n where n is either 1, a prime, or a power of an odd prime not dividing the discriminant, then there is a symmetry of the topograph that takes one region labeled n to the other. Similarly, for positive discriminants and for the same numbers n, if there is one region labeled n and another labeled $-n$ then there is a skew symmetry taking one region to the other.*

Proof: Suppose first that there are two regions having the same label n. As we saw in the proof of the preceding proposition, each of these regions is adjacent to an edge with the same label h and hence the labels q across these edges are also the same. This means there is a symmetry taking one region labeled n to the other.

The other case is that one region is labeled n and the other $-n$. The topographs of the given form Q and its negative $-Q$ then each have a region labeled n so there is an equivalence from Q to $-Q$ taking the n-region for Q to the n-region for $-Q$. This equivalence can be regarded as a skew symmetry of Q taking the n-region to the $-n$-region. □

For the last result in this section we will use a variant of Euclid's proof that there are infinitely many primes to prove the following general statement:

Proposition 6.18. *For each discriminant Δ the set of primes represented in discriminant Δ is infinite.*

Proof: In each discriminant Δ there is a form $Q(x,y) = x^2 + bxy + cy^2$ representing 1. We can assume c is nonzero since in the topograph of Q there will always be at least one region adjacent to the 1 region that is not labeled by 0. (Only parabolic and 0-hyperbolic forms can have a 0 region and they have at most two 0 regions.) Let $p_1, \cdots, p_k$ be any finite list of primes. We allow repetitions on this list so we can make k as large as we like just by repeating some p_i often enough. Let P be the product $p_1 \cdots p_k$ and consider the number $n = Q(1,P) = 1 + bP + cP^2$. This is represented by Q since $(1,P)$ is a primitive pair. If k is large enough we will have $|n| > 1$ since $|cP^2|$ will be much larger than $|1 + bP|$. Any prime p dividing n will also be represented by some form of discriminant Δ. This p must be different from any of the primes p_i on the initial list since dividing p_i into $n = 1 + P + cP^2$ gives a remainder of 1, whereas p divides n evenly. Thus we have shown that for any finite list of primes there is another prime not on the list that is represented in discriminant Δ. Hence the set of primes represented in discriminant Δ must be infinite. $\square$

Exercises

1. Determine discriminants Δ for which there exists a quadratic form of discriminant Δ that represents 5, and also the discriminants for which there does not exist a form representing 5. When 5 is represented, find a form that gives the representation.

2. The following is a generalization of Lemma 6.4. Let $P(x)$ be a polynomial with integer coefficients and let n be an integer. Show that if the congruence $P(x) \equiv n$ has a solution mod m_1 and also a solution mod m_2 where m_1 and m_2 are coprime, then it has a solution mod $m_1 m_2$. Give an example where this fails without the coprimeness condition.

3. Verify that the statement of quadratic reciprocity is true for the following pairs of primes (p,q): $(3,5)$, $(3,7)$, $(3,13)$, $(5,13)$, $(7,11)$, and $(13,17)$.

4. Evaluate the following Legendre symbols: $\left(\frac{30}{101}\right)$, $\left(\frac{99}{101}\right)$, $\left(\frac{506}{967}\right)$.

5. Show that $\left(\frac{a}{p}\right)$ can always be computed just from the four basic properties of Legendre symbols.

6. Determine which numbers in the range from 40 to 50 are squares mod 132.

7. (a) Using quadratic reciprocity determine which primes are represented by some form of discriminant 17.

(b) Show that all forms of discriminant 17 are equivalent to the form $x^2 + xy - 4y^2$.

(c) Draw enough of the topograph of $x^2 + xy - 4y^2$ to show all values between -70 and 70, and verify that the primes that occur are precisely the ones predicted by your answer in part (a).

8. Determine which primes are represented by at least one form of the following discriminants: (a) 21 (b) -19 (c) -20 (d) -24.

9. Show that every prime is represented by at least one of the forms x^2+y^2, x^2+2y^2, and $x^2 - 2y^2$.

10. Consider forms $Q = ax^2+bxy+cy^2$ of discriminant Δ. Show that the following three conditions are equivalent:

(1) The coefficients a, b, and c of Q are all odd.

(2) Q represents only odd numbers.

(3) $\Delta \equiv 5 \bmod 8$.

11. For which fundamental discriminants Δ is there a form of discriminant Δ representing $|\Delta|$? What about nonfundamental discriminants?

12. In terms of their prime factorizations, which numbers are sums of two nonzero squares? Which squares are sums of two nonzero squares?

13. Show that if the form $x^2 + ny^2$ represents 2^k with n odd and $k > 0$ then $n \equiv 7 \bmod 8$ except when $(n, k) = (1, 1)$ and $(3, 2)$.

14. Show that for each prime p dividing the conductor for discriminant Δ there is at least one primitive form of discriminant Δ that represents a power of p. *Hint*: Use induction on the highest power of p dividing the conductor, along with Theorem 6.11 and Propositions 6.13 and 6.14.

15. This exercise involves using quadratic reciprocity to apply Legendre's Theorem (Theorem 2.6) on rational points on quadratic curves.

(a) Determine the values of n for which the curve $2x^2 + ny^2 = 1$ contains rational points, assuming n is odd and squarefree. For each of the first three positive values of n for which the curve contains rational points find two of these rational points that lie in the first quadrant.

(b) For the same equation show that the case that n is even and squarefree reduces to the case n is odd and squarefree.

(c) Determine the values of n for which the curve $3x^2 + ny^2 = 1$ contains rational points, assuming n is odd, squarefree, and coprime to 3.

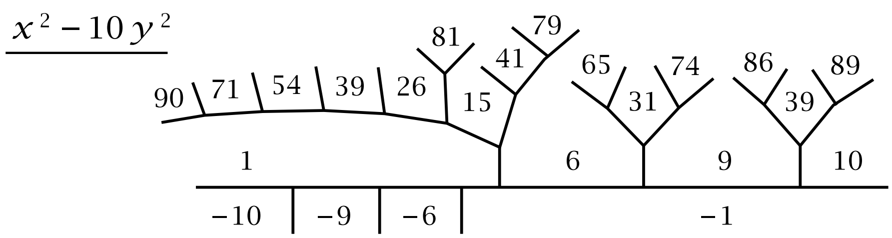
6.3 Genus and Characters

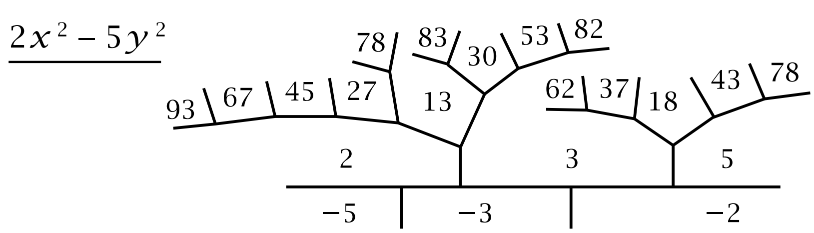
In the previous section we obtained a reasonably complete answer to the question of which numbers are represented by at least one form of a given discriminant. Legendre symbols determine which primes are represented, and in a fairly simple way this determines which nonprimes are represented. For discriminants of class number 1 this gives a complete answer to the question of which numbers are represented by a given form.

The main goal of the present section is to see how Legendre symbols, along with a few extensions of them for the special prime 2, can give additional information when the class number is not 1. In particular, in favorable cases we will be able to determine fully which forms represent which primes. Underlying this method is the following basic result:

Proposition 6.19. *Let Q be a form of discriminant Δ and let p be an odd prime dividing Δ. Then the Legendre symbol $\left(\frac{n}{p}\right)$ has the same value for all numbers n in the topograph of Q that are not divisible by p.*

Before proving this let us see how it applies in the case $\Delta = 40$ with $p = 5$. The class number here is 2 corresponding to the forms $x^2 - 10y^2$ and $2x^2 - 5y^2$.





According to the proposition, for each of the two forms the value of $\left(\frac{n}{5}\right)$ must be the same for all numbers n in the topograph not divisible by 5. To determine the value of $\left(\frac{n}{5}\right)$ for each form it therefore suffices to compute it for a single number n. The simplest thing is just to compute it for $(x,y) = (1,0)$ or $(0,1)$. Choosing $(1,0)$, for $x^2 - 10y^2$ we have $\left(\frac{1}{5}\right) = +1$ and for $2x^2 - 5y^2$ we have $\left(\frac{2}{5}\right) = -1$. The proposition then says that all numbers n in the topograph of $x^2 - 10y^2$ not divisible by 5 have $\left(\frac{n}{5}\right) = +1$, hence $n \equiv \pm 1 \bmod 5$, while for $2x^2 - 5y^2$ we have $\left(\frac{n}{5}\right) = -1$, hence $n \equiv \pm 2 \bmod 5$. Thus the last digits of the numbers in the topograph of $x^2 - 10y^2$ must be 0, 1, 4, 5, 6, or 9 and for $2x^2 - 5y^2$ the last digits must be 0, 2, 3, 5, 7, or 8.

Note that the congruences $n \equiv \pm 1$ and $n \equiv \pm 2 \bmod 5$ are consistent with the fact that for both forms the negative values are just the negatives of the positive values. (The proposition holds for negative as well as positive numbers in topographs.)

We know that $\left(\frac{40}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{p}{5}\right)$ must equal $+1$ for primes $p \neq 2, 5$ represented by either form, so for $x^2 - 10y^2$ this product must be $(+1)(+1)$ while for $2x^2 - 5y^2$ it must be $(-1)(-1)$.

	1	3	7	9	11	13	17	19	21	23	27	29	31	33	37	39
$\left(\frac{2}{p}\right)$	+1	−1	+1	+1	−1	−1	+1	−1	−1	+1	−1	−1	+1	+1	−1	+1
$\left(\frac{p}{5}\right)$	+1	−1	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1
	Q_1	Q_2		Q_1		Q_2					Q_2		Q_1		Q_2	Q_1

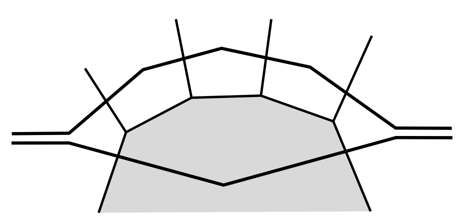
From the table we can see exactly which primes each of these two forms represents, namely $x^2 - 10y^2$ represents primes $p \equiv 1, 9, 31, 39 \bmod 40$ while $2x^2 - 5y^2$ represents primes $p \equiv 3, 13, 27, 37 \bmod 40$.

Proof of Proposition 6.19: For an edge in the topograph labeled h with adjacent regions labeled n and k we have $\Delta = h^2 - 4nk$. If p is a prime dividing Δ this implies that $4nk \equiv h^2 \bmod p$. Thus if neither n nor k is divisible by p and p is odd then the Legendre symbol $\left(\frac{4nk}{p}\right)$ is defined and $\left(\frac{4nk}{p}\right) = \left(\frac{h^2}{p}\right) = +1$. Since $\left(\frac{4nk}{p}\right) = \left(\frac{4}{p}\right)\left(\frac{n}{p}\right)\left(\frac{k}{p}\right)$ and $\left(\frac{4}{p}\right) = +1$ this implies $\left(\frac{n}{p}\right) = \left(\frac{k}{p}\right)$. In other words, the symbol $\left(\frac{n}{p}\right)$ takes the same value on any two adjacent regions of the topograph of Q labeled by numbers not divisible by p. To finish the proof we will use the following fact:

Lemma 6.20. *Given a form Q and a prime p dividing the discriminant of Q, then any two regions in the topograph of Q where the value of Q is not divisible by p can be connected by a path passing only through such regions.*

Assuming this, Proposition 6.19 easily follows since we have seen that the value of $\left(\frac{n}{p}\right)$ is the same for any two adjacent regions with label not divisible by p. □

Proof of the Lemma: Let us call regions in the topograph of Q whose label is not divisible by p *good* regions, and the other regions *bad* regions. We can assume that at least one region is good, otherwise there is nothing to prove. What we will show is that no two bad regions can be adjacent. Thus a path in the topograph from one good region to another cannot pass through two consecutive bad regions, and if it does pass through a bad region then a detour around this region allows this bad region to be avoided, creating a new path passing through one fewer bad region as in the figure at the right. By repeating this detouring process as often as necessary we eventually obtain a path avoiding bad regions entirely, still starting and ending at the same two given good regions.



To see that no two adjacent regions are bad, suppose this is false, so there are two adjacent regions whose Q values n and k are both divisible by p. If the edge separating these two regions is labeled h then we have an equation $\Delta = h^2 - 4nk$, and since we assume p divides Δ this implies that p divides h as well as n and k. Thus the form $nx^2 + hxy + ky^2$, which is equivalent to Q, is equal to p times another form. This implies that all regions in the topograph of Q are bad. This contradicts an earlier assumption so we conclude that there are no adjacent bad regions. $\square$

A useful observation is that the value of $\left(\frac{n}{p}\right)$ for numbers n in the topograph of a form $ax^2 + bxy + cy^2$ with discriminant divisible by p can always be determined just by looking at the coefficients a and c. This is because a and c appear in adjacent regions of the topograph, so if both these coefficients were divisible by p, this would imply that b was also divisible by p since p divides $b^2 - 4ac$, so the whole form would be divisible by p. Excluding this uninteresting possibility, we see that at least one of a and c is not divisible by p and we can use this to compute $\left(\frac{n}{p}\right)$.

Let us look at another example, the discriminant $\Delta = -84 = -2^2 \cdot 3 \cdot 7$ with three different prime factors. For this discriminant there are four equivalence classes of forms: $Q_1 = x^2 + 21y^2$, $Q_2 = 3x^2 + 7y^2$, $Q_3 = 2x^2 + 2xy + 11y^2$, and $Q_4 = 5x^2 + 4xy + 5y^2$. The topographs of these forms were shown in Section 6.1. To see which odd primes are represented in discriminant -84 we compute:

$$\left(\frac{-84}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right)\left(\frac{4}{p}\right)\left(\frac{7}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right)\left(\frac{7}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{3}\right)\left(\frac{p}{7}\right)$$

As in the example of $\Delta = 40$ we can make a table of the values of these Legendre symbols for the 24 numbers mod 84 that are not divisible by the prime divisors $2, 3, 7$ of 84. Using the fact that the squares mod 3 are $(\pm 1)^2 = 1$ and the squares mod 7 are $(\pm 1)^2 = 1$, $(\pm 2)^2 = 4$, and $(\pm 3)^2 \equiv 2$, we obtain the following table:

	1	5	11	13	17	19	23	25	29	31	37	41
$\left(\frac{-1}{p}\right)$	+1	+1	−1	+1	+1	−1	−1	+1	+1	−1	+1	+1
$\left(\frac{p}{3}\right)$	+1	−1	−1	+1	−1	+1	−1	+1	−1	+1	+1	−1
$\left(\frac{p}{7}\right)$	+1	−1	+1	−1	−1	−1	+1	+1	+1	−1	+1	−1
	Q_1	Q_4	Q_3		Q_4	Q_2	Q_3	Q_1		Q_2	Q_1	Q_4

	43	47	53	55	59	61	65	67	71	73	79	83
$\left(\frac{-1}{p}\right)$	−1	−1	+1	−1	−1	+1	+1	−1	−1	+1	−1	−1
$\left(\frac{p}{3}\right)$	+1	−1	−1	+1	−1	+1	−1	+1	−1	+1	+1	−1
$\left(\frac{p}{7}\right)$	+1	−1	+1	−1	−1	−1	+1	+1	+1	−1	+1	−1
				Q_2					Q_3			

The twelve cases when the product $\left(\frac{-1}{p}\right)\left(\frac{p}{3}\right)\left(\frac{p}{7}\right)$ is $+1$ give the congruence classes of primes not dividing Δ that are represented by one of the four forms, and we can determine which form it is by looking at the values of $\left(\frac{p}{3}\right)$ and $\left(\frac{p}{7}\right)$ for each of the four forms. As noted earlier, these values can be computed directly from the coefficients of x^2 and y^2 that are not divisible by 3 for $\left(\frac{p}{3}\right)$ or by 7 for $\left(\frac{p}{7}\right)$. For example, for $Q_2 = 3x^2 + 7y^2$ the coefficient of y^2 tells us that $\left(\frac{p}{3}\right) = \left(\frac{7}{3}\right) = +1$ and the coefficient of x^2 tells us that $\left(\frac{p}{7}\right) = \left(\frac{3}{7}\right) = -1$. Thus the pair $\left(\frac{p}{3}\right), \left(\frac{p}{7}\right)$ is $+1, -1$ for Q_2. In a similar way we find that $\left(\frac{p}{3}\right), \left(\frac{p}{7}\right)$ is $+1, +1$ for $Q_1 = x^2 + 21y^2$, while it is $-1, +1$ for $Q_3 = 2x^2 + 2xy + 11y^2$ and $-1, -1$ for $Q_4 = 5x^2 + 4xy + 5y^2$. This allows us to determine which congruence classes of primes are represented by which form, as indicated in the table, since the product $\left(\frac{-1}{p}\right)\left(\frac{p}{3}\right)\left(\frac{p}{7}\right)$ must be $+1$.

Another case we looked at was $\Delta = -56$ where there were three inequivalent forms $Q_1 = x^2 + 14y^2$, $Q_2 = 2x^2 + 7y^2$, and $Q_3 = 3x^2 + 2xy + 5y^2$. Here we have $\left(\frac{-56}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{7}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{p}{7}\right)$. The table of values for these Legendre symbols for congruence classes of numbers mod 56 not divisible by 2 or 7 is:

	1	3	5	9	11	13	15	17	19	23	25	27
$\left(\frac{2}{p}\right)$	$+1$	-1	-1	$+1$	-1	-1	$+1$	$+1$	-1	$+1$	$+1$	-1
$\left(\frac{p}{7}\right)$	$+1$	-1	-1	$+1$	$+1$	-1	$+1$	-1	-1	$+1$	$+1$	-1
	$\binom{Q_1}{Q_2}$	Q_3	Q_3	$\binom{Q_1}{Q_2}$		Q_3	$\binom{Q_1}{Q_2}$		Q_3	$\binom{Q_1}{Q_2}$	$\binom{Q_1}{Q_2}$	Q_3

	29	31	33	37	39	41	43	45	47	51	53	55
$\left(\frac{2}{p}\right)$	-1	$+1$	$+1$	-1	$+1$	$+1$	-1	-1	$+1$	-1	-1	$+1$
$\left(\frac{p}{7}\right)$	$+1$	-1	-1	$+1$	$+1$	-1	$+1$	-1	-1	$+1$	$+1$	-1
					$\binom{Q_1}{Q_2}$			Q_3				

From the table we see that $\left(\frac{2}{p}\right)\left(\frac{p}{7}\right)$ is $(+1)(+1)$ for $p \equiv 1, 9, 15, 23, 25, 39$ mod 56 and $(-1)(-1)$ for $p \equiv 3, 5, 13, 19, 27, 45$ mod 56. Thus the primes that are represented in discriminant -56 are the primes in these twelve congruence classes, along with 2 and 7, the prime divisors of 56. Moreover, since $\left(\frac{p}{7}\right)$ has the value $+1$ for numbers in the topographs of Q_1 and Q_2 not divisible by 7, and the value -1 for numbers in the topograph of Q_3 not divisible by 7, we can deduce that primes $p \equiv 1, 9, 15, 23, 25, 39$ mod 56 are represented by Q_1 or Q_2 while primes $p \equiv 3, 5, 13, 19, 27, 45$ mod 56 are represented by Q_3. However the values of the Legendre symbols in the table do not allow us to distinguish between Q_1 and Q_2.

Each row in one of the tables above can be regarded as a function assigning a number ± 1 to each congruence class of numbers n coprime to the discriminant Δ. Such a function is called a ***character*** and the table is called a ***character table***. There is one column in the table for each congruence class of numbers coprime to Δ so the number of columns is $\varphi(|\Delta|)$ where φ is the Euler phi function from Section 2.3. For each odd prime p dividing Δ there is a character given by the Legendre symbol $\left(\frac{n}{p}\right)$. There is sometimes also a character associated to the prime 2 in a somewhat less transparent way. In the example $\Delta = -84$ this is the character defined by the first row of the table, which assigns the values $+1$ to numbers $n = 4k + 1$ and -1 to numbers $n = 4k + 3$. We will denote this character by χ_4 to indicate that its values $\chi_4(n) = \pm 1$ depend only on the value of n mod 4. Thus $\chi_4(p) = \left(\frac{-1}{p}\right)$ when p is an odd prime, but $\chi_4(n)$ is defined for all odd numbers n, not just primes. One can check that an explicit formula for χ_4 is $\chi_4(n) = (-1)^{(n-1)/2}$ although we will not be needing this formula.

In the example with $\Delta = -56$ the character corresponding to the prime 2 is given by the row labeled $\left(\frac{2}{p}\right)$. This character associates the value $+1$ to an odd number $n \equiv \pm 1$ mod 8 and the value -1 when $n \equiv \pm 3$ mod 8. We will denote it by χ_8 since its values $\chi_8(n) = \pm 1$ depend only on n mod 8. We have $\chi_8(p) = \left(\frac{2}{p}\right)$ for all odd primes p, but $\chi_8(n)$ is defined for all odd numbers n. There is again an explicit formula $\chi_8(n) = (-1)^{(n^2-1)/8}$ that we will not use.

By analogy we can also introduce the notation χ_p for the earlier character defined by $\chi_p(n) = \left(\frac{n}{p}\right)$ for p an odd prime and n not divisible by p.

As another example illustrating the use of characters let us determine which powers of 2 are represented by the two forms $x^2 + 15y^2$ and $3x^2 + 5y^2$ of discriminant -60. This is not a fundamental discriminant since it is 4 times the fundamental discriminant -15, so the conductor is 2 which is why the question of determining the forms representing powers of 2 is more subtle, as we saw in the previous section. In both the discriminants -15 and -60 we have the characters χ_3 and χ_5 and we can use either one of these for this application so we will use χ_3.

First consider discriminant -15 where the class number is 2 corresponding to the two forms $x^2 + xy + 4y^2$ and $2x^2 + xy + 2y^2$. The second form represents 2 which does not divide the discriminant -15 so all powers of 2 are represented by one or the other of these two forms. To determine which form it is for each power we use the character χ_3. This has the value $+1$ on numbers not divisible by 3 in the topograph of $x^2 + xy + 4y^2$ since 1 is one of these numbers and $\chi_3(1) = +1$. Similarly χ_3 has the value -1 for the other form $2x^2 + xy + 2y^2$ since 2 appears in the topograph of this form and $\chi_3(2) = -1$. We have $\chi_3(2^k) = (-1)^k$ since $\chi_3(2^k) = \left(\frac{2^k}{3}\right) = \left(\frac{2}{3}\right)^k$. Hence $x^2 + xy + 4y^2$ represents only the even powers of 2 and $2x^2 + xy + 2y^2$ represents only the odd powers.

For discriminant -60 the class number is also 2, corresponding to the forms $x^2 + 15y^2$ and $3x^2 + 5y^2$. Obviously neither of these forms represents 2 or 4. However by Proposition 6.13 each power 2^k with $k \geq 3$ is represented by at least one of the two forms since all powers 2^k with $k \geq 1$ are represented by one of the forms of discriminant -15. The value of χ_3 for $x^2 + 15y^2$ is $+1$ since this form represents 1 and $\chi_3(1) = +1$, and the value of χ_3 for $3x^2 + 5y^2$ is -1 since this form represents 5 and $\chi_3(5) = -1$. From this it follows as before that $x^2 + 15y^2$ represents just the even powers of 2 starting with 2^4 and $3x^2 + 5y^2$ represents just the odd powers starting with 2^3. This is the answer that was given in the large table in the preceding section.

Characters for the Prime 2

Let us consider now how characters can be associated to the prime 2 in general. Since characters arise from primes that divide the discriminant, this means we are interested in even discriminants, and the characters we are looking for should assign a value ± 1 to each number not divisible by 2, that is, to each odd number. We would like the analogue of Proposition 6.19 to hold, so characters for the prime 2 should take the same value on all odd numbers in the topograph of a form of the given discriminant. By Lemma 6.20 this just means that the characters should have the same value for odd numbers in adjacent regions of the topographs.

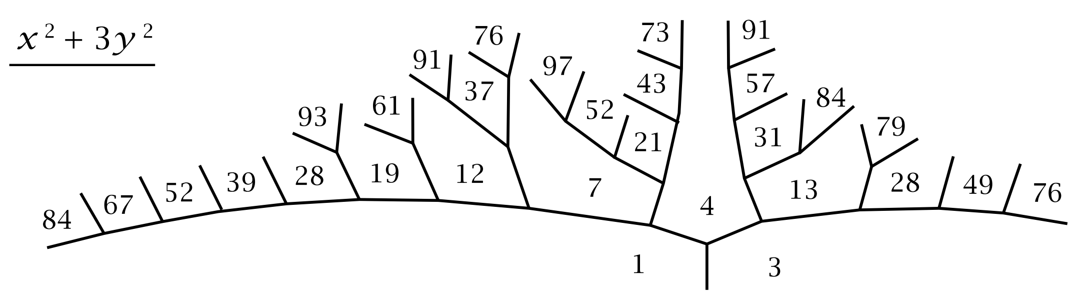
Even discriminants are multiples of 4 so can be written as $\Delta = 4\delta$. For adjacent regions in a topograph with labels n and k we have $\Delta = h^2 - 4nk$ where h is the label on the edge between the two regions. Since Δ is even, so is h and we can write $h = 2l$. The discriminant equation then becomes $4\delta = 4l^2 - 4nk$ or just $\delta = l^2 - nk$.

There will be six different cases. The first two are when δ is odd, which means that Δ is divisible by 4 but not 8. In these two cases we consider congruences mod 4, the highest power of 2 dividing Δ. Since δ is odd and both n and k are odd, the equation $\delta = l^2 - nk$ implies that l must be even, so $l^2 \equiv 0 \bmod 4$ and we have $nk \equiv -\delta \bmod 4$. Multiplying both sides of this congruence by k, we get $n \equiv -\delta k \bmod 4$ since $k^2 \equiv 1 \bmod 4$, k being odd. Multiplying the congruence $n \equiv -\delta k$ by k again gives the previous congruence $nk \equiv -\delta$ so the two congruences are equivalent.

Case 1: $\delta = 4m-1$. The congruence condition $n \equiv -\delta k \bmod 4$ is then $n \equiv k \bmod 4$. Thus Lemma 6.20 implies that the character χ_4 assigning $+1$ to integers $4s+1$ and -1 to integers $4s-1$ has the same value for all odd numbers in the topograph of a form of discriminant $\Delta = 4(4m-1)$. We might try reversing the values of χ_4, assigning the value $+1$ to integers $4s-1$ and -1 to integers $4s+1$, but this just gives the function $-\chi_4$ which does not really give any new information that χ_4 does not give. In practice χ_4 turns out to be more convenient to use than $-\chi_4$ would be.

An example for the case $\delta = 4m-1$ is the discriminant $\Delta = -84$ considered earlier, where the first row of the character table gave the values for χ_4.

Case 2: $\delta = 4m+1$. The difference from the previous case is that the congruence condition is now $n \equiv -k \bmod 4$. This means the mod 4 value of odd numbers in the topograph is not constant, and so we do not get a character for the prime 2. As an example, consider the form $x^2 + 3y^2$ with $\Delta = -12$ and $\delta = -3$.



Here there are odd numbers in the topograph congruent to both 1 and 3 mod 4. The situation is not improved by considering odd numbers mod 8 instead of mod 4 since the topograph contains numbers congruent to each of $1, 3, 5, 7$ mod 8. Trying congruences modulo higher powers of 2 does not help either.

The absence of a character for the prime 2 when $\delta = 4m + 1$ could perhaps have been predicted from the calculation of $\left(\frac{\Delta}{p}\right)$. Since δ is odd we have $\Delta = 4\delta = 4p_1 \cdots p_r$ for odd primes $p_1, \cdots, p_r$ and so $\left(\frac{\Delta}{p}\right) = \left(\frac{p_1}{p}\right) \cdots \left(\frac{p_r}{p}\right)$. This equals $\left(\frac{p}{p_1}\right) \cdots \left(\frac{p}{p_r}\right)$ since the number of primes p_i congruent to 3 mod 4 is even when $\delta = 4m + 1$. Thus the value of $\left(\frac{\Delta}{p}\right)$ depends only on the characters associated to the odd prime factors of Δ.

There remain the cases that δ is even. The next two cases are when Δ is divisible by 8 but not by 16. After that is the case that Δ is divisible by 16 but not by 32, and finally the case that Δ is divisible by 32. In all these cases we will consider congruences mod 8, so the equation $\delta = l^2 - nk$ becomes $\delta \equiv l^2 - nk$ mod 8. Since δ is now even while n and k are still odd, this congruence implies l is odd, and so $l^2 \equiv 1$ mod 8 and the congruence can be written as $nk \equiv 1 - \delta$ mod 8. Since $k^2 \equiv 1$ mod 8 when k is odd, we can multiply both sides of the congruence $nk \equiv 1 - \delta$ by k to obtain the equivalent congruence $n \equiv (1 - \delta)k$ mod 8.

Case 3: $\delta \equiv 2$ mod 8. The congruence is then $n \equiv -k$ mod 8. It follows that in the topograph of a form of discriminant $\Delta = 4(8m + 2)$ either the odd numbers must all be congruent to ± 1 mod 8 or they must all be congruent to ± 3 mod 8. Thus the character χ_8 which takes the value $+1$ on numbers $8s \pm 1$ and -1 on numbers $8s \pm 3$ has a constant value, either $+1$ or -1, for all odd numbers in the topograph.

An example for this case is $\Delta = 40$. Here the two rows of the character table computed earlier in this section gave the values for χ_8 and χ_5.

Case 4: $\delta \equiv 6$ mod 8. Now the congruence $n \equiv (1 - \delta)k$ mod 8 becomes $n \equiv -5k$, or equivalently $n \equiv 3k$ mod 8. This implies that all odd numbers in the topograph of a form of discriminant $\Delta = 4(8m + 6)$ must be congruent to 1 or 3 mod 8, or they must all be congruent to 5 or 7 mod 8. The character associated to the prime 2 in this case has the value $+1$ on numbers $8s + 1$ and $8s + 3$, and the value -1 on numbers $8s + 5$ and $8s + 7$. We have not encountered this character previously, so let us give it the new name χ_8'. However, it is not entirely new since it is actually just the product $\chi_4\chi_8$ as one can easily check by evaluating this product on $1, 3, 5$, and 7.

A simple example is $\Delta = -8$ with class number 1. Here we have $\left(\frac{\Delta}{p}\right) = \left(\frac{-8}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)$ which equals $+1$ for $p \equiv 1, 3$ mod 8 and -1 for $p \equiv 5, 7$ mod 8 so this is just the character χ_8'.

Another example is $\Delta = 24$ where there are the two forms $Q_1 = x^2 - 6y^2$ and $Q_2 = 6x^2 - y^2$. We have $\left(\frac{\Delta}{p}\right) = \left(\frac{24}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{3}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{-1}{p}\right)\left(\frac{p}{3}\right)$. The character table has the following form:

	1	5	7	11	13	17	19	23
χ_8'	$+1$	-1	-1	$+1$	-1	$+1$	$+1$	-1
χ_3	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1

Thus Q_1 represents primes $p \equiv 1, 19 \bmod 24$ and Q_2 represents primes $p \equiv 5, 23 \bmod 24$.

Case 5: $\delta \equiv 4 \bmod 8$. Now we have the congruence $n \equiv -3k \bmod 8$. Thus in the topograph of a form of discriminant $\Delta = 4(8m+4)$ all odd numbers must be congruent to 1 or 5 mod 8, or they must all be congruent to 3 or 7 mod 8. More simply, one can say that all odd numbers in the topograph must be congruent to 1 mod 4 or they must all be congruent to 3 mod 4. Thus we obtain the character χ_4 again.

An example is $\Delta = -48$ where we have the two forms $Q_1 = x^2 + 12y^2$ and $Q_2 = 3x^2 + 4y^2$ as well as a pair of nonprimitive forms $Q_3 = 2x^2 + 6y^2$ and $Q_4 = 4x^2 + 4xy + 4y^2$. We have $\left(\frac{\Delta}{p}\right) = \left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right)$. This is the character χ_3. We also have the character χ_4 that we just described. Here is the character table:

	1	5	7	11	13	17	19	23	25	29	31	35	37	41	43	47
χ_4	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1
χ_3	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1
	Q_1		Q_2		Q_1		Q_2		Q_1		Q_2		Q_1		Q_2	

The columns repeat every four columns since $\left(\frac{-1}{p}\right)$ and $\left(\frac{p}{3}\right)$ are determined by the value of p mod 12. In contrast with earlier examples, the representability of a prime $p > 3$ in discriminant -48 is determined by one character, χ_3, and the other character χ_4 serves only to decide which of the forms Q_1 and Q_2 achieves the representation. The character χ_4 says nothing about the nonprimitive forms Q_3 and Q_4 whose values are all even. On the other hand, from χ_3 we can deduce that all values of Q_3 not divisible by 3 must be congruent to 2 mod 3 while for Q_4 they must be congruent to 1 mod 3. This could also have been deduced from applying χ_3 to the associated primitive forms $x^2 + 3y^2$ and $x^2 + xy + y^2$.

Case 6: $\delta \equiv 0 \bmod 8$, so Δ is a multiple of 32. In this case the congruence $n \equiv (1-\delta)k \bmod 8$ becomes simply $n \equiv k \bmod 8$. Thus all odd numbers in the topograph of a form of discriminant $\Delta = 32m$ must lie in the same congruence class mod 8. The two characters χ_4 and χ_8 can now both occur independently, as shown in the following chart listing their values on the four classes $1, 3, 5, 7 \bmod 8$:

	1	3	5	7
χ_4	$+1$	-1	$+1$	-1
χ_8	$+1$	-1	-1	$+1$

As an example consider the discriminant $\Delta = -32$. Here there are two primitive forms $Q_1 = x^2 + 8y^2$ and $Q_2 = 3x^2 + 2xy + 3y^2$ along with one nonprimitive form $Q_3 = 2x^2 + 4y^2$. We have $\left(\frac{\Delta}{p}\right) = \left(\frac{-2}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)$ with the two factors being the two independent characters for the prime 2. The full character table is then just a four-fold repetition of the previous shorter table:

	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
χ_4	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1	$+1$	-1
χ_8	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$
	Q_1	Q_2			Q_1	Q_2			Q_1	Q_2			Q_1	Q_2		

This finishes the analysis of the six cases for characters associated to the prime 2. To summarize we have:

Proposition 6.21. *The characters associated to the prime* 2 *are given in the following table*:

Δ	$4(4m+1)$	$4(4m+3)$	$8(4m+1)$	$8(4m+3)$	$16(2m+1)$	$32m$
χ	—	χ_4	χ_8	$\chi_8' = \chi_4\chi_8$	χ_4	χ_4, χ_8

We have now defined a set of characters for each discriminant Δ, with one character for each odd prime dividing Δ and either zero, one, or two characters for the prime 2 when Δ is even. The character table for discriminant Δ has one row for each of these characters.

If one restricts attention to fundamental discriminants then the only relevant columns in the table in the preceding proposition are the second, third, and fourth columns on the right. Thus the characters for the prime 2 that arise in the three cases of fundamental discriminants are exactly χ_4, χ_8, and χ_8'.

A nice property satisfied by characters is that they are multiplicative, so $\chi(mn) = \chi(m)\chi(n)$ for all m and n for which χ is defined. For the characters χ_p associated to odd primes p this is just the basic property $\left(\frac{mn}{p}\right) = \left(\frac{m}{p}\right)\left(\frac{n}{p}\right)$ of Legendre symbols. For the prime 2 the characters χ_4 and χ_8 are multiplicative as well. For χ_4 this holds since $\chi_4(1\cdot 1) = +1 = \chi_4(1)\chi_4(1)$, $\chi_4(1\cdot 3) = -1 = \chi_4(1)\chi_4(3)$, and $\chi_4(3\cdot 3) = +1 = \chi_4(3)\chi_4(3)$. Similarly for χ_8 we have $\chi_8(\pm 1\cdot \pm 1) = +1 = \chi_8(\pm 1)\chi_8(\pm 1)$, $\chi_8(\pm 1\cdot \pm 3) = -1 = \chi_8(\pm 1)\chi_8(\pm 3)$, and $\chi_8(\pm 3\cdot \pm 3) = +1 = \chi_8(\pm 3)\chi_8(\pm 3)$. The multiplicativity of χ_8' follows since $\chi_8' = \chi_4\chi_8$.

In fact χ_4, χ_8, and χ_8' are the only multiplicative functions from the odd integers mod 8 to $\{\pm 1\}$, apart from the trivial function assigning $+1$ to all four of $1, 3, 5, 7$. To see this, note first that each of $3, 5, 7$ has square equal to 1 mod 8 and the product of any two of $3, 5, 7$ is the third, mod 8. This means that a multiplicative function χ from odd integers mod 8 to $\{\pm 1\}$ is completely determined by the two values $\chi(3)$ and $\chi(5)$ since $\chi(1) = \chi(3)\chi(3)$ and $\chi(7) = \chi(3)\chi(5)$. For χ_4 the values on 3 and 5 are $-1, +1$, for χ_8 they are $-1, -1$, and for $\chi_8' = \chi_4\chi_8$ they are $+1, -1$. The only other possibility is $+1, +1$ but this leads to the trivial character.

As we know, an odd prime p is represented in discriminant Δ exactly when $\left(\frac{\Delta}{p}\right) = +1$. This criterion can also be expressed in terms of characters via the following restatement of Proposition 6.9 in different notation:

Proposition 6.22. $\left(\frac{\Delta}{p}\right) = X_\Delta(p)$ *for* X_Δ *the product of characters given in the table below, where* $\Delta = \varepsilon 2^s p_1 \cdots p_k$ *for* $\varepsilon = \pm 1$ *with each* p_i *an odd prime.* □

Δ	$\left(\frac{\Delta}{p}\right)$	X_Δ
$2^{2l}(4m+1)$	$\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$	$\chi_{p_1}\cdots\chi_{p_k}$
$2^{2l}(4m+3)$	$\left(\frac{-1}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$	$\chi_4\chi_{p_1}\cdots\chi_{p_k}$
$2^{2l+1}(4m+1)$	$\left(\frac{2}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$	$\chi_8\chi_{p_1}\cdots\chi_{p_k}$
$2^{2l+1}(4m+3)$	$\left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{p}{p_1}\right)\cdots\left(\frac{p}{p_k}\right)$	$\chi_8'\chi_{p_1}\cdots\chi_{p_k}$

The value $X_\Delta(n) = \pm 1$ is defined whenever n is coprime to Δ. If n is represented in discriminant Δ then $X_\Delta(n) = +1$ since each prime factor p of n is then represented, so $X_\Delta(p) = +1$, and $X_\Delta(n)$ is the product of these terms $X_\Delta(p)$ since X_Δ is multiplicative, being a product of multiplicative functions. If n is not a prime it can happen that $X_\Delta(n) = +1$ even when n is not represented in discriminant Δ. For example for $\Delta = -4$ we have $X_\Delta(21) = \chi_4(21) = \chi_4(3)\chi_4(7) = (-1)(-1) = +1$ but 21 is not represented by the form $x^2 + y^2$, the only form in this discriminant up to equivalence.

Next let us verify that some of the special features of the character tables in the earlier examples hold in general.

Proposition 6.23. (a) *The columns of a character table contain all possible combinations of* $+1$ *and* -1*, and each such combination occurs in the same number of columns.*
(b) *If the discriminant* Δ *is not a square then half of the columns have* $X_\Delta(n) = +1$ *and half have* $X_\Delta(n) = -1$ *for numbers* n *in the congruence class corresponding to the column.*

For example, if Δ is a fundamental discriminant then X_Δ is just the product of all the characters in the character table, so the combinations of ± 1's that give $X_\Delta = +1$ in these cases are the combinations with an even number of -1's. This need not be true for nonfundamental discriminants as the earlier example $\Delta = -48$ shows.

From statement (b) in the proposition we immediately deduce:

Corollary 6.24. *For hyperbolic and elliptic forms, the primes not dividing the discriminant* Δ *that are represented in discriminant* Δ *are the primes in exactly half of the congruence classes mod* Δ *of numbers coprime to* Δ*.*

For the proof of Proposition 6.23 we will need the following fact:

Lemma 6.25. *For a power p^r of an odd prime p exactly half of the $p^r - p^{r-1}$ congruence classes mod p^r of numbers a not divisible by p satisfy $\left(\frac{a}{p}\right) = +1$.*

Proof: First we do the case $r = 1$. The $p-1$ nonzero congruence classes mod p are $\pm 1, \pm 2, \cdots, \pm \frac{1}{2}(p-1)$. The two numbers $+a$ and $-a$ in each pair $\pm a$ have the same square, so there are at most $\frac{1}{2}(p-1)$ different nonzero squares mod p. In fact there are exactly this many since if $a^2 \equiv b^2$ mod p then p divides $a^2 - b^2 = (a-b)(a+b)$, so since p is prime it must divide either $a-b$ or $a+b$ which means that either $a \equiv b$ or $a \equiv -b$ mod p. Thus exactly half of the $p-1$ nonzero congruence classes mod p are squares, so the lemma is proved when $r = 1$.

Now suppose $r > 1$. The value of $\left(\frac{a}{p}\right)$ depends only on the congruence class of a mod p so there are the same number of numbers a with $\left(\frac{a}{p}\right) = +1$ in each of the intervals $[0, p]$, $[p, 2p]$, $[2p, 3p]$, etc. There are p^{r-1} of these intervals in $[0, p^r]$. Thus half of the $p^{r-1}(p-1) = p^r - p^{r-1}$ congruence classes mod p^r of numbers a not divisible by p have $\left(\frac{a}{p}\right) = +1$ and half have $\left(\frac{a}{p}\right) = -1$. □

Proof of Proposition 6.23: Let us write $\Delta = \varepsilon\, 2^s p_1^{r_1} \cdots p_k^{r_k}$ where $\varepsilon = \pm 1$, $s \ge 0$, and the p_i's are the distinct odd prime divisors of Δ. Thus the characters for this discriminant are $\chi_{p_1}, \cdots, \chi_{p_k}$ and either zero, one, or two characters associated to the prime 2 when $s > 0$.

To prove statement (a) choose numbers a_i realizing any combination of preassigned values $\chi_{p_i}(a_i) = \pm 1$. When $s > 0$ we also choose a number 1, 3, 5, or 7 to realize any preassigned pair of values for χ_4 and χ_8, hence for any preassigned values for the characters associated to the prime 2. By the Chinese Remainder Theorem there is a number a congruent to each a_i mod $p_i^{r_i}$ and to the chosen number $1, 3, 5, 7$ mod 8. The number a is coprime to Δ since it is nonzero mod p_i for each i and is odd when $s > 0$. Thus the column in the character table corresponding to a realizes the chosen values for all the characters.

To prove the second half of statement (a) we will count the number of columns in the character table realizing a given combination of values ± 1 and see that this number does not depend on which combination is chosen. By the preceding lemma the number of choices for a_i mod $p_i^{r_i}$ in the previous paragraph is $\frac{1}{2} p_i^{r_i - 1}(p_i - 1)$, so the Chinese Remainder Theorem implies that when $s = 0$ the number of congruence classes mod Δ realizing a given combination of values ± 1 is the product of these numbers $\frac{1}{2} p_i^{r_i-1}(p_i - 1)$. When $s > 0$ but there is no character for the prime 2, the product of the numbers $\frac{1}{2} p_i^{r_i-1}(p_i - 1)$ is multiplied by 2^{s-1} since this is the number of odd congruence classes mod 2^s. If there is one character for the prime 2 the number 2^{s-1} is cut in half, and if there are two characters for the prime 2 it is cut in half again. Thus in all cases the number of columns realizing a given combination of ± 1's is independent of the combination.

For (b), consider the definition of X_Δ which has four different cases depending on the prime factorization of Δ. If Δ is a square then the applicable formula is the first of the four formulas since an odd square is $1 \bmod 4$, and in fact the formula degenerates to just the constant $+1$ since its terms all cancel out, as each prime factor of Δ occurs to an even power. When Δ is not a square, the terms in the first of the four formulas do not all cancel out, and in the other three formulas there is also at least one term remaining after cancellations, either χ_4, χ_8, or χ_8'.

In view of property (a), to prove (b) it will suffice to show that when Δ is not a square, the set of combinations of values ± 1 in columns of the character table that give $X_\Delta = +1$ has the same number of elements as the set of combinations that give $X_\Delta = -1$. But this is obviously true since we can interchange these two sets by choosing one term in the formula for X_Δ that remains after cancellation and switching the sign of the value ± 1 for this term, keeping the values for the other characters unchanged. □

Genus

Recall the concept of genus that was introduced informally in Section 6.1. The idea was that if two forms of the same discriminant cannot be distinguished by looking only at their values modulo the discriminant then they should be regarded as having the same genus. Here it is best to restrict attention just to primitive forms. We can now give this notion a more precise definition by saying that two primitive forms of discriminant Δ have the same ***genus*** if each character for discriminant Δ takes the same value on the two forms, where the value of a character on a form means its value on all numbers in the topograph not divisible by the prime associated to the character.

In fact there is always a single number in the topograph that can be used to evaluate all the characters, according to the following general result:

Proposition 6.26. *Given a positive integer n and a primitive form Q that represents at least one positive number, then Q represents a positive number coprime to n.*

For the application to evaluating characters we choose $n = |\Delta|$ for Δ the discriminant of Q, which we assume is nonzero.

Proof: Let $Q = ax^2 + bxy + cy^2$. We can replace Q by any equivalent form so we can arrange that $a > 0$ and $c > 0$ by choosing two adjacent regions in the topograph of Q with positive labels a and c. We can also assume $b \geq 0$ since changing the sign of b produces an equivalent form.

The case $n = 1$ is trivial since every positive number is coprime to 1, so we may assume $n > 1$. Suppose first that n is a prime p. One of the following three cases will apply:

(1) If p does not divide a let (x, y) be a primitive pair with p dividing y but not x. Then p will not divide $ax^2 + bxy + cy^2$. For example we could take $(x, y) = (1, p)$.

(2) If p divides a but not c let (x, y) be a primitive pair with p dividing x but not y. Then p will not divide $ax^2 + bxy + cy^2$. For example we could take $(x, y) = (p, 1)$.

(3) If p divides both a and c then it will not divide b since Q is primitive. In this case let (x, y) be a primitive pair with neither x nor y divisible by p. Then p will not divide $ax^2 + bxy + cy^2$. For example we could take $(x, y) = (1, 1)$.

This finishes the proof when n is prime. For a general n let $p_1, \cdots, p_k$ be its distinct prime divisors. For each p_i let (x_i, y_i) be $(1, p_i)$, $(p_i, 1)$, or $(1, 1)$ according to which of the three cases above applies to p_i. Now let $x = x_1 \cdots x_k$ and $y = y_1 \cdots y_k$. Then x and y are coprime since no p_i is a factor of both x and y. If the number $ax^2 + bxy + cy^2$ is not coprime to n it will be divisible by some p_i. If case (1) applies to p_i then p_i divides y but not x so p_i does not divide $ax^2 + bxy + cy^2$. Likewise if cases (2) or (3) apply to p_i then p_i does not divide $ax^2 + bxy + cy^2$. Thus no p_i can divide $ax^2 + bxy + cy^2$. Finally, $ax^2 + bxy + cy^2$ is positive since x and y are positive as are the coefficients except possibly b which is either positive or zero. □

The number of genera in discriminant Δ is at most 2^κ where κ is the number of characters in discriminant Δ. In all the character tables we have looked at, only half of the 2^κ possible combinations of ± 1's were actually realized by forms, and in fact this is true generally:

Theorem 6.27. *If Δ is not a square then the number of genera of primitive forms of discriminant Δ is $2^{\kappa-1}$ where κ is the number of characters in discriminant Δ.*

This turns out to be fairly hard to prove. The original proof by Gauss required a somewhat lengthy digression into the theory of quadratic forms in three variables. An exposition of this proof can be found in the book by Flath listed in the Bibliography. We will give a different proof that deduces the result rather quickly from things we have already done, together with Dirichlet's Theorem about primes in arithmetic progressions discussed at the end of Section 6.1, which we will not prove. We will not need the full strength of Dirichlet's Theorem, and in fact all we will actually need is that each congruence class of numbers $x \equiv b \bmod a$ contains at least one prime greater than 2 if a and b are coprime. One might think this would be easier to prove than that there are infinitely many primes in the congruence class, but this seems not to be the case.

Proof of Theorem 6.27 using Dirichlet's Theorem: We have seen that for each primitive form Q of discriminant Δ there is a number n coprime to Δ that is represented by Q. Then $X_\Delta(n)$ is defined, and we saw when we defined X_Δ that $X_\Delta(n) = +1$ when n is represented by a form of discriminant Δ. In the proof of Proposition 6.23 we showed that exactly half of the 2^κ possible combinations of ± 1's have $X_\Delta = +1$, so the number of genera of forms is at most $2^{\kappa-1}$.

To show that the number of genera is at least $2^{\kappa-1}$ consider a combination of ± 1's with $X_\Delta = +1$. By Proposition 6.23 this combination occurs in some column of the character table. This column corresponds to some number n coprime to Δ. By Dirichlet's Theorem there exists a prime p congruent to n mod Δ. We have $X_\Delta(p) = +1$, so since p is prime this implies that p is represented by some form of discriminant Δ. This form must be primitive, otherwise every number it represents would be divisible by some number $d > 1$ dividing Δ so it could not represent p which is coprime to Δ. Thus every combination of ± 1's with $X_\Delta = +1$ is realized by some primitive form, so the number of genera is at least $2^{\kappa-1}$. □

From this theorem we can deduce two very strong corollaries.

Corollary 6.28. *For a nonsquare discriminant the number of genera is equal to the number of equivalence classes of primitive forms that have mirror symmetry.*

This may seem a little surprising since there is no apparent connection between genera and mirror symmetry. A possible explanation might be that each genus contains exactly one equivalence class of primitive forms with mirror symmetry, but this is not always true. For example when $\Delta = -56$ we saw in Section 6.1 that there are two genera and two equivalence classes of mirror symmetric forms, but both these forms belong to the same genus. The true explanation will come in Chapter 7 when we study the class group.

Proof: For a nonsquare discriminant the number of equivalence classes of primitive forms with mirror symmetry was computed in Theorem 5.9 to be 2^{k-1} in most cases, where k is the number of distinct prime divisors of Δ. The exceptions are discriminants $\Delta = 4(4m+1)$ when 2^{k-1} is replaced by 2^{k-2}, and $\Delta = 32m$ when 2^{k-1} is replaced by 2^k. In the nonexceptional cases we have $k = \kappa$, the number of characters in discriminant Δ since there is one character for each prime dividing Δ. When $\Delta = 4(4m+1)$ there is no character for the prime 2 so $\kappa = k-1$, and when $\Delta = 32m$ there are two characters for the prime 2 so $\kappa = k+1$. The result follows. □

Corollary 6.29. *For a nonsquare discriminant, each genus of primitive forms consists of a single equivalence class of forms if and only if all the topographs of primitive forms have mirror symmetry.*

Proof: Let $E(\Delta)$ be the set of equivalence classes of primitive forms of discriminant Δ and let $G(\Delta)$ be the set of genera of primitive forms of discriminant Δ. There is a natural function $\Phi : E(\Delta) \to G(\Delta)$ assigning to each equivalence class of forms the genus of these forms. The function Φ is onto since there is at least one form in each genus, by the definition of genus. If all primitive forms of discriminant Δ have mirror symmetry then Corollary 6.28 says that the sets $E(\Delta)$ and $G(\Delta)$ have the same number of elements. Then since Φ is onto it must also be one-to-one. This means that each genus consists of a single equivalence class of forms.

Conversely, if each genus consists of a single equivalence class then Φ is one-to-one. Since Φ is also onto, this means it is a one-to-one correspondence so $E(\Delta)$ and $G(\Delta)$ have the same number of elements. By Corollary 6.28 this means that the equivalence classes of primitive forms with mirror symmetry account for all the elements of $E(\Delta)$, and the proof is complete. □

Exercises

1. For the following discriminants determine the class number and a form in each class, then use a character table to determine which primes are represented by each of the forms, at least to the extent that this can be determined by characters. Also determine the various genera.
(a) -24 (b) 24 (c) -39 (d) -96

2. Determine which primes are represented by each of the following forms:
(a) $x^2 + 8y^2$ (b) $x^2 + 9y^2$ (c) $x^2 + 25y^2$ (d) $x^2 - 12y^2$ and $12x^2 - y^2$

3. Show that each genus consists of a single equivalence class of forms for the following discriminants: (a) -168 (b) -660 (c) 105

4. Find the smallest positive discriminant for which the number of genera is 16. How does the answer change if only fundamental discriminants are allowed?

5. Show that for a positive nonsquare discriminant Δ, if the principal form represents -1 then all odd primes p dividing Δ must satisfy $p \equiv 1 \bmod 4$. *Hint*: Use χ_p.

6. Use Propositions 6.1 and 6.26 to show that in each nonzero discriminant there exists a form that represents an infinite number of primes.

6.4 Proof of Quadratic Reciprocity

First let us show that quadratic reciprocity can be expressed more concisely as a single formula:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$$

Here p and q are distinct odd primes. Since they are odd, the fractions $\frac{p-1}{2}$ and $\frac{q-1}{2}$ are integers. The only way the exponent $\frac{p-1}{2}\cdot\frac{q-1}{2}$ can be odd is for both factors to be odd, so $\frac{p-1}{2} = 2k+1$ and $\frac{q-1}{2} = 2l+1$, which is equivalent to saying $p = 4k+3$ and $q = 4l+3$. Thus the only time that the right side of the formula shown above is -1 is when p and q are both congruent to $3 \bmod 4$, and quadratic reciprocity is the assertion that the left side of the formula has exactly this property.

There will be three main steps in the proof of quadratic reciprocity. The first is to derive an explicit algebraic formula for $\left(\frac{a}{p}\right)$ due originally to Euler. The second step is to use this formula to give a somewhat more geometric interpretation of $\left(\frac{a}{p}\right)$ in terms of the number of dots in a certain triangular pattern. Then the third step is the actual proof of quadratic reciprocity using symmetry properties of the patterns of dots. This proof is due to Eisenstein, first published in 1844, simplifying an earlier proof by Gauss who was the first to give a full proof of quadratic reciprocity.

Step 1. In what follows we will always use p to denote an odd prime, and the symbol a will always denote an arbitrary nonzero integer not divisible by p. When we write a congruence such as $a \equiv b$ this will always mean congruence mod p, even if we do not explicitly say that the modulus is p.

Euler's formula is

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p$$

For example, for $p = 11$ Euler's formula says $\left(\frac{2}{11}\right) \equiv 2^5 = 32 \equiv -1 \mod 11$ and $\left(\frac{3}{11}\right) \equiv 3^5 = 243 \equiv +1 \mod 11$. These are the correct values since the squares mod 11 are $(\pm 1)^2 = 1$, $(\pm 2)^2 = 4$, $(\pm 3)^2 = 9$, $(\pm 4)^2 \equiv 5$, and $(\pm 5)^2 \equiv 3$.

Euler's formula determines the value of $\left(\frac{a}{p}\right)$ uniquely since $+1$ and -1 are not congruent mod p if $p > 2$. It is not immediately obvious that the number $a^{\frac{p-1}{2}}$ should always be congruent to either $+1$ or -1 mod p, but when we prove Euler's formula we will see that this has to be true.

As a special case, taking $a = -1$ in Euler's formula gives the calculation of $\left(\frac{-1}{p}\right)$:

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} +1 & \text{if } p = 4k+1 \\ -1 & \text{if } p = 4k+3 \end{cases}$$

Before proving Euler's formula we will need to derive a few preliminary facts about congruences modulo a prime p. First let us note that each of the numbers $a = 1, 2, \cdots, p-1$ has a multiplicative inverse mod p. This is a special case of the fact that each number coprime to a number n has a multiplicative inverse mod n as we saw in Section 2.3. (This was because the equation $ax + ny = 1$ has an integer solution (x, y) whenever a and n are coprime.) Any two choices for an inverse to a mod p are congruent mod p since if $ax \equiv 1$ and $ax' \equiv 1$ then multiplying both sides of $ax' \equiv 1$ by x gives $xax' \equiv x$, and $xa \equiv 1$ so we conclude that $x \equiv x'$.

Which numbers equal their own inverse mod p? If $a \cdot a \equiv 1$, then we can rewrite this as $a^2 - 1 \equiv 0$, or equivalently $(a+1)(a-1) \equiv 0$. This is certainly a valid congruence if $a \equiv \pm 1$, so suppose that $a \not\equiv \pm 1$. The factor $a + 1$ is then not congruent to 0 mod p so it has a multiplicative inverse mod p, and if we multiply the congruence $(a+1)(a-1) \equiv 0$ by this inverse, we get $a - 1 \equiv 0$ so $a \equiv 1$, contradicting the assumption that $a \not\equiv \pm 1$. This argument shows that the only numbers among $1, 2, \cdots, p-1$ that are congruent to their inverses mod p are 1 and $p-1$.

An application of this fact is a result known as ***Wilson's Theorem***:

$$(p-1)! \equiv -1 \mod p \quad \text{whenever } p \text{ is prime.}$$

To see why this is true, observe that in the product $(p-1)! = (1)(2)\cdots(p-1)$ each factor other than 1 and $p-1$ can be paired with its multiplicative inverse mod p and these two terms multiply together to give 1 mod p, so the whole product is congruent to just $(1)(p-1)$ mod p. Since $p-1 \equiv -1$ mod p this gives Wilson's Theorem.

Now let us prove the following congruence known as ***Fermat's Little Theorem***:

$$a^{p-1} \equiv 1 \quad \text{mod } p \ \text{ whenever } p \text{ is an odd prime not dividing } a.$$

To show this, note first that the numbers $a, 2a, 3a, \cdots, (p-1)a$ are all distinct mod p since we know that a has a multiplicative inverse mod p, so in a congruence $ma \equiv na$ we can multiply both sides by the inverse of a to deduce that $m \equiv n$. Let us call this property that $ma \equiv na$ implies $m \equiv n$ the *cancellation property* for congruences mod p.

It follows from the cancellation property that the set $\{a, 2a, 3a, \cdots, (p-1)a\}$ is the same mod p as $\{1, 2, 3, \cdots, p-1\}$ since both sets have $p-1$ elements and neither set contains numbers that are 0 mod p. (If $ma \equiv 0$ then multiplying by the inverse of a gives $m \equiv 0$.) If we take the product of all the numbers in each of these two sets we obtain the following congruence:

$$(a)(2a)(3a)\cdots(p-1)a \equiv (1)(2)(3)\cdots(p-1) \quad \text{mod } p$$

We can cancel the factors $2, 3, \cdots, p-1$ from both sides by repeated applications of the cancellation property. The result is the congruence $a^{p-1} \equiv 1$ claimed by Fermat's Little Theorem.

Now we can prove Euler's formula for $\left(\frac{a}{p}\right)$. The first case is that $\left(\frac{a}{p}\right) = +1$. Then $a \equiv x^2$ for some $x \not\equiv 0$ and $a^{\frac{p-1}{2}} \equiv x^{p-1}$ so by Fermat's Little Theorem we have $a^{\frac{p-1}{2}} \equiv 1$. Thus Euler's formula $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}}$ is valid in this case since both sides are $+1$.

The other case is that $\left(\frac{a}{p}\right) = -1$ so a is not a square mod p. Observe first that the congruence $xy \equiv a$ has a solution y mod p for each $x \not\equiv 0$ since x has an inverse x^{-1} mod p so we can take $y = x^{-1}a$. Moreover the solution y is unique mod p since $xy_1 \equiv xy_2$ implies $y_1 \equiv y_2$ by the cancellation property. Since we are in the case that a is not a square mod p the solution y of $xy \equiv a$ satisfies $y \not\equiv x$. Thus the numbers $1, 2, 3, \cdots, p-1$ are divided up into $\frac{p-1}{2}$ pairs $\{x_1, y_1\}$, $\{x_2, y_2\}$, $\cdots$, $\{x_{\frac{p-1}{2}}, y_{\frac{p-1}{2}}\}$ with $x_i y_i \equiv a$ for each i. Multiplying these $\frac{p-1}{2}$ pairs together, we get:

$$a^{\frac{p-1}{2}} \equiv x_1 y_1 x_2 y_2 \cdots x_{\frac{p-1}{2}} y_{\frac{p-1}{2}}$$

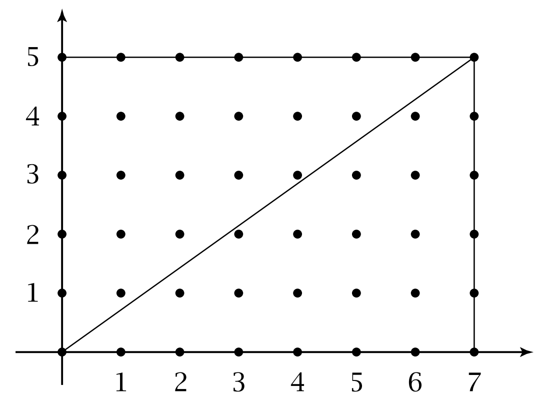
The product on the right is just a rearrangement of $(1)(2)(3)\cdots(p-1)$, and Wilson's Theorem says that this product is congruent to -1 mod p. Thus we see that Euler's formula $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}}$ holds also when $\left(\frac{a}{p}\right) = -1$, completing the proof in both cases.

A consequence of Euler's formula is the multiplicative property of Legendre symbols that we stated and used earlier in the chapter:

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

This holds since $(ab)^{\frac{p-1}{2}} = a^{\frac{p-1}{2}} b^{\frac{p-1}{2}}$.

Step 2. Our goal here will be to express the Legendre symbol $\left(\frac{a}{p}\right)$ in more geometric terms. To begin, consider a rectangle in the first quadrant of the xy-plane that is p units wide and a units high, with one corner at the origin and the opposite corner at the point (p,a). The picture at the right shows the case $(p,a) = (7,5)$. We will be interested in points that lie strictly in the interior of the rectangle and whose coordinates are integers. Points satisfying the latter condition are called *lattice points*. The number of lattice points in the interior is then $(p-1)(a-1)$ since their x-coordinates can range from 1 to $p-1$ and their y-coordinates from 1 to $a-1$, independently.



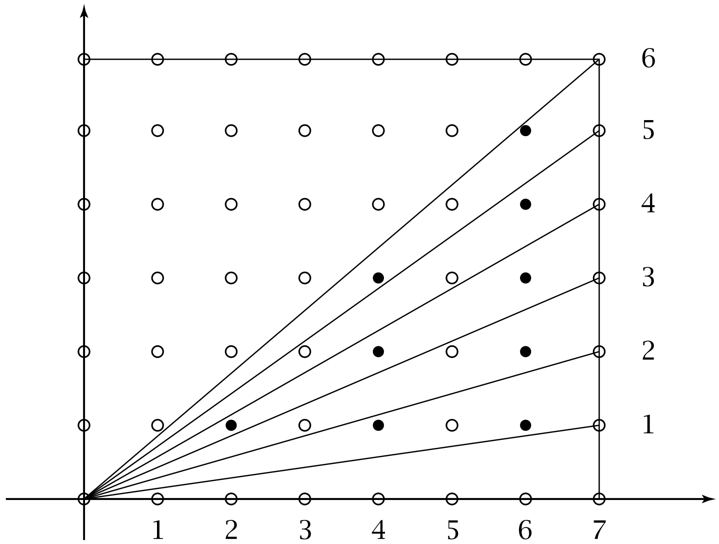
The diagonal of the rectangle from $(0,0)$ to (p,a) does not pass through any of these interior lattice points since we assume that the prime p does not divide a, so the fraction a/p, which is the slope of the diagonal, is in lowest terms. (If there were an interior lattice point on the diagonal, the slope of the diagonal would be a fraction with numerator and denominator smaller than a and p.) Since there are no interior lattice points on the diagonal, exactly half of the lattice points inside the rectangle lie on each side of the diagonal, so the number of lattice points below the diagonal is $\frac{1}{2}(p-1)(a-1)$. This is an integer since p is odd, which makes $p-1$ even.

A more refined question one can ask is how many lattice points below the diagonal have even x-coordinate and how many have odd x-coordinate. Here there is no guarantee that these two numbers must be equal, and indeed if they were equal then both numbers would have to be $\frac{1}{4}(p-1)(a-1)$ but this fraction need not be an integer, for example when $p = 7$ and $a = 4$.

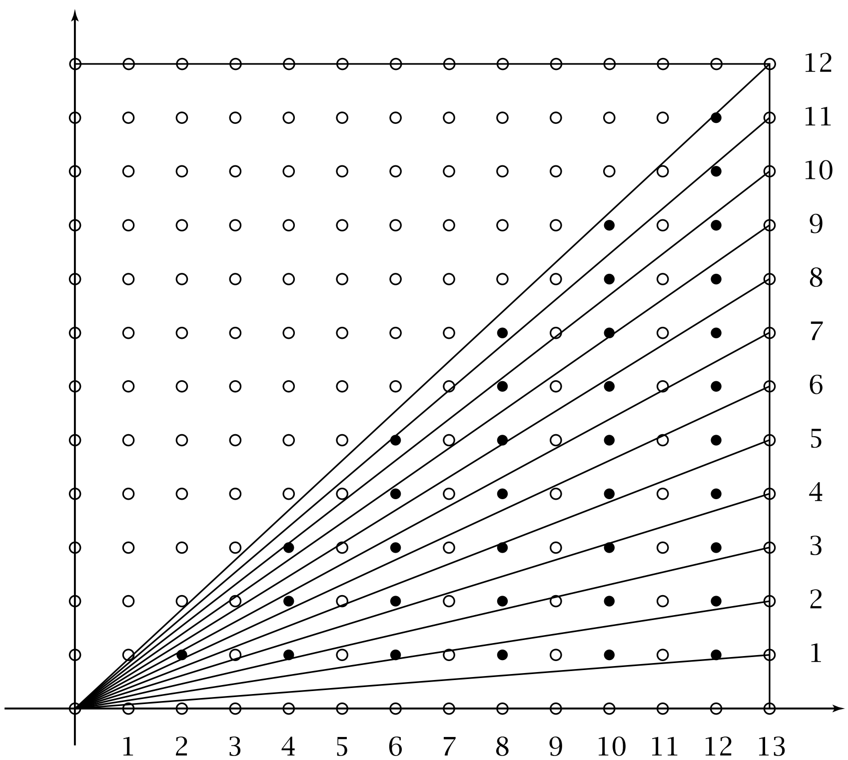
We denote the number of lattice points that are below the diagonal and have even x-coordinate by the letter e. The cases $p = 7$ and $p = 13$ are illustrated in the figures on the next page, with a ranging from 1 to 6 when $p = 7$ and from 1 to 12 when $p = 13$. The corresponding values of e count the number of black dots below the line from the origin to the point (p,a). The values of $\left(\frac{a}{p}\right)$ are also listed. The way that e varies with a seems somewhat unpredictable, but one can observe that $\left(\frac{a}{p}\right)$ is $+1$ when e is even and -1 when e is odd in these examples with $p = 7$ and $p = 13$.

We will show that this simple relationship between e and $\left(\frac{a}{p}\right)$ holds in general:

$$\left(\frac{a}{p}\right) = (-1)^e$$



a	e	$\left(\frac{a}{7}\right)$
6	9	−1
5	7	−1
4	6	+1
3	3	−1
2	2	+1
1	0	+1



a	e	$\left(\frac{a}{13}\right)$
12	36	+1
11	33	−1
10	30	+1
9	26	+1
8	23	−1
7	21	−1
6	15	−1
5	13	−1
4	10	+1
3	6	+1
2	3	−1
1	0	+1

To prove the formua $\left(\frac{a}{p}\right) = (-1)^e$ we first derive a formula for e. The segment of the vertical line $x = u$ between the x-axis and the diagonal has length $u \cdot {}^a\!/_p = {}^{ua}\!/_p$ since the slope of the diagonal is ${}^a\!/_p$. If u is a positive integer, the number of lattice points on this line segment is $\lfloor {}^{ua}\!/_p \rfloor$, the greatest integer $n \leq {}^{ua}\!/_p$. If we add up these numbers of lattice points for u running through the set of even numbers $E = \{2, 4, \cdots, p-1\}$ we get:

$$e = \sum_E \lfloor {}^{ua}\!/_p \rfloor$$

The way to compute $\lfloor {}^{ua}\!/_p \rfloor$ is to apply the division algorithm for integers, dividing p into ua to obtain $\lfloor {}^{ua}\!/_p \rfloor$ as the quotient with a remainder that we denote $r(u)$.

Thus we have:

$$ua = p\lfloor {}^{ua}/_{p} \rfloor + r(u) \tag{1}$$

The formula $ua = p\lfloor {}^{ua}/_{p} \rfloor + r(u)$ implies that $\lfloor {}^{ua}/_{p} \rfloor$ has the same parity as $r(u)$ since u is even and p is odd. Hence $\sum_E \lfloor {}^{ua}/_{p} \rfloor$ has the same parity as $\sum_E r(u)$. Since $e = \sum_E \lfloor {}^{ua}/_{p} \rfloor$, this implies that the number $(-1)^e$ that we are interested in can be computed as:

$$(-1)^e = (-1)^{\sum_E r(u)} \tag{2}$$

With this last expression in mind we will focus our attention on the remainders $r(u)$.

The number $r(u)$ lies strictly between 0 and p and can be either even or odd, but in both cases we can say that $(-1)^{r(u)} r(u)$ is congruent to an even number in the interval $(0, p)$ since if $r(u)$ is odd, so is $(-1)^{r(u)} r(u)$ and then adding p to this gives an even number between 0 and p. Thus there is always an even number $s(u)$ between 1 and p that is congruent to $(-1)^{r(u)} r(u)$ mod p. Obviously $s(u)$ is unique since no two numbers in the interval $(0, p)$ are congruent mod p.

A key fact about these even numbers $s(u)$ is that they are all distinct as u varies over the set E. For suppose we have $s(u) = s(v)$ for another even number v in E. Thus $r(u) \equiv \pm r(v)$ mod p, which implies $au \equiv \pm av$ mod p in view of the equation (1) above. We can cancel the a from both sides of the congruence $au \equiv \pm av$ to get $u \equiv \pm v$. However we cannot have $u \equiv -v$ because the number between 0 and p that is congruent to $-v$ is $p - v$, so we would have $u = p - v$ which is impossible since u and v are even while p is odd. Thus we must have $u \equiv +v$, hence $u = v$ since these are numbers strictly between 0 and p. This shows that the numbers $s(u)$ are all distinct.

Now consider the product of all the numbers $(-1)^{r(u)} r(u)$ as u ranges over the set E. Written out, this is:

$$\Big[(-1)^{r(2)} r(2)\Big]\Big[(-1)^{r(4)} r(4)\Big] \cdots \Big[(-1)^{r(p-1)} r(p-1)\Big] \tag{3}$$

By equation (1) we have $r(u) \equiv ua$ mod p, so this product is congruent mod p to:

$$\Big[(-1)^{r(2)} 2a\Big]\Big[(-1)^{r(4)} 4a\Big] \cdots \Big[(-1)^{r(p-1)} (p-1)a\Big]$$

On the other hand, by the definition of the numbers $s(u)$ the product (3) is congruent mod p to $[s(2)][s(4)] \cdots [s(p-1)]$. There are $\frac{1}{2}(p-1)$ factors here and they are all distinct even numbers in the interval $(0, p)$ as we showed in the previous paragraph, so they are just a rearrangement of the numbers $2, 4, \cdots, p-1$. Thus we have the following congruence:

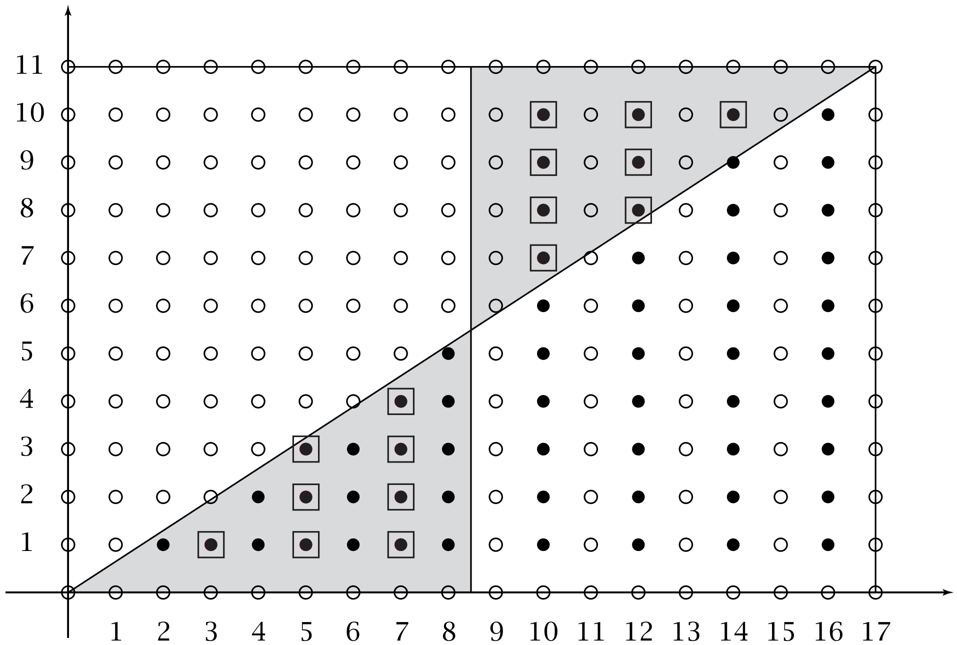
$$\Big[(-1)^{r(2)} 2a\Big]\Big[(-1)^{r(4)} 4a\Big] \cdots \Big[(-1)^{r(p-1)} (p-1)a\Big] \equiv (2)(4) \cdots (p-1) \mod p$$

Canceling the factors $2, 4, \cdots, p-1$ from both sides of this congruence gives:

$$(-1)^{\sum_E r(u)} a^{\frac{p-1}{2}} \equiv 1 \mod p$$

Both the factors $(-1)^{\sum_E r(u)}$ and $a^{\frac{p-1}{2}}$ are ± 1 mod p and their product is 1 so they must be equal mod p (using the fact that 1 and -1 are not congruent modulo an odd prime). By Euler's formula we have $a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right)$ mod p, so from the earlier formula (2) we conclude that $\left(\frac{a}{p}\right) = (-1)^e$. This finishes Step 2.

Step 3. Now we specialize the value of a to be an odd prime q distinct from p. As in Step 2 we consider lattice points in the interior of a $p \times q$ rectangle.



From Step 2 we know that $\left(\frac{q}{p}\right) = (-1)^e$ where e is the number of lattice points with even x-coordinate inside the rectangle and below the diagonal. Suppose that we divide the rectangle into two equal halves separated by the vertical line $x = p/2$ which does not pass through any lattice points since p is odd. This vertical line cuts off two smaller triangles from the two large triangles above and below the diagonal of the rectangle. In the figure above, these smaller triangles are the shaded triangles. Call the lower small triangle L and the upper one U, and let l and u denote the number of lattice points with even x-coordinate in the interiors of L and U respectively. Note that u has the same parity as the number of lattice points with even x-coordinate in the interior of the quadrilateral below U in the right half of the rectangle since each column of lattice points inside the rectangle has $q-1$ points, an even number. Thus e has the same parity as $l+u$, hence $(-1)^e = (-1)^{l+u}$.

The next thing to notice is that rotating the triangle U by 180 degrees about the center of the rectangle carries it onto the triangle L. This rotation takes the lattice points inside U with even x-coordinate onto the lattice points inside L with odd x-coordinate. Thus we obtain the formula $\left(\frac{q}{p}\right) = (-1)^t$ where t is the total number of lattice points inside the triangle L.

Reversing the roles of p and q, we can also say that $\left(\frac{p}{q}\right) = (-1)^{t'}$ where t' is the number of lattice points inside the triangle L' with edges on the diagonal of the rectangle, the horizontal line $y = {}^q\!/_2$, and the y-axis. Then $t + t'$ is the number of lattice points in the interior of the small rectangle formed by L and L' together. This number is just $\frac{p-1}{2} \cdot \frac{q-1}{2}$. Thus we have

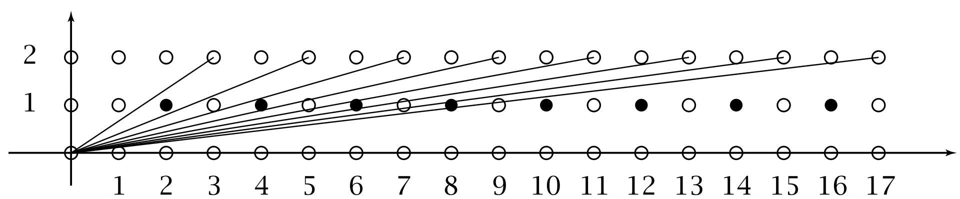
$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^t(-1)^{t'} = (-1)^{t+t'} = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$$

which finally finishes the proof of quadratic reciprocity. □

We can also use the geometric interpretation of $\left(\frac{a}{p}\right)$ to prove the formula for $\left(\frac{2}{p}\right)$ that was given in Section 6.2, namely:

$$\left(\frac{2}{p}\right) = \begin{cases} +1 & \text{if } p = 8k \pm 1 \\ -1 & \text{if } p = 8k \pm 3 \end{cases}$$

We have shown that $\left(\frac{2}{p}\right) = (-1)^e$ where e is the number of lattice points inside a $p \times 2$ rectangle lying below the diagonal and having even x-coordinate, as indicated in the following figure which shows the diagonals for $p = 3, 5, 7, \cdots, 17$:



p	3	5	7	9	11	13	15	17
e	1	1	2	2	3	3	4	4

Another way to describe e is to say that it is equal to the number of even integers in the interval from ${}^p\!/_2$ to p. We do not need to assume that p is prime in order to count these points below the diagonals, just that p is odd. One can see what the pattern is just by looking at the figure: Each time p increases by 2 there is one more even number at the right end of the interval $({}^p\!/_2, p)$, and there may or may not be one fewer even number at the left end of the interval, depending on whether p is increasing from $4k - 1$ to $4k + 1$ or from $4k + 1$ to $4k + 3$. It follows that the parity of e depends only on the value of p mod 8 as in the table for $p \le 17$, so e is even for $p \equiv \pm 1$ mod 8 and e is odd for $p \equiv \pm 3$ mod 8.

Exercises

1. As a sort of converse to Wilson's Theorem, show that if n is not a prime then $(n-1)!$ is not congruent to -1 mod n. More precisely, when $n > 4$ and n is not prime, show that n divides $(n-1)!$, so $(n-1)! \equiv 0$ mod n. What happens when $n = 4$?

2. In Step 2 of the proof of quadratic reciprocity there were figures depicting the geometric interpretation of $\left(\frac{a}{7}\right)$ and $\left(\frac{a}{13}\right)$. Draw analogous figures for $\left(\frac{a}{5}\right)$ and $\left(\frac{a}{11}\right)$.

3. Show that the calculation of the Legendre symbol $\left(\frac{-1}{p}\right)$ can also be obtained using the method in the proof of quadratic reciprocity involving counting certain lattice points in a $(p-1) \times p$ rectangle.

7 The Class Group for Quadratic Forms

In the previous chapter we determined which numbers n are represented by at least one form of a given discriminant, where the numbers represented by a form Q are the numbers that appear in the topograph of Q, so we consider only the values $Q(x,y)$ for primitive pairs (x,y). The answer was in terms of certain congruence conditions on the prime divisors of n. We could also determine the genus of the forms representing n via congruence conditions.

What one would really like to do is refine these results to determine which equivalence classes of forms represent n, and for this it is natural to consider only primitive forms. The hardest part of the problem is determining which primes each primitive form represents. Much is known about this, but it requires considerably deeper mathematics than we can cover in this book so we will say nothing more about representing primes beyond what we have already discussed concerning genus. Instead, what we will do in the present chapter is study the question for nonprimes, assuming one already knows which primes each form represents. For fundamental discriminants we will obtain a fairly complete picture, while for nonfundamental discriminants there will remain certain ambiguities, with examples showing the extra complication in these cases.

The main tool will be a method for multiplying forms of a given discriminant that corresponds to multiplying the numbers represented by these forms. This multiplication of forms gives rise to a commutative group structure on the set of proper equivalence classes of primitive forms of a given discriminant. This group, called the *class group* and denoted $CG(\Delta)$ for discriminant Δ, also has other uses besides determining the forms representing nonprimes. For example we will use it to give a good explanation for why the number of genera in a given discriminant is equal to the number of equivalence classes of primitive forms in that discriminant whose topographs have mirror symmetry.

In this chapter we will restrict attention entirely to forms of nonsquare discriminant, which means elliptic and hyperbolic forms. For elliptic forms we only consider those with positive values, as usual.

7.1 Multiplication of Forms

Since we will often be dealing with several different forms at a time it will be convenient to shorten the notation by writing a form $ax^2 + bxy + cy^2$ simply as $[a, b, c]$, retaining only the essential information of the coefficients. We are restricting attention to discriminants that are not squares so the outer coefficients a and c must always be nonzero.

Recall that a number a is represented by a form Q if and only if a appears in the topograph of Q, and this in turn is equivalent to a appearing as the leading coefficient of a form $[a, b, c]$ equivalent to Q. A simple observation is that if a factors as $a = a_1a_2$ then the forms $[a_1a_2, b, c]$, $[a_1, b, a_2c]$, and $[a_2, b, a_1c]$ all have the same discriminant. This shows that if a number a is represented in discriminant Δ then so is each divisor of a, as we saw in Proposition 6.1.

A form $[a_1a_2, b, c]$ can thus be split into two forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ of the same discriminant. One might wonder about the reverse process of combining or "multiplying" the two forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ to obtain the form $[a_1a_2, b, c]$. For example the product of $[2, 0, 15]$ and $[3, 0, 10]$ would be $[6, 0, 5]$. The main goal in this section will be to show that this simple way to multiply certain special pairs of forms is nevertheless sufficiently general to give a well-defined multiplication operation on the set of proper equivalence classes of primitive forms of a given discriminant.

A pair of forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ is said to be ***concordant***. For two forms to be concordant is obviously a very strong condition since not only are the second coefficients of the two forms equal, but also the first coefficient of each form divides the third coefficient of the other form. Furthermore, the discriminants of the two forms are equal. Conversely, suppose that two forms $[a_1, b, c_1]$ and $[a_2, b, c_2]$ with the same middle coefficient have the same discriminant. Then $a_1c_1 = a_2c_2$, so if a_1 divides c_2, say $c_2 = a_1c$ for some integer c, then $a_1c_1 = a_2c_2 = a_2a_1c$ so in particular $a_1c_1 = a_2a_1c$, and since a_1 is nonzero we can cancel it from this equation to get $c_1 = a_2c$. The two forms are thus $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ so they are concordant. This argument shows in fact that for two forms $[a_1, b, c_1]$ and $[a_2, b, c_2]$ of the same discriminant, if a_1 divides c_2 then it automatically follows that a_2 divides c_1.

Since we wish to consider only primitive forms the following result will be useful:

Lemma 7.1. *If the concordant forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ are primitive then so is their product $[a_1a_2, b, c]$. If a_1 and a_2 are coprime then the converse is also true: If $[a_1a_2, b, c]$ is primitive then so are $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$.*

An extra condition is needed in the converse since for example the primitive form $[4,0,1]$ factors as the product of the nonprimitive concordant forms $[2,0,2]$ and $[2,0,2]$.

Proof: If the coefficients of $[a_1a_2,b,c]$ have a common divisor then they have a common prime divisor, which will divide either a_1 or a_2, as well as b and c, so one of the forms $[a_1,b,a_2c]$ and $[a_2,b,a_1c]$ will not be primitive. This gives the first statement. For the second, if one of $[a_1,b,a_2c]$ and $[a_2,b,a_1c]$ is not primitive, say $[a_1,b,a_2c]$, then its coefficients will be divisible by some prime p. If a_1 and a_2 are coprime, then p dividing a_1 and a_2c implies that p divides c. Thus p divides all three coefficients of $[a_1a_2,b,c]$, making it nonprimitive. □

Proposition 7.2. *For each pair of primitive forms Q_1 and Q_2 of discriminant Δ there is a pair of primitive forms $Q_1' = [a_1,b,a_2c]$ and $Q_2' = [a_2,b,a_1c]$ which are concordant to each other and properly equivalent to Q_1 and Q_2 respectively. The forms Q_1' and Q_2' can be chosen so that $a_1 > 0$ and $a_2 > 0$.*

For the proof we will need the following result which will be useful on other occasions as well:

Lemma 7.3. *For each pair of forms $Q_1 = [a_1,b_1,c_1]$ and $Q_2 = [a_2,b_2,c_2]$ of the same discriminant with a_1 and a_2 coprime there exists a pair of forms $[a_1,b,a_2c]$ and $[a_2,b,a_1c]$ that are concordant to each other and properly equivalent to Q_1 and Q_2 respectively.*

Proof: The main step will be to find two forms properly equivalent to Q_1 and Q_2 that have the same first coefficients as Q_1 and Q_2 and have equal second coefficients. To do this we begin by recalling that the edges in the topograph of a form have integer labels, with the sign of a label changing when the orientation of the edge is reversed. For a region in the topograph of Q_1 labeled a_1 let us orient the edges bordering this region all in the same direction so that the region lies to the left as we move along the edges in the direction specified by their orientation. The edge labels then form an arithmetic progression with increment $2a_1$. One of these edges is labeled b_1, so the other edge labels are $b_1 + 2a_1m$ for m varying over all integers. Similarly, in the topograph of Q_2 we have a region labeled a_2 whose bordering edges have labels $b_2 + 2a_2n$ for all integers n.

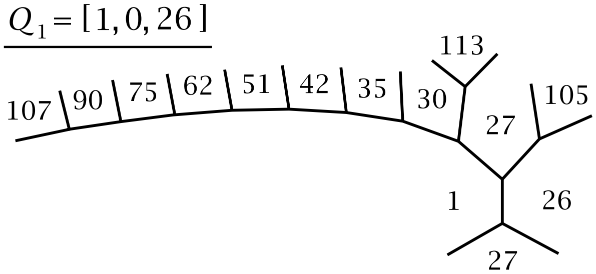
We would like one of the edge labels $b_1 + 2a_1m$ to equal one of the edge labels b_2+2a_2n. This means we would like to find integers m and n satisfying the equation $b_1+2a_1m = b_2+2a_2n$, or equivalently $a_1m-a_2n = (b_2-b_1)/2$. Note that the right side of this equation is an integer since the edge labels in a topograph always have the same parity as the discriminant, which is the same for both forms by assumption. From Section 2.3 we know the equation $a_1m-a_2n = (b_2-b_1)/2$ always has an integer solution (m,n) if a_1 and a_2 are coprime. Thus we can find edges bordering the a_1 and a_2 regions with the same label b. The two given forms are therefore equivalent to forms $[a_1, b, c_1']$ and $[a_2, b, c_2']$, and in fact properly equivalent because of the way we have oriented the edges bordering the a_1 and a_2 regions.

Equating the discriminants of these two forms $[a_1, b, c_1']$ and $[a_2, b, c_2']$ leads to the equation $a_1c_1' = a_2c_2'$. Since a_1 and a_2 are coprime this implies that a_1 divides c_2', so $c_2' = a_1c$ for some integer c. The equation $a_1c_1' = a_2c_2'$ then becomes $a_1c_1' = a_2a_1c$, which implies that $c_1' = a_2c$ since a_1 is nonzero. Thus we have two concordant forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ properly equivalent to the original forms $[a_1, b_1, c_1]$ and $[a_2, b_2, c_2]$. □

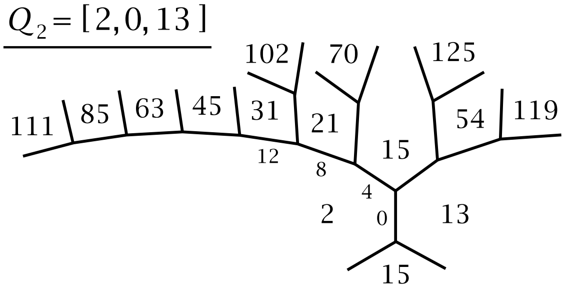
Proof of Proposition 7.2: Choose a number $a_1 > 0$ in the topograph of Q_1. By Proposition 6.26 the topograph of Q_2 contains some number $a_2 > 0$ coprime to a_1. Thus Q_1 and Q_2 are properly equivalent to forms $[a_1, b_1, c_1]$ and $[a_2, b_2, c_2]$, and then Lemma 7.3 finishes the proof. □

To illustrate how to multiply forms let us look at a few examples in the case $\Delta = -104$. Here there are four equivalence classes of forms:

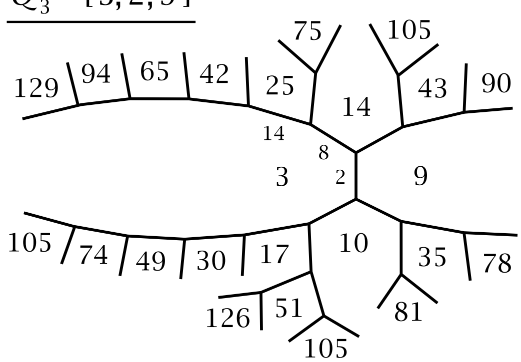
$Q_1 = [1, 0, 26]$



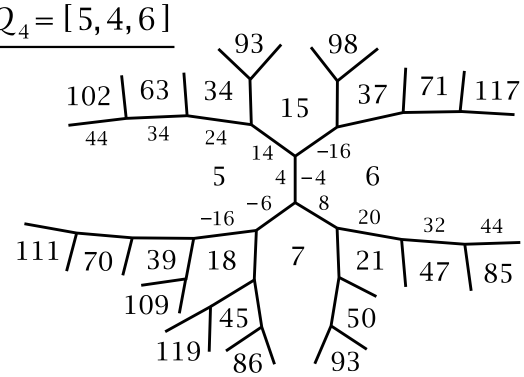
$Q_2 = [2, 0, 13]$



$Q_3 = [3, 2, 9]$



$Q_4 = [5, 4, 6]$



Since only the first two forms have mirror symmetry, the class number is 6. We will be somewhat free with the notation and use the same symbol Q_i to denote any form properly equivalent to the original form Q_i.

Let us compute the product of Q_2 and Q_3 using the method in the proof of Lemma 7.3. To begin we need regions in the topographs of Q_2 and Q_3 with coprime labels, so the simplest thing is to use the region labeled 2 in the topograph of Q_2 and the region labeled 3 in the topograph of Q_3. For the Q_2 topograph the edge between the 2 and 13 regions is labeled 0 so the next edges bordering the 2 region are labeled $4, 8, 12, \cdots$. For the 3 region in the topograph of Q_3 the bordering edges are labeled $2, 8, 14, \cdots$ starting with the edge adjacent to the 9 region. The number 8 is in both these arithmetic progressions so we choose this for b. In the Q_2 topograph this edge labeled 8 is between the regions labeled 2 and 21 so the form we want is $[2,8,21]$. For Q_3 the edge labeled 8 is between the 3 and 14 regions so the form corresponding to this edge is $[3,8,14]$. The product of these two concordant forms is then $[6,8,7]$. The values of this form at $(x,y)=(0,1)$, $(1,0)$, and $(1,1)$ are 6, 7, and 21 so from the topograph of Q_4 we see that this form is properly equivalent to Q_4. Thus we have $Q_2Q_3=Q_4$.

The product Q_4Q_4, or in other words Q_4^2, can be computed in the same way using the regions in the topograph of Q_4 with the coprime labels 5 and 6. For the edges bordering the 5 region the labels starting with the edge between the 5 and 6 regions are $4,14,24,\cdots$. For the edges bordering the 6 region we can start with the same edge but now this edge must be oriented in the opposite direction in order to have the 6 region on our left as we move forward. The edge labels are then $-4,8,20,\cdots$. Continuing these arithmetic progressions a little farther we find the common label 44 on the edge between the 5 and 102 regions, and on the edge between the 6 and 85 regions. Thus we have the concordant forms $[5,44,102]$ and $[6,44,85]$, with product $[30,44,17]$. The coefficients 30 and 17 appear in adjacent regions in the topograph of Q_3 so Q_4^2 is properly equivalent to either Q_3 or the mirror image form. We can determine which by evaluating $[30,44,17]$ at $(x,y)=(-1,1)$, giving the value 3. Thus in the topograph of $[30,44,17]$ the values $30,17,3$ appear in clockwise order around a vertex, while in the topograph of Q_3 they are in counterclockwise order, so these two topographs are mirror images and hence Q_4^2 is properly equivalent to the mirror image form of Q_3.

In these examples there were a number of choices made in order to compute the products Q_iQ_j. Thus for computing Q_2Q_3 we first chose the regions labeled 2 and 3 in the topographs of Q_2 and Q_3, but we could have chosen any region in one topograph and then chosen any of the infinitely many regions in the other topograph with a label coprime to the label of the first region. After choosing the regions labeled 2 and 3 we then chose edges bordering these regions having the same label b, and there are infinitely many possibilities to choose from here too. For the 2 region the edge labels are the integers $8+4k$ and for the 3 region they are the integers $8+6k$ so the common edge labels are the integers $8+12k$, which are in fact all the edge labels for the 6 region in the topograph of Q_4. It is not at all obvious that the various choices that were made for the two topographs always lead to the same result that $Q_2Q_3=Q_4$. Our next task will be to prove that this is true not just for this calculation but in general.

What we wish to prove is the crucial fact that multiplication of proper equivalence classes of primitive forms by choosing a concordant pair of forms in these classes does not depend on which concordant pair we choose. This can be phrased in the following way:

Proposition 7.4. *For a fixed discriminant let Q_1, Q_2 be a pair of concordant primitive forms and let Q_1', Q_2' be another such pair properly equivalent to Q_1 and Q_2 respectively. Then the products Q_1Q_2 and $Q_1'Q_2'$ are properly equivalent.*

The proof will involve a certain amount of calculation, and to ease the burden it will be convenient to express quadratic forms in terms of matrices. This is based on the simple observation that a form $ax^2 + bxy + cy^2$, regarded as a 1×1 matrix $(ax^2 + bxy + cy^2)$, can be obtained as a product of a 1×2 matrix, a 2×2 matrix, and a 2×1 matrix:

$$\begin{aligned}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} ax + by/2 & bx/2 + cy \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} ax^2 + bxy + cy^2 \end{pmatrix}\end{aligned}$$

Thus we are expressing the form $ax^2 + bxy + cy^2$ as a matrix $\begin{pmatrix} a & \underline{b} \\ \underline{b} & c \end{pmatrix}$ where $\underline{b} = b/2$. The entries $\underline{b}$ might not be integers, but this will not matter for our purposes.

When we do a change of variables by means of a matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ with determinant $ps - qr = 1$, replacing $\begin{pmatrix} x \\ y \end{pmatrix}$ by $\begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} px+qy \\ rx+sy \end{pmatrix}$, then the product $(x\ y)\begin{pmatrix} a & \underline{b} \\ \underline{b} & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$ becomes $(x\ y)\begin{pmatrix} p & r \\ q & s \end{pmatrix}\begin{pmatrix} a & \underline{b} \\ \underline{b} & c \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$, with the second matrix being the transpose of the fourth matrix. Thus the matrix $\begin{pmatrix} a & \underline{b} \\ \underline{b} & c \end{pmatrix}$ for the form $ax^2 + bxy + cy^2$ is replaced by the matrix $\begin{pmatrix} a' & \underline{b}' \\ \underline{b}' & c' \end{pmatrix} = \begin{pmatrix} p & r \\ q & s \end{pmatrix}\begin{pmatrix} a & \underline{b} \\ \underline{b} & c \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ for the new form $a'x^2 + b'xy + c'y^2$. We can write this last equation as

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix}\begin{pmatrix} a & \underline{b} \\ \underline{b} & c \end{pmatrix} = \begin{pmatrix} a' & \underline{b}' \\ \underline{b}' & c' \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{-1} = \begin{pmatrix} a' & \underline{b}' \\ \underline{b}' & c' \end{pmatrix}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$$

where this last matrix is the inverse of $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ since $ps - qr = 1$.

Proof of Proposition 7.4: We will use the notation $Q \approx Q'$ to mean that the forms Q and Q' are properly equivalent.

Let $Q_1 = [a_1, b, a_2c]$ and $Q_2 = [a_2, b, a_1c]$, with $Q_1' = [a_1', b', a_2'c']$ and $Q_2' = [a_2', b', a_1'c']$. To begin the proof we look at the special case that $Q_1 = Q_1'$, so $a_1 = a_1'$, $b = b'$, and $a_2c = a_2'c'$. We assume $Q_2 \approx Q_2'$ so by the remarks preceding the proof there is an integer matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ of determinant 1 such that:

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix}\begin{pmatrix} a_2 & \underline{b} \\ \underline{b} & a_1c \end{pmatrix} = \begin{pmatrix} a_2' & \underline{b} \\ \underline{b} & a_1c' \end{pmatrix}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$$

Multiplied out, this becomes:

$$\begin{pmatrix} a_2p + \underline{b}r & \underline{b}p + a_1cr \\ a_2q + \underline{b}s & \underline{b}q + a_1cs \end{pmatrix} = \begin{pmatrix} a_2's - \underline{b}r & \underline{b}p - a_2'q \\ \underline{b}s - a_1c'r & a_1c'p - \underline{b}q \end{pmatrix} \qquad (*)$$

To show $Q_1Q_2 \approx Q_1Q_2'$ we would like to find an integer matrix $\begin{pmatrix} p' & q' \\ r' & s' \end{pmatrix}$ of determinant 1 such that:

$$\begin{pmatrix} p' & r' \\ q' & s' \end{pmatrix}\begin{pmatrix} a_1a_2 & \underline{b} \\ \underline{b} & c \end{pmatrix} = \begin{pmatrix} a_1a_2' & \underline{b} \\ \underline{b} & c' \end{pmatrix}\begin{pmatrix} s' & -q' \\ -r' & p' \end{pmatrix}$$

This becomes:

$$\begin{pmatrix} a_1a_2p' + \underline{b}r' & \underline{b}p' + cr' \\ a_1a_2q' + \underline{b}s' & \underline{b}q' + cs' \end{pmatrix} = \begin{pmatrix} a_1a_2's' - \underline{b}r' & \underline{b}p' - a_1a_2'q' \\ \underline{b}s' - c'r' & c'p' - \underline{b}q' \end{pmatrix} \qquad (**)$$

We can convert the upper left entries in the two matrices in $(*)$ to the corresponding entries in $(**)$ by multiplying by a_1 if we choose $p' = p$, $s' = s$, and $r' = a_1r$. Then the equality of the upper left entries in $(*)$ will imply equality of the upper left entries in $(**)$. If we further choose $q' = q/a_1$ then the upper right entries in $(*)$ will be equal to the corresponding entries in $(**)$, and the same will be true for the lower left entries. The lower right entries in $(*)$ will be a_1 times those in $(**)$ so the lower right entries in $(**)$ will be equal as well. Thus we hope to define $\left(\begin{smallmatrix} p' & q' \\ r' & s' \end{smallmatrix}\right)$ by:

$$\begin{pmatrix} p' & q' \\ r' & s' \end{pmatrix} = \begin{pmatrix} p & q/a_1 \\ a_1r & s \end{pmatrix}$$

Note that this matrix has the same determinant as $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$. The only problem is that the entry $q' = q/a_1$ will only be an integer if a_1 divides q. To guarantee that it does, observe that the equality of the upper right entries in $(*)$ implies that $a_1cr = -a_2'q$, so if a_1 is coprime to a_2' then a_1 will divide q. Thus we have proved the proposition in the special case $Q_1 = Q_1'$ provided that a_1 and a_2' are coprime.

In the case just considered we assumed $Q_1 = Q_1'$ which implied that $b = b'$. Now let us assume merely that $b = b'$ along with the previous hypothesis that a_1 and a_2' are coprime. Under these conditions the desired equivalence $Q_1Q_2 \approx Q_1'Q_2'$ will be obtained as the combination of two equivalences $Q_1Q_2 \approx Q_1Q_2' \approx Q_1'Q_2'$, but first we have to check that Q_1 and Q_2' are concordant so that Q_1Q_2' is defined. Since $b = b'$ and the determinants of Q_1 and Q_1' are equal, we have $a_1a_2c = a_1'a_2'c'$. Since a_1 and a_2' are coprime it follows that a_1 divides $a_1'c'$. As we saw earlier, this implies that the forms $Q_1 = [a_1, b, a_2c]$ and $Q_2' = [a_2', b, a_1'c']$ are concordant.

Assuming that a_1 and a_2' are coprime, the previous case $Q_1 = Q_1'$ now gives an equivalence $Q_1Q_2 \approx Q_1Q_2'$. Switching the roles of Q_1 and Q_2' as well as Q_1' and Q_2, this argument also shows $Q_1Q_2' \approx Q_1'Q_2'$ using the same assumption that a_2' and a_1 are coprime. We conclude that $Q_1Q_2 \approx Q_1'Q_2'$ when a_1 and a_2' are coprime and $b = b'$.

Next we consider how to arrange that $b = b'$. The hypothesis that will allow this is that a_1a_2 is coprime to $a_1'a_2'$, which is equivalent to saying that each of a_1 and a_2 is coprime to each of a_1' and a_2'. If a_1a_2 and $a_1'a_2'$ are coprime, we know by an argument in the proof of Lemma 7.3 that the arithmetic progressions $b + a_1a_2m$ and $b' + a_1'a_2'n$ have a common value B. This will also be a value in each of the arithmetic progressions $b + a_1m$, $b + a_2n$, $b' + a_1'm$, and $b' + a_2'n$. Thus we have forms $\widetilde{Q}_i = [a_i, B, \widetilde{c}_i] \approx Q_i$ for $i = 1, 2$, and similarly $\widetilde{Q}_i' = [a_i', B, \widetilde{c}_i'] \approx Q_i'$.

Let us check that $\widetilde{Q}_1$ and $\widetilde{Q}_2$ are concordant. This will be true if the first coefficient of one form divides the third coefficient of the other, say a_2 divides $\widetilde{c}_1$. The forms Q_1 and $\widetilde{Q}_1$ have the same discriminant so $b^2-4a_1a_2c = B^2-4a_1\widetilde{c}_1$. Substituting $B = b+2a_1a_2m$ and simplifying, we get $-a_1a_2c = a_1a_2bm+a_1^2a_2^2m^2-a_1\widetilde{c}_1$. After canceling a factor of a_1 from both sides, this becomes $-a_2c = a_2bm+a_1a_2^2m^2-\widetilde{c}_1$ which implies that a_2 divides $\widetilde{c}_1$. Thus $\widetilde{Q}_1$ and $\widetilde{Q}_2$ are concordant, and by the same reasoning $\widetilde{Q}_1'$ and $\widetilde{Q}_2'$ are concordant, so we can form the products $\widetilde{Q}_1\widetilde{Q}_2$ and $\widetilde{Q}_1'\widetilde{Q}_2'$.

We have $Q_1Q_2 \approx \widetilde{Q}_1\widetilde{Q}_2$ since the label B occurs on an edge bordering the region labeled a_1a_2 in the topographs of both of these product forms, which is obvious for $\widetilde{Q}_1\widetilde{Q}_2 = [a_1a_2, B, -]$ while for $Q_1Q_2 = [a_1a_2, b, -]$ it follows from the definition of B. Similarly $Q_1'Q_2' \approx \widetilde{Q}_1'\widetilde{Q}_2'$. We can now apply the previous case $b = b'$ to the four forms $\widetilde{Q}_1, \widetilde{Q}_2, \widetilde{Q}_1', \widetilde{Q}_2'$ since the leading coefficients a_1 and a_2' of the first and fourth forms are coprime. Thus we have $\widetilde{Q}_1\widetilde{Q}_2 \approx \widetilde{Q}_1'\widetilde{Q}_2'$ and hence $Q_1Q_2 \approx \widetilde{Q}_1\widetilde{Q}_2 \approx \widetilde{Q}_1'\widetilde{Q}_2' \approx Q_1'Q_2'$. This proves the proposition under the assumption that a_1a_2 is coprime to $a_1'a_2'$.

Now we can finish the proof by reducing to the case just considered, that a_1a_2 is coprime to $a_1'a_2'$. Choose a number A_1 represented by Q_1 coprime to $a_1a_2a_1'a_2'$, and then choose a number A_2 represented by Q_2 and coprime to $A_1a_1a_2a_1'a_2'$. Since A_1 and A_2 are coprime, Lemma 7.3 implies that there exist concordant forms $\widehat{Q}_1 = [A_1, B, A_2C]$ and $\widehat{Q}_2 = [A_2, B, A_1C]$ with $\widehat{Q}_1 \approx Q_1$ and $\widehat{Q}_2 \approx Q_2$. Since A_1A_2 is coprime to a_1a_2 the previous case implies that $Q_1Q_2 \approx \widehat{Q}_1\widehat{Q}_2$. The previous case also implies that $\widehat{Q}_1\widehat{Q}_2 \approx Q_1'Q_2'$ since A_1A_2 is coprime to $a_1'a_2'$ and we have $\widehat{Q}_1 \approx Q_1 \approx Q_1'$ and $\widehat{Q}_2 \approx Q_2 \approx Q_2'$. Thus $Q_1Q_2 \approx \widehat{Q}_1\widehat{Q}_2 \approx Q_1'Q_2'$ and we are done. ☐

For proper equivalence classes of primitive forms of a fixed discriminant we have seen that if two classes represent coprime numbers, then the product class represents the product of the two numbers. The next proposition says that we can drop the coprimeness condition on the two numbers if we allow "representations" $Q(x,y) = n$ with nonprimitive pairs (x,y).

Proposition 7.5. *If Q_1 and Q_2 are concordant forms with product Q_1Q_2 then each product $Q_1(x_1,y_1)Q_2(x_2,y_2)$ can be expressed as $Q_1Q_2(X,Y)$ where X and Y are certain explicit functions of (x_1,y_1) and (x_2,y_2) given in terms of the coefficients of Q_1 and Q_2.*

Proof: Let $Q_1(x,y) = a_1x^2+bxy+a_2cy^2$ and $Q_2(x,y) = a_2x^2+bxy+a_1cy^2$. It will suffice to express a product $Q_1(x_1,y_1)Q_2(x_2,y_2)$ as $a_1a_2X^2+bXY+cY^2$ where X and Y are given in terms of the coefficients a_1,a_2,c and the variables x_1,y_1,x_2,y_2. First we compute $Q_1(x_1,y_1)Q_2(x_2,y_2)$:

$$\begin{aligned}(a_1x_1^2+bx_1y_1+a_2cy_1^2)&(a_2x_2^2+bx_2y_2+a_1cy_2^2)\\ &= \underbrace{a_1a_2x_1^2x_2^2}_{(1)} + \underbrace{a_1bx_1^2x_2y_2}_{(2)} + \underbrace{a_1^2cx_1^2y_2^2}_{(3)} + \underbrace{a_2bx_1x_2^2y_1}_{(4)} + \underbrace{b^2x_1x_2y_1y_2}_{(5)}\\ &\quad + \underbrace{a_1bcx_1y_1y_2^2}_{(6)} + \underbrace{a_2^2cx_2^2y_1^2}_{(7)} + \underbrace{a_2bcx_2y_1^2y_2}_{(8)} + \underbrace{a_1a_2c^2y_1^2y_2^2}_{(9)}\end{aligned}$$

There are nine terms here and we label them (1)–(9) as shown. We want to choose X and Y so that the sum of these nine terms is equal to $a_1a_2X^2 + bXY + cY^2$. Only the terms (1) and (9) contain the factor a_1a_2 appearing in $a_1a_2X^2$ so to get (1) it is reasonable to start with $X = x_1x_2$. Then to get (9) we expand this to:

$$X = x_1x_2 \pm cy_1y_2$$

Here we allow a sign $\pm$ for flexibility later in the calculation. Now we have:

$$a_1a_2X^2 = \underbrace{a_1a_2x_1^2x_2^2}_{(1)} \pm 2a_1a_2cx_1x_2y_1y_2 + \underbrace{a_1a_2c^2y_1y_2}_{(9)}$$

This gives (1) and (9) but the middle term does not appear among (1)–(9) so we will have to have something that cancels it later.

Next, to get the term (2) we start with $Y = a_1x_1y_2$ so that bXY starts with $a_1bx_1^2x_2y_2$ which is (2). For symmetry let us expand $Y = a_1x_1y_2$ to:

$$Y = a_1x_1y_2 + a_2x_2y_1$$

This gives:

$$bXY = \underbrace{a_1bx_1^2x_2y_2}_{(2)} + \underbrace{a_2bx_1x_2^2y_1}_{(4)} \pm \underbrace{a_1bcx_1y_1y_2^2}_{(6)} \pm \underbrace{a_2bcx_2y_1^2y_2}_{(8)}$$

$$\text{and} \quad cY^2 = \underbrace{a_1^2cx_1^2y_2^2}_{(3)} + 2a_1a_2cx_1x_2y_1y_2 + \underbrace{a_2^2cx_2^2y_1^2}_{(7)}$$

If we choose the sign $\pm$ in X to be minus then the middle term of cY^2 cancels the middle term of $a_1a_2X^2$, but this gives the terms (6) and (8) in bXY the wrong sign so we will need other terms to compensate for this. We have also not yet accounted for the term (5). To get this let us add another term to Y so that X and Y are now:

$$X = x_1x_2 - cy_1y_2$$
$$Y = a_1x_1y_2 + a_2x_2y_1 + by_1y_2$$

Then we have:

$$a_1a_2X^2 = \underbrace{a_1a_2x_1^2x_2^2}_{(1)} - 2a_1a_2cx_1x_2y_1y_2 + \underbrace{a_1a_2c^2y_1y_2}_{(9)}$$

$$bXY = \underbrace{a_1bx_1^2x_2y_2}_{(2)} + \underbrace{a_2bx_1x_2^2y_1}_{(4)} - \underbrace{a_1bcx_1y_1y_2^2}_{(6)} - \underbrace{a_2bcx_2y_1^2y_2}_{(8)}$$
$$+ \underbrace{b^2x_1x_2y_1y_2}_{(5)} - b^2cy_1^2y_2^2$$

$$cY^2 = \underbrace{a_1^2cx_1^2y_2^2}_{(3)} + 2a_1a_2cx_1x_2y_1y_2 + \underbrace{a_2^2cx_2^2y_1^2}_{(7)}$$
$$+ b^2cy_1^2y_2^2 + 2\underbrace{a_1bcx_1y_1y_2^2}_{(6)} + 2\underbrace{a_2bcx_2y_1^2y_2}_{(8)}$$

Now when we add everything up, the unlabeled terms cancel and the remaining terms combine to give precisely the terms (1)–(9). □

As a very simple illustration let us consider the case $\Delta = -24$ where there are the two reduced forms $[1,0,6]$ and $[2,0,3]$. The form $[1,0,6]$ is concordant to itself and we have $[1,0,6][1,0,6] = [1,0,6]$. Also $[1,0,6]$ is concordant to $[2,0,3]$ and we have $[1,0,6][2,0,3] = [2,0,3]$. However $[2,0,3]$ is not concordant to itself, although it is concordant to $[3,0,2]$ which is equivalent to $[2,0,3]$ and in fact properly equivalent to it since both forms have mirror symmetry. Thus we have $[2,0,3][3,0,2] = [6,0,1]$ which is properly equivalent to $[1,0,6]$. If we apply the preceding proposition with $Q_1 = [2,0,3]$ and $Q_2 = [3,0,2]$ then we have $a_1 = 2$, $a_2 = 3$, $b = 0$, and $c = 1$, so the formulas for X and Y are $X = x_1x_2 - y_1y_2$ and $Y = 2x_1y_2 + 3x_2y_1$. The calculations in the proof then give:

$$(2x_1^2 + 3y_1^2)(3x_2^2 + 2y_2^2) = 6X^2 + Y^2 = 6(x_1x_2 - y_1y_2)^2 + (2x_1y_2 + 3x_2y_1)^2$$

To express this in terms of the original two forms $[1,0,6]$ and $[2,0,3]$ we change variables by switching x_2 and y_2 and then we interchange the two terms on the right to get:

$$(2x_1^2 + 3y_1^2)(2x_2^2 + 3y_2^2) = (2x_1x_2 + 3y_1y_2)^2 + 6(x_1y_2 - x_2y_1)^2$$

This shows explicitly that the product of two numbers $2x^2+3y^2$ is a number x^2+6y^2. In a similar way we can obtain formulas for the other products:

$$(x_1^2 + 6y_1^2)(x_2^2 + 6y_2^2) = (x_1x_2 - 6y_1y_2)^2 + 6(x_1y_2 + x_2y_1)^2$$
$$(x_1^2 + 6y_1^2)(2x_2^2 + 3y_2^2) = 2(x_1x_2 - 3y_1y_2)^2 + 3(x_1y_2 + 2x_2y_1)^2$$

Other discriminants can be handled in the same way although the calculations can become complicated. One would start with a list of forms, one for each proper equivalence class of forms of the given discriminant. For each pair of forms on the list one would find a properly equivalent pair of concordant forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$, with suitable changes of variables to convert the given pair of forms to the concordant pair. Then one would apply the formulas for X and Y in the proof of the preceding proposition, and finally one would do another change of variables to convert $a_1a_2X^2 + bXY + cY^2$ to a form on the original list.

Exercises

1. In discriminant $\Delta = -56$ we have the forms $Q_2 = [2,0,7]$ and $Q_3 = [3,2,5]$. Compute Q_2Q_3 and Q_3^2 by finding suitable pairs of concordant forms.

2. Find all the concordant pairs of forms $[3, b, c_1]$ and $[5, b, c_2]$ of discriminant -120.

7.2 The Class Group for Forms

In the previous section we defined a method for multiplying any two elements of the set $CG(\Delta)$ of proper equivalence classes of primitive forms of discriminant Δ, which was to choose a pair of concordant forms $Q_1 = a_1x^2 + bxy + a_2cy^2$ and $Q_2 = a_2x^2 + bxy + a_1cy^2$ in the two proper equivalence classes, and then the product of these two classes is the class containing the form $Q_1Q_2 = a_1a_2x^2 + bxy + cy^2$. Note that the form Q_1Q_2 is the same as Q_2Q_1 since $a_1a_2 = a_2a_1$ so this multiplication operation in $CG(\Delta)$ is commutative.

The multiplication operation in $CG(\Delta)$ has a few other simple properties. A form $[a,b,c]$ is concordant to $[1,b,ac]$ and $[a,b,c][1,b,ac] = [a,b,c]$. Since $[1,b,ac]$ represents 1 it is equivalent to the principal form, hence properly equivalent to it since the principal form has mirror symmetry. Thus the class of the principal form in $CG(\Delta)$ is an identity element for the multiplication.

Each form $[a,b,c]$ is concordant to its mirror image form $[c,b,a]$, and their product is $[ac,b,1]$ which represents 1 hence is properly equivalent the principal form. Thus all elements of $CG(\Delta)$ have inverses for the multiplication operation, obtained by taking mirror image forms.

Forms whose topographs have mirror symmetry give elements of $CG(\Delta)$ that are equal to their inverses. The converse is also true since if a topograph is properly equivalent to its mirror image, this says it has an orientation-reversing symmetry and all such symmetries are mirror reflections by Proposition 5.8.

Another basic property of the multiplication operation in $CG(\Delta)$ is that it is associative, although proving this takes a little more work. To do this we start with three forms Q_1, Q_2, Q_3 giving three classes in $CG(\Delta)$. Choose a number a_1 in the topograph of Q_1, then a number a_2 in the topograph of Q_2 coprime to a_1, then a number a_3 in the topograph of Q_3 coprime to a_1a_2. Each Q_i is then properly equivalent to a form $[a_i, b_i, c_i]$. Since each a_i is coprime to the other two, the Chinese Remainder Theorem guarantees that there is a number b congruent to b_i mod a_i for each i. We would like these congruences to be mod $2a_i$ instead of just mod a_i. To arrange this we go back and first choose a_1 coprime to 2, then a_2 coprime to $2a_1$, then a_3 coprime to $2a_1a_2$, so each a_i is odd. Next, when we apply the Chinese Remainder Theorem we find b congruent to each b_i mod a_i and also congruent to Δ mod 2, hence also congruent to each b_i mod 2. Then b will be congruent to each b_i mod $2a_i$ since 2 and a_i are coprime.

Having chosen b in this way, each form $[a_i, b_i, c_i]$ is properly equivalent to a form $[a_i, b, c_i']$. Equating discriminants of the first two of these new forms, we see that $a_1c_1' = a_2c_2'$ so a_2 divides a_1c_1' and hence it divides c_1' since a_1 and a_2 are coprime. Similarly a_3 divides c_1'. Since a_2 and a_3 are coprime this means that a_2a_3

divides c_1' and we can write $c_1' = a_2a_3c$ for some integer c. Equating discriminants then gives $c_2' = a_1a_3c$ and $c_3' = a_1a_2c$. Thus we have the three forms $[a_1, b, a_2a_3c]$, $[a_2, b, a_1a_3c]$, and $[a_3, b, a_1a_2c]$, and each pair of these forms is concordant. If we multiply the first two forms we get $[a_1a_2, b, a_3c]$, and then multiplying this by the third form $[a_3, b, a_1a_2c]$ gives $[a_1a_2a_3, b, c]$. We get the same result if we first multiply the second and third forms and then multiply their product by the first form. This proves associativity.

We have now shown the following basic fact:

Proposition 7.6. *$CG(\Delta)$ is a group, that is, the multiplication is associative, there is an identity element whose product with any element is that element, and each element has an inverse, so that the product of an element and its inverse is the identity element.*

For general groups the multiplication operation is not required to be commutative, and this complicates the definition slightly. The identity element is required to act as an identity when it is multiplied on both the right and the left. Thus there must be an element e such that both $ge = g$ and $eg = g$ for all elements g in the group. Similar, inverses are required to be inverses for both multiplication on the right and on the left, so each element g must have an inverse element g^{-1} satisfying both $gg^{-1} = e$ and $g^{-1}g = e$. Noncommutative groups often arise quite naturally, and we have in fact already made extensive use of one, the group of linear fractional transformations $LF(\mathbb{Z})$. This differs from $CG(\Delta)$ not just in being noncommutative, but also by having an infinite number of elements, while the number of elements of $CG(\Delta)$ is the class number h_Δ which is always finite.

We should observe that the identity element in a group is always unique since if two elements g and h both act as the identity then $gh = h$ since g is an identity, but we also have $gh = g$ since h is an identity, so $g = h$. Another general fact is that each element g in a group has a unique inverse since if h and h' are two possibly different inverses for g, so both gh and gh' are the identity, then we have $gh = gh'$ so after multiplying both sides of this equation on the left by any inverse g^{-1} we get $h = h'$.

We can now re-examine some of the examples in Section 6.1 to verify that the conjectured group structures on $CG(\Delta)$ are in fact correct.

First consider the case $\Delta = 40$. Here there were two equivalence classes of forms, given by $Q_1 = x^2 - 10y^2$ and $Q_2 = 2x^2 - 5y^2$. Both topographs have mirror symmetry so proper equivalence is the same as equivalence. Thus the group $CG(\Delta)$ has two elements, and we will use the same symbols Q_1 and Q_2 for these elements of $CG(\Delta)$. The identity element of $CG(\Delta)$ is Q_1 since this is the principal form. Since $Q_2 = Q_2^{-1}$ by the mirror symmetry of its topograph, we have $Q_2Q_2 = Q_1$, the identity element of $CG(\Delta)$. This determines the group structure in $CG(\Delta)$ completely, and it agrees with what we predicted from the topographs in Section 6.1.

Next consider the case $\Delta = -84$ where there were four equivalence classes of forms Q_1, Q_2, Q_3, and Q_4. All four topographs have mirror symmetry so $CG(\Delta)$ has four elements. The principal form Q_1 gives the identity element, and $Q_iQ_i = Q_1$ for each i by the mirror symmetry. It remains to determine the products Q_2Q_3, Q_2Q_4, and Q_3Q_4. For Q_2Q_3, this cannot be Q_1 otherwise Q_3 would be Q_2^{-1}. Also Q_2Q_3 cannot be Q_2 otherwise Q_3 would be the identity element Q_1. Similarly, Q_2Q_3 cannot be Q_3. Therefore we must have $Q_2Q_3 = Q_4$. The same reasoning shows that $Q_2Q_4 = Q_3$ and $Q_3Q_4 = Q_2$.

In more complicated cases it can be helpful to use the fact that if two primitive forms Q_1 and Q_2 of the discriminant Δ represent coprime numbers a_1 and a_2 then their product Q_1Q_2 represents a_1a_2. This is a consequence of results in the previous section, particularly Lemma 7.3. For example in the preceding case $\Delta = -84$ we could also show that $Q_2Q_3 = Q_4$ by looking at the topographs to see that Q_2 represents 3 and Q_3 represents 2 so Q_2Q_3 must represent 6. The only element of $CG(\Delta)$ whose topograph contains 6 is Q_4, so $Q_2Q_3 = Q_4$. Similarly one sees that $Q_2Q_4 = Q_3$ using the numbers 3 and 5, and $Q_3Q_4 = Q_2$ using 2 and 5. We could also deduce the last two formulas from $Q_2Q_3 = Q_4$ by multiplying both sides by Q_2 or Q_3.

The next example from Section 6.1 is $\Delta = -56$ where there were three equivalence classes of forms Q_1, Q_2, and Q_3. For Q_1 and Q_2 the topographs have mirror symmetry but not for Q_3 so there is another form Q_4 whose topograph is the mirror image of the one for Q_3, with $Q_4 = Q_3^{-1}$ in $CG(\Delta)$. Again we have Q_1 the identity in $CG(\Delta)$ and we have $Q_2Q_2 = Q_1$ by mirror symmetry. However it is not so easy to determine Q_3Q_3. The topograph of Q_3 contains 3 and 5 so the topograph of Q_3Q_3 must contain 15, but 15 is in the topographs of both Q_1 and Q_2 so this is inconclusive. The same thing happens for other pairs of primes in the topograph of Q_3 such as $3, 13$ or $5, 19$. However, since the topograph of Q_3 does not have mirror symmetry, we know that Q_3 is not Q_3^{-1} hence Q_3Q_3 is not Q_1 so it must be Q_2. Thus all four elements of $CG(\Delta)$ are powers of Q_3, namely Q_3, $Q_3^2 = Q_2$, $Q_3^4 = Q_2^2 = Q_1$, and $Q_3^3 = Q_4$ since $Q_3^4 = Q_1$ implies $Q_3^3 = Q_3^{-1}$ which is Q_4. This determines the structure of $CG(\Delta)$ completely. For example $Q_2Q_4 = Q_3^2Q_3^3 = Q_3^5 = Q_3$ since $Q_3^4 = Q_1$.

In the preceding examples the group $CG(\Delta)$ was small enough that its structure could be determined just from the topographs. This is not always the case in more complicated examples, however. One difficulty is that a form Q and its inverse Q^{-1} have mirror image topographs containing exactly the same numbers, so from the topographs one may be able to compute a product $Q_iQ_j = Q_k^{\pm 1}$ but one cannot always tell which exponent $+1$ or -1 is correct. Another problem is that some numbers can appear in more than one topograph.

We illustrate these difficulties with an example, discriminant $\Delta = -104$ where we showed the topographs of the four equivalence classes of forms in the previous section. Since the first two forms Q_1 and Q_2 have mirror symmetry while the second two Q_3 and Q_4 do not, the group $CG(\Delta)$ has six elements, with the principal form Q_1 the identity and $Q_2^2 = Q_1$. From the product $3 \cdot 17 = 51$ we see that Q_3^2 is Q_1, Q_3, or Q_3^{-1}, but Q_1 is ruled out since the topograph of Q_3 does not have mirror symmetry, and Q_3 is ruled out since $Q_3^2 = Q_3$ would imply $Q_3 = Q_1$. Thus $Q_3^2 = Q_3^{-1}$, or equivalently, $Q_3^3 = Q_1$. Similarly, we can try to compute Q_4^2 from the product $5 \cdot 7 = 35$ which appears in the topographs of Q_1 and Q_3. The possibility that Q_4^2 is Q_1 is ruled out since Q_4 does not have mirror symmetry. Thus $Q_4^2 = Q_3^{\pm 1}$, but we cannot tell which exponent is correct from the topographs and the argument we used to compute Q_3^2 does not work here. In fact we computed Q_4^2 in the previous section by finding a pair of concordant forms properly equivalent to Q_4, and it turned out that Q_4^2 was Q_3^{-1}, the mirror image of Q_3.

Let us see what the higher powers of Q_4 are. Note first that $Q_4^6 = (Q_4^2)^3 = (Q_3^{-1})^3 = Q_1$ since $(Q_3^{-1})^3$ is the inverse of $Q_3^3 = Q_1$. From $Q_4^6 = Q_1$ we obtain $Q_4^5 = Q_4^{-1}$ and $Q_4^4 = Q_4^{-2} = Q_3$. For Q_4^3 we have $(Q_4^3)^2 = Q_4^6 = Q_1$ so Q_4^3 has mirror symmetry making it either Q_1 or Q_2, but $Q_4^3 = Q_1$ is impossible since it would say that Q_4^2 is Q_4^{-1} rather than Q_3^{-1}. Thus $Q_4^3 = Q_2$ and so the six elements of $CG(\Delta)$ are the powers Q_4^i for $i = 1, 2, 3, 4, 5, 6$ with Q_4^6 the identity. This determines the multiplication in $CG(\Delta)$ completely. We will see in Section 7.3 that a group with six elements and commutative multiplication always contains an element whose first through sixth powers are all the elements of the group.

The Representation Theorem

Now we come to our first application of the class group, which is to the problem of determining which primitive forms of a given discriminant Δ represent a given number n. It will suffice to consider only the case that n is positive. This is no restriction when $\Delta < 0$ since there is no need to consider elliptic forms with negative values. When $\Delta > 0$, if we know which forms represent positive n then the negatives of these forms will be the forms representing $-n$. The only forms representing 1 are the forms equivalent to the principal form so we can assume $n > 1$.

Here is the theorem, where for convenience we continue to use the same symbol for a primitive form and for the element of $CG(\Delta)$ that it determines:

Theorem 7.7. (a) *Let a number $n > 1$ be factored as $n = p_1^{e_1} \cdots p_k^{e_k}$ for distinct primes p_i with $e_i > 0$ for each i. Then the primitive forms of discriminant Δ that represent n are the products $Q_1 \cdots Q_k$ where Q_i is a primitive form representing p^{e_i}.*

(b) *The forms of discriminant Δ representing a power p^e of a prime p not dividing Δ are primitive and are exactly the forms $Q^{\pm e}$ where Q is a form representing p.*

If p divides Δ but not the conductor then the only power of p represented in discriminant Δ is p itself, and it is represented by a primitive form.

The theorem says nothing about the primitive forms that represent powers of a prime dividing the conductor, and indeed this is a delicate question as the examples in the large table in Section 6.2 show. In the first statement in (b) the form Q is unique up to equivalence by Proposition 6.15. It may or may not have mirror symmetry, so Q and Q^{-1} may be different elements of $CG(\Delta)$ and the same is true of Q^e and Q^{-e}. In the second statement of (b) a form Q representing p is unique up to equivalence and is symmetric by Proposition 6.17 so there is no need to consider Q^{-1}.

If Δ is a fundamental discriminant then the conductor is 1 so the theorem gives a full reduction of the representation problem for nonprimes to the corresponding problem for primes: The forms representing $p_1^{e_1}\cdots p_k^{e_k}$ are the products $Q_1^{\pm e_1}\cdots Q_k^{\pm e_k}$ where Q_i represents p_i and $e_i = 1$ if p_i divides Δ. For nonfundamental discriminants one obtains all primitive forms representing $p_1^{e_1}\cdots p_k^{e_k}$ by modifying the previous statement to allow some of the primes p_i to divide the conductor, replacing the corresponding terms $Q_i^{\pm e_i}$ by any primitive forms Q_i that represent $p_i^{e_i}$.

As a special case, the only forms representing a power p^e of a prime p not dividing the discriminant are Q^e and Q^{-e} where Q represents p. Since Q^{-e} is the inverse of Q^e in $CG(\Delta)$, these two forms are equivalent so there is only one equivalence class of forms representing p^e. When p is odd this was proved in Proposition 6.15, and now we see that it holds also for $p = 2$.

When there are two or more distinct prime factors p_i the choices between Q^{e_i} and Q^{-e_i} can lead to nonequivalent forms representing the same number. For example for a product p_1p_2 of two different primes there can be four different proper equivalence classes $Q_1^{\pm 1}Q_2^{\pm 1}$ for the four choices of signs, and these can give two different equivalence classes, even if $Q_1 = Q_2$.

Proof of Theorem 7.7: If n is represented by a form Q then Q is properly equivalent to a form $[n,b,c]$. If n factors as $n = a_1a_2\cdots a_k$ then $[n,b,c]$ factors as $[n,b,c] = [a_1,b,nc/a_1][n/a_1,b,a_1c]$ with the latter two forms being concordant. If $k = 2$ this gives $[a_1a_2,b,c] = [a_1,b,a_2c][a_2,b,a_1c]$. If $k > 2$ we can factor $[n/a_1,b,a_1c]$ further as $[a_2,b,nc/a_2][n/a_1a_2,b,a_1a_2c]$. Continuing in this way, we eventually get:

$$[n,b,c] = [a_1,b,nc/a_1][a_2,b,nc/a_2]\cdots[a_k,b,nc/a_k]$$

Here any two forms in the product on the right are concordant.

In particular for the prime factorization $n = p_1^{e_1}\cdots p_k^{e_k}$ we have $[n,b,c] = Q_1\cdots Q_k$ for $Q_i = [p_i^{e_i},b,nc/p_i^{e_i}]$, a form representing $p_i^{e_i}$. By Lemma 7.1 the form $[n,b,c]$ is primitive if and only if each Q_i is primitive since the primes p_i are assumed to be distinct. This proves half of statement (a), that each primitive form representing n can be expressed as a product $Q_1\cdots Q_k$ with Q_i a primitive form representing $p_i^{e_i}$. The other half is the statement that a product $Q_1 \cdots Q_k$ is primitive and represents n if each Q_i is primitive and represents $p_i^{e_i}$. This follows by applying Lemma 7.3 repeatedly, first to forms $[p_1^{e_1}, b_1, c_1]$ and $[p_2^{e_2}, b_2, c_2]$ properly equivalent to Q_1 and Q_2, then to the product of the two resulting forms and a form $[p_3^{e_3}, b_3, c_3]$ properly equivalent to Q_3, and so on.

For part (b) of the theorem, a form representing p^e is properly equivalent to a form $[p^e, b, c]$. As above, this factors as $[p^e, b, c] = [p, b, p^{e-1}c]^e$. If p does not divide the conductor then the forms $Q = [p, b, p^{e-1}c]$ representing p and $Q^e = [p^e, b, c]$ representing p^e are primitive by Proposition 6.14. Since forms representing primes are unique up to equivalence, any form representing p must be properly equivalent to Q or Q^{-1}. Hence the form we started with that represents p^e is properly equivalent to the e^{th} power of Q or Q^{-1}, that is, to Q^e or Q^{-e}.

If p divides Δ but not the conductor then Proposition 6.7 says that p is represented by a form of discriminant Δ but no higher power of p is represented. The form representing p is primitive by Proposition 6.14. □

Let us look at a few examples. For $\Delta = -56$, a fundamental discriminant, we have already determined the group structure of $CG(\Delta)$ which has four elements, but we can use the preceding Theorem 7.7 to quickly rederive the group structure from the topographs which were shown in Section 6.1. For this it suffices to look just at how the powers of 3 are represented. Since 3 is represented by $Q_3 = [3, 2, 5]$ it follows that 3^i is represented by $Q_3^{\pm i}$. The topographs show that 3^2 is represented by $Q_2 = [2, 0, 7]$ so $Q_3^2 = Q_2^{\pm 1}$, but $Q_2 = Q_2^{-1}$ since the topograph of Q_2 has mirror symmetry, so we have $Q_3^2 = Q_2$. Next, 3^3 is represented by Q_3 so $Q_3^3 = Q_3^{\pm 1}$, but $Q_3^3 = Q_3$ would imply $Q_3^2 = Q_1$ contradicting the fact that $Q_3^2 = Q_2$, so $Q_3^3 = Q_3^{-1}$. And finally 3^4 is represented by $Q_1 = [1, 0, 14]$ so $Q_3^4 = Q_1^{\pm 1} = Q_1$. Thus we see again that $CG(\Delta)$ consists of the powers of Q_3, with Q_3^4 the identity.

From this we can determine which forms represent a number $n = p_1^{e_1} \cdots p_k^{e_k}$, with $e_i \le 1$ for $p_i = 2, 7$. Changing notation for convenience, let Q be the form $[3, 2, 5]$ previously called Q_3, so the other three forms are powers of Q. According to the theorem, the forms representing n are the products $(Q^{q_1})^{\pm e_1} \cdots (Q^{q_k})^{\pm e_k}$ where Q^{q_i} is the power of Q representing p_i. We may assume each q_i is 0, 1, or 2 since $Q^3 = Q^{-1}$ represents the same numbers as Q. The product $(Q^{q_1})^{\pm e_1} \cdots (Q^{q_k})^{\pm e_k}$ is then a power Q^e where only the value of e mod 4 matters. Primes p_i represented by Q^4, the identity in $CG(\Delta)$, can be ignored. Then we have

$$e = \sum_{Q} \pm e_i + \sum_{Q^2} \pm 2e_i$$

where the first sum is over subscripts i such that p_i is represented by Q and similarly for the second sum with Q^2 in place of Q. The sign $\pm$ in the second sum can be ignored since $Q^2 = Q^{-2}$. As we saw in Section 6.3, the forms Q^0 and Q^2 make up one genus while Q and the equivalent form $Q^3 = Q^{-1}$ make up the other genus. The parity of e thus determines the genus of the forms representing n. (Recall that forms representing a given number all belong to the same genus.) From the formula for e we can deduce that n is represented by both Q^0 and Q^2 exactly when e is even and at least one e_i in the first sum is odd since this is the only time when the choice of the signs $\pm$ matters.

As another example, when $\Delta = -104$ we computed $CG(\Delta)$ to have six elements, the first through sixth powers of the form $Q_4 = [5,4,6]$ with $Q_4^6 = Q_1$, the identity in $CG(\Delta)$. We can obtain most of this structure a little more efficiently now using Theorem 7.7. Looking at the topographs, we see that 5, 5^2, and 5^3 are represented by Q_4, Q_3, and Q_2 so $Q_4^2 = Q_3^{\pm 1}$ and $Q_4^3 = Q_2^{\pm 1}$ which is Q_2 since the topograph of Q_2 has mirror symmetry. Since $Q_2^2 = Q_1$ it follows that $Q_4^6 = Q_2^2 = Q_1$ so $Q_4^5 = Q_4^{-1}$ and $Q_4^4 = Q_4^{-2} = Q_3^{\mp 1}$. We cannot determine which sign in $Q_4^2 = Q_3^{\pm 1}$ is correct just from the topographs, but we showed that $Q_4^2 = Q_3^{-1}$ earlier.

The forms representing a number $n = p_1^{e_1} \cdots p_k^{e_k}$ when $\Delta = -104$ can be described in a similar way to the preceding example with $\Delta = -56$. For $\Delta = -104$ the exceptional primes p_i with $e_i \leq 1$ are 2 and 13. The forms representing n are the products $(Q^{q_1})^{\pm e_1} \cdots (Q^{q_k})^{\pm e_k}$ where Q^{q_i} is the power of $Q = Q_4$ representing p_i with q_i either 0, 1, 2, or 3. Writing this product as Q^e where only the value of e mod 6 matters, the formula for e now has another term:

$$e = \sum_{Q} \pm e_i + \sum_{Q^2} \pm 2e_i + \sum_{Q^3} \pm 3e_i$$

The parity of e again determines the genus, with one genus consisting of Q^0 and Q^2 (which is equivalent to Q^4) and the other genus consisting of Q and Q^3 (with $Q^5 = Q^{-1}$ equivalent to Q). From the formula for e one could work out when a number is represented by both forms within a genus and when it is represented by only one form. Note that for the formula above it does not matter whether Q_4^2 is Q_3 or Q_3^{-1} since both these forms represent the same numbers.

Exercises

1. For discriminant $\Delta = -47$ show the class number is 5 and determine the multiplication rules for the five proper equivalence classes of forms.

2. Determine the numbers represented by each of the two forms $[1,1,6]$ and $[2,1,3]$.

3. Show that the numbers represented by $x^2 + 4y^2$ are the numbers $2^m p_1 \cdots p_k$ where m is 0, 2, or 3 and each p_i is a prime congruent to 1 mod 4.

4. Show that if two forms Q_1 and Q_2 in the class group $CG(\Delta)$ represent coprime numbers n_1 and n_2 then their product $Q_1 Q_2$ represents $n_1 n_2$. Give an example where this fails without the coprimeness assumption, even if n_1 and n_2 are coprime to Δ.

5. For a fixed discriminant Δ consider the set S_Δ of primes that do not divide the conductor and are represented by primitive forms with mirror symmetry. Show that numbers that are products of primes in S_Δ are represented by at most one form of discriminant Δ, up to equivalence, and this form has mirror symmetry.

7.3 Finite Abelian Groups

A group whose multiplication operation is commutative is usually referred to as an ***abelian*** group, after the mathematician Niels Henrik Abel (1802–1829), although the term "commutative group" is sometimes used as well. The aim of this section is to explain the structure of abelian groups with finitely many elements. This structure is far simpler than for finite nonabelian groups which can be extremely complicated, with no hope of being completely classified.

The number of elements in a group G is called the ***order*** of G. This can be finite or infinite, but for the class group $CG(\Delta)$ it is always finite since it is just the class number for discriminant Δ.

For an element g in a group G the smallest positive integer n such that g^n is the identity is called the ***order*** of g if such an n exists, and otherwise the order of g is said to be infinite. Each element g in a finite group G has finite order since the powers $g, g^2, g^3, \cdots$ cannot all be distinct elements of G, so we must have $g^m = g^n$ for some $m \neq n$, say $m < n$, and then if we multiply both g^m and g^n by g^{-m}, the inverse of g^m, we see that g^{n-m} is the identity. Thus some positive power of g is the identity, and the smallest such power is the order of g. The identity element of a group always has order 1 and is obviously the only element of order 1.

If an element g of a group G has order n then all the powers $g, g^2, g^3, \cdots, g^n$ must be distinct elements of G, otherwise if two of these powers g^i and g^j were equal with $i < j$ we would have g^{j-i} equal to the identity, with $j - i < n$, contrary to the assumption that g has order n. If g has order n then the higher powers $g^{n+1}, g^{n+2}, \cdots$ just cycle through the powers $g, g^2, \cdots g^n$ repeatedly. In particular the only powers of g that are the identity element of G are the powers g^{kn} for integers k. The negative powers of g are just the inverses of the positive powers, and these cycle through the same sequence $g, g^2, \cdots, g^n$ in reverse order since $g^{-1} = g^{n-1}$, $g^{-2} = g^{n-2}$, and so on.

If g has order n then the order of each power g^k can be determined in the following way. The order of g^k is the number m such that mk is the smallest multiple of k that is also a multiple of n. The smallest common multiple of k and n is kn/d where d is the greatest common divisor of k and n, as one can see by comparing the prime factorizations of k and n. Thus $mk = kn/d$ so $m = n/d$ and the order of g^k is n/d.

In particular if g has order $n = kl$, then g^k has order l. This means that for each divisor l of n there is a power of g having order l.

For example if g has order 6 then g^2 and g^4 have order 3, g^3 has order 2, and g^5 has order 6. Similarly, if g has order 12 then g^2 and g^{10} have order 6, g^3 and g^9 have order 4, g^4 and g^8 have order 3, and g^5, g^7, and g^{11} have order 12.

A finite group G is called ***cyclic*** if there is an element $g \in G$ such that every element of G is a power of g, so the elements $g, g^2, g^3, \cdots$ cycle through all the elements of G. The element g is then called a ***generator*** of G. Cyclic groups are automatically abelian since $g^k g^l$ and $g^l g^k$ both equal g^{k+l}. If a generator g of a cyclic group G has order n, then this is also the order of G since all the powers $g, g^2, g^3, \cdots, g^n$ must be distinct, as noted earlier. Thus a group of order n is cyclic exactly when it contains an element of order n. In a cyclic group there are generally a number of different choices for a generator since if g is one generator of order n then g^k is a generator exactly when it has order n, which is equivalent to k being coprime to n. The number of different generators is thus $\varphi(n)$ where φ is the Euler phi function.

Among the groups $CG(\Delta)$ that we computed in the previous section, $CG(\Delta)$ is cyclic of order 4 for $\Delta = -56$ and cyclic of order 6 for $\Delta = -104$, but for $\Delta = -84$ the group is not cyclic since it has order 4 but each element other than the identity has order 2.

Cyclic groups are easy to understand, and our next goal is to see that all finite abelian groups are built from cyclic groups by a fairly simple procedure. Given two groups G_1 and G_2, the ***product group*** $G_1 \times G_2$ is defined to be the set of all pairs (g_1, g_2) with $g_1 \in G_1$ and $g_2 \in G_2$. The multiplication operation in $G_1 \times G_2$ is defined by $(g_1, g_2) \cdot (g_1', g_2') = (g_1 g_1', g_2 g_2')$, so the coordinates are multiplied separately. The identity element of $G_1 \times G_2$ is the pair (g_1, g_2) with g_1 the identity in G_1 and g_2 the identity in G_2. The inverse of an element (g_1, g_2) is (g_1^{-1}, g_2^{-1}). More generally one can define products $G_1 \times \cdots \times G_k$ of any collection of groups $G_1, \cdots, G_k$, with the elements of this product group being k-tuples $(g_1, \cdots, g_k)$ with $g_i \in G_i$ for each i. One can also iterate the process of forming products of groups but this gives nothing new since for example $(G_1 \times G_2) \times G_3$ is really the same as $G_1 \times G_2 \times G_3$ by rewriting its elements $((g_1, g_2), g_3)$ as (g_1, g_2, g_3).

If G_1 and G_2 are finite groups of orders n_1 and n_2, then $G_1 \times G_2$ has order $n_1 n_2$ since the two coordinates g_1 and g_2 of pairs (g_1, g_2) in $G_1 \times G_2$ vary independently over G_1 and G_2. For an element (g_1, g_2) in $G_1 \times G_2$, if g_1 has order n_1 and g_2 has order n_2 then the order of (g_1, g_2) is the least common multiple of n_1 and n_2 since a power $(g_1, g_2)^n = (g_1^n, g_2^n)$ is the identity exactly when n is a multiple of both n_1 and n_2, so the order of (g_1, g_2) is the smallest such multiple. In particular, if n_1 and n_2 are coprime then (g_1, g_2) has order $n_1 n_2$. This leads to the following interesting fact:

Proposition 7.8. *If G_1 and G_2 are cyclic of coprime orders n_1 and n_2 then $G_1 \times G_2$ is cyclic of order $n_1 n_2$.*

Proof: If g_1 is a generator of G_1 of order n_1 and g_2 is a generator of G_2 of order n_2 then (g_1, g_2) has order $n_1 n_2$ if n_1 and n_1 are coprime, as we saw above. The group $G_1 \times G_2$ is therefore cyclic since it contains an element whose order equals the order of the group. $\square$

Now we come to the main result in this section, the basic structure theorem for finite abelian groups:

Theorem 7.9. *Every finite abelian group is a product $G_1 \times \cdots \times G_k$ of cyclic groups $G_1, \cdots, G_k$, with the possibility $k = 1$ allowed when the group itself is cyclic.*

For the proof we will use the notation $o(g)$ for the order of an element $g \in G$. The identity element of G will be written simply as 1. We need two preliminary lemmas.

Lemma 7.10. *If two elements g_1 and g_2 of a finite abelian group have coprime orders $o(g_1)$ and $o(g_2)$ then their product $g_1 g_2$ has order $o(g_1)o(g_2)$.*

This need not be true if $o(g_1)$ and $o(g_2)$ are not coprime. As an extreme example take g_2 to be g_1^{-1}. Another example would be to take g_1 to be an element of maximal order in G and g_2 any element with $o(g_2) > 1$.

Proof: Let $n_1 = o(g_1)$ and $n_2 = o(g_2)$. Then $(g_1 g_2)^{n_1 n_2} = g_1^{n_1 n_2} g_2^{n_1 n_2} = 1$ so it will suffice to show that if $(g_1 g_2)^n = 1$ then n is a multiple of $n_1 n_2$.

Suppose $(g_1 g_2)^n = 1$ and let $g = g_1^n = g_2^{-n}$. Then $g^{n_1} = g_1^{n n_1} = (g_1^{n_1})^n = 1$ so $o(g)$ divides n_1. Similarly, $g^{n_2} = g_2^{-n n_2} = (g_2^{n_2})^{-n} = 1$ so $o(g)$ divides n_2. Since n_1 and n_2 are assumed to be coprime, this means $o(g) = 1$ and hence $g = 1$. Thus $g_1^n = 1$ and $g_2^{-n} = 1$, which implies $g_2^n = 1$. Since $g_1^n = 1$ it follows that n is a multiple of n_1, and n is also a multiple of n_2 since $g_2^n = 1$. As n_1 and n_2 are coprime, this implies that n is a multiple of $n_1 n_2$. $\square$

Lemma 7.11. *For a finite abelian group G let m be the maximal order of elements of G. Then the order of each element of G is a divisor of m.*

Proof: Suppose this is false, so there is an element g such that $o(g)$ does not divide the maximal order m. This means there is some prime power p^k dividing $o(g)$ such that the highest power p^l dividing m has $l < k$. Since p^k divides $o(g)$ there is a power of g having order p^k. Let g_1 be this power of g and let g_2 be an element of G of order m/p^l, for example h^{p^l} where h is an element of order m. Then by the preceding lemma the product $g_1 g_2$ has order $p^k(m/p^l)$ which is greater than m since $k > l$. This contradicts the maximality of m, so we conclude that $o(g)$ divides m for all $g \in G$. $\square$

Proof of Theorem 7.9: Let g_1 be an element of G of maximal order n_1. If every element of G is a power of g_1 then G is cyclic and there is nothing more to prove. If there are elements of G that are not powers of g_1 then we proceed by induction to find further elements $g_2, \cdots, g_q$ satisfying the following two properties:

(1_q) The elements $g_1, g_2, \cdots, g_q$ have orders $n_1, n_2, \cdots, n_q$ where $n_i > 1$ for each i and n_i is divisible by n_{i+1} for each $i < q$.

(2_q) If $g_1^{k_1} \cdots g_q^{k_q} = g_1^{k_1'} \cdots g_q^{k_q'}$ then $k_i \equiv k_i' \bmod n_i$ for each i. Since each g_i has order n_i an equivalent statement is that if $g_1^{k_1} \cdots g_q^{k_q} = g_1^{k_1'} \cdots g_q^{k_q'}$ with $0 \le k_i < n_i$ and $0 \le k_i' < n_i$ for each i, then $k_i = k_i'$ for each i.

If we have elements $g_1, \cdots, g_q$ satisfying (1_q) and (2_q) such that their products $g_1^{k_1} \cdots g_q^{k_q}$ give all the elements of G, then by rewriting each product $g_1^{k_1} \cdots g_q^{k_q}$ as a q-tuple $(g_1^{k_1}, \cdots, g_q^{k_q})$ we see that G is a product of cyclic groups of orders $n_1, \cdots, n_q$ and the proof will be complete.

If the products $g_1^{k_1} \cdots g_q^{k_q}$ do not account for all elements of G then we will show how to find another element g_{q+1} of order n_{q+1} so that the conditions (1_{q+1}) and (2_{q+1}) are satisfied. This process can be iterated until all elements of G are exhausted since at each step the number of products $g_1^{k_1} \cdots g_q^{k_q}$ increases, at least doubling in fact, and G has only finitely many elements.

Assume inductively that we have already chosen elements $g_1, \cdots, g_q$ satisfying (1_q) and (2_q). To find g_{q+1} we consider congruence classes of elements of G mod $g_1, \cdots, g_q$, which means that we consider each element g as congruent to all the products $gg_1^{k_1} \cdots g_q^{k_q}$ for arbitrary exponents k_i. Let $[g]_q$ denote the congruence class of g, the set of all the elements $gg_1^{k_1} \cdots g_q^{k_q}$. In particular $[g]_q$ includes g itself by choosing each k_i to be 0. It is not hard to see that these congruence classes $[g]_q$ form an abelian group with the product defined by $[g]_q[g']_q = [gg']_q$. Let this group of congruence classes $[g]_q$ be denoted $[G]_q$. In particular when $q = 0$ we start with $[G]_0 = G$ before we have chosen any of the elements g_i. We then start the induction by choosing g_1 to be an element of $G = [G]_0$ of maximal order n_1. Conditions (1_1) and (2_1) are then obviously satisfied.

For the induction step, if there are elements of G that are not products $g_1^{k_1} \cdots g_q^{k_q}$ then $[G]_q$ has more than one element. Let $[g_{q+1}]_q$ be an element of $[G]_q$ of maximal order n_{q+1} in $[G]_q$. First we check that n_{q+1} divides n_q. Since $[g_{q+1}]_q^{n_{q+1}} = [1]_q$ we have $g_{q+1}^{n_{q+1}} = g_1^{k_1} \cdots g_q^{k_q}$ for some exponents k_i. Then in $[G]_{q-1}$ we have $[g_{q+1}]_{q-1}^{n_q} = [1]_{q-1}$ since Lemma 7.11 implies that all elements of $[G]_{q-1}$ have order dividing the maximal order, which is n_q by the inductive definition of n_q. The equation $[g_{q+1}]_{q-1}^{n_q} = [1]_{q-1}$ means that $g_{q+1}^{n_q}$ is a product of powers of $g_1, \cdots, g_{q-1}$, so it is certainly a product of powers of $g_1, \cdots, g_q$ which means $[g_{q+1}]_q^{n_q} = [1]_q$. Thus n_q is a multiple of n_{q+1}, the order of $[g_{q+1}]_q$ in $[G]_q$, as we wanted to show. Since (1_q) holds by inductive assumption, it follows that n_{q+1} divides each n_i with $i \le q$.

It is also true that n_{q+1} divides each k_i in the formula $g_{q+1}^{n_{q+1}} = g_1^{k_1} \cdots g_q^{k_q}$. To see this, consider the power $g_{q+1}^{n_i}$. We can write this as $g_{q+1}^{n_i} = (g_{q+1}^{n_{q+1}})^{n_i/n_{q+1}} =$

$(g_1^{k_1}\cdots g_q^{k_q})^{n_i/n_{q+1}}$ with n_i/n_{q+1} an integer since n_{q+1} divides n_i. We can also write $g_{q+1}^{n_i}$ as a product $g_1^{l_1}\cdots g_{i-1}^{l_{i-1}}$ since $[g_{q+1}^{n_i}]_{i-1} = [g_{q+1}]_{i-1}^{n_i} = [1]_{i-1}$ as a consequence of the definition of n_i as the maximal order of elements of $[G]_{i-1}$, so all elements of $[G]_{i-1}$ have order dividing n_i by Lemma 7.11. Since the two expressions $(g_1^{k_1}\cdots g_q^{k_q})^{n_i/n_{q+1}}$ and $g_1^{l_1}\cdots g_{i-1}^{l_{i-1}}$ for $g_{q+1}^{n_i}$ are equal with g_i not appearing in the second expression, the property (2_q) implies that the exponent $k_i n_i/n_{q+1}$ on g_i in the first expression must be a multiple of n_i, the order of g_i by (1_i). Thus we have $k_i n_i/n_{q+1} = mn_i$ for some integer m. Canceling n_i from this equation, we get $k_i/n_{q+1} = m$ so n_{q+1} divides k_i.

Next we would like to find an element $g_{q+1}g_1^{x_1}\cdots g_q^{x_q}$ congruent to g_{q+1} mod $g_1,\cdots,g_q$ and having order n_{q+1} in G. The order of $g_{q+1}g_1^{x_1}\cdots g_q^{x_q}$ cannot be less than n_{q+1} since it determines the same element of $[G]_q$ as g_{q+1} and $[g_{q+1}]_q$ has order n_{q+1} in $[G]_q$. This means that we just need to find exponents x_i so that $(g_{q+1}g_1^{x_1}\cdots g_q^{x_q})^{n_{q+1}} = 1$. Since $g_{q+1}^{n_{q+1}} = g_1^{k_1}\cdots g_q^{k_q}$ we have:

$$(g_{q+1}g_1^{x_1}\cdots g_q^{x_q})^{n_{q+1}} = g_{q+1}^{n_{q+1}}g_1^{x_1 n_{q+1}}\cdots g_q^{x_q n_{q+1}} = g_1^{k_1+x_1 n_{q+1}}\cdots g_q^{k_q+x_q n_{q+1}}$$

This will be 1 if $k_i + x_i n_{q+1} = 0$ for each i. Solving $k_i + x_i n_{q+1} = 0$ for x_i gives $x_i = -k_i/n_{q+1}$ with x_i an integer since we have shown that n_{q+1} divides k_i.

Having found an element $g_{q+1}g_1^{x_1}\cdots g_q^{x_q}$ of order n_{q+1}, we replace g_{q+1} by this element, so the new g_{q+1} has order n_{q+1} in G. It remains to check condition (2_{q+1}). If $g_1^{k_1}\cdots g_q^{k_q}g_{q+1}^{k_{q+1}} = g_1^{k_1'}\cdots g_q^{k_q'}g_{q+1}^{k_{q+1}'}$ then in $[G]_q$ we have $[g_{q+1}]_q^{k_{q+1}} = [g_{q+1}]_q^{k_{q+1}'}$. Since the order of $[g_{q+1}]_q$ in $[G]_q$ is n_{q+1} this implies that $k_{q+1} \equiv k_{q+1}' \bmod n_{q+1}$, hence $g_{q+1}^{k_{q+1}} = g_{q+1}^{k_{q+1}'}$ in G since g_{q+1} has order n_{q+1}. We can then cancel $g_{q+1}^{k_{q+1}}$ and $g_{q+1}^{k_{q+1}'}$ from the equation $g_1^{k_1}\cdots g_q^{k_q}g_{q+1}^{k_{q+1}} = g_1^{k_1'}\cdots g_q^{k_q'}g_{q+1}^{k_{q+1}'}$ to get $g_1^{k_1}\cdots g_q^{k_q} = g_1^{k_1'}\cdots g_q^{k_q'}$. Since condition (2_q) holds by induction, we have $k_i \equiv k_i' \bmod n_i$ for each $i \le q$. Thus (2_{q+1}) holds and we are done. □

To illustrate how the preceding proof works, suppose we start with the group $G = H_1 \times H_2$ where H_1 is cyclic of order 4 generated by an element h_1 and H_2 is cyclic of order 2 generated by an element h_2 of order 2. In this case we already know that G is a product of cyclic groups, but suppose we forget this and just follow the proof through. At the first step we choose an element g_1 in G of maximal order, so let us choose $g_1 = (h_1, 1)$ which has order 4 in G. There are then two congruence classes of elements of G mod g_1, namely the class consisting of the elements $(h_1^{k_1}, h_2^{k_2})$ with $k_2 = 0$ and the class with $k_2 = 1$, so the group $[G]_1$ of congruence classes mod g_1 has order 2. Intuitively, taking congruence classes mod g_1 amounts just to ignoring the first coordinates of pairs $(h_1^{k_1}, h_2^{k_2})$ since we are free to change this coordinate arbitrarily by multiplying $(h_1^{k_1}, h_2^{k_2})$ by any element $(h_1^{l_1}, 1)$. Next we choose an element g_2 of maximal order in $[G]_1$. For this we can choose $g_2 = (h_1^{k_1}, h_2)$ for any k_1. If we choose k_1 to be 1 or 3 then g_2 will have order 4, which is larger than the maximal order of elements of $[G]_1$ which is 2. The next-to-last paragraph of the proof gives a procedure for rechoosing g_2 to have order equal to 2 rather than 4, so in the present example this would amount to choosing k_1 to be 0 or 2 rather than 1 or 3. Either choice $k_1 = 0$ or $k_1 = 2$ will work, but if we choose $k_1 = 0$ then the element g_2 becomes simply $(1, h_2)$ and a general product $g_1^{l_1} g_2^{l_2}$ becomes the general element $(h_1^{l_1}, h_2^{l_2})$ of $H_1 \times H_2$.

From the preceding theorem we can deduce a general fact:

Corollary 7.12. *Each element of a finite abelian group has order dividing the order of the group.*

An equivalent statement is that if a finite abelian group G has order n then $g^n = 1$ for each $g \in G$. This is because if $g^n = 1$ then the order of g divides n and conversely.

Proof: By the theorem a finite abelian group G is a product $G_1 \times \cdots \times G_k$ of cyclic groups G_i. If the order of G_i is n_i then the order of G is $n = n_1 \cdots n_k$. Each element g_i in G_i is a power of a generator of G_i which has order n_i so $g_i^{n_i} = 1$ and hence $g_i^n = 1$. For any element $g = (g_1, \cdots, g_k)$ of G we then have $g^n = 1$. □

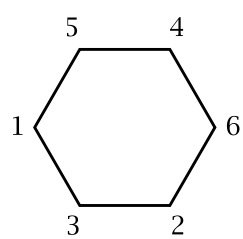
Fermat's Little Theorem, which we encountered in the proof of quadratic reciprocity in Section 6.4, is a special case of this corollary, the case that the group is the group of congruence classes mod p of integers coprime to p, for p an odd prime. The group operation is multiplication of congruence classes, and integers coprime to p have multiplicative inverses mod p so one does indeed have a group. The order of the group is $p - 1$, so each element has order dividing $p - 1$ which implies that $a^{p-1} \equiv 1 \bmod p$ for each integer a coprime to p, as Fermat's Little Theorem asserts.

The proof we gave for Fermat's Little Theorem in Section 6.4 extends easily to give a simple proof of the corollary for any finite abelian group G. To see this, suppose G has order n, with the elements of G being $g_1, \cdots, g_n$. For an arbitrary element g in G the multiples $gg_1, \cdots, gg_n$ are all distinct since if $gg_i = gg_j$ then multiplying both sides of this equation by g^{-1} gives $g_i = g_j$. Thus the sets $\{g_1, \cdots, g_n\}$ and $\{gg_1, \cdots, gg_n\}$ are equal. Taking the product of all the elements in each of these two sets and using commutativity of the multiplication operation, we have $g_1 \cdots g_n = g^n g_1 \cdots g_n$ which implies $g^n = 1$.

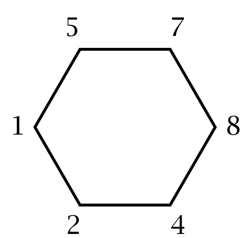
Fermat's Little Theorem was generalized by Euler to replace the prime p by any number n. Here one takes the group of congruence classes mod n of numbers coprime to n. As we know, these numbers have multiplicative inverses mod n so we again have a group. Its order is given by Euler's function $\varphi(n)$, the number of positive integers less than n and coprime to n. The statement is then that $a^{\varphi(n)} \equiv 1 \bmod n$ for every a coprime to n.

There are several different notations commonly used for the group of congruence classes mod n of integers coprime to n. We will write it as $\mathbb{Z}_n^*$ with $\mathbb{Z}_n$ denoting the set of congruence classes of integers mod n and the star indicating that we are only taking congruence classes of integers coprime to n. One might wonder what the structure of $\mathbb{Z}_n^*$ is as a product of cyclic groups. The first step in understanding this is to apply the Chinese Remainder Theorem. As we saw in Section 2.3, if the prime factorization of n is $p_1^{r_1}\cdots p_k^{r_k}$ for distinct primes p_i, then specifying the congruence class mod n of an integer coprime to n is equivalent to specifying its congruence class mod $p_i^{r_i}$ for each i, with the latter classes being coprime to $p_i^{r_i}$ (which is the same as being coprime to p_i). This amounts to saying that $\mathbb{Z}_n^*$ is the product $\mathbb{Z}_{p_1^{r_1}}^*\times\cdots\times\mathbb{Z}_{p_k^{r_k}}^*$.

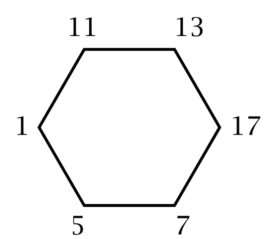
This gives a reduction to the case of a prime power p^r. When p is an odd prime the group $\mathbb{Z}_{p^r}^*$ is cyclic, while $\mathbb{Z}_{2^r}^*$ is cyclic when $r\le 2$ but for larger r it is the product of two cyclic groups, one of order 2^{r-2} and the other of order 2. These facts will not be needed in the rest of the book so we will not prove them but will instead just look at a few examples. Some cases when $\mathbb{Z}_n^*$ is cyclic are shown in the following figures where the elements of $\mathbb{Z}_n^*$ label the vertices of a polygon and multiplication by a generator of $\mathbb{Z}_n^*$ rotates the polygon, taking each vertex to the next vertex.



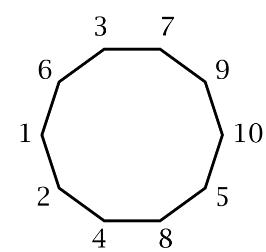
$\mathbb{Z}_7^*$



$\mathbb{Z}_9^*$



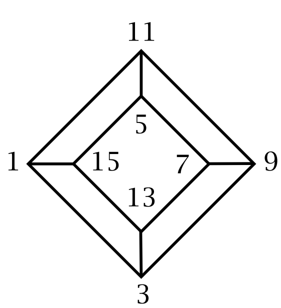
$\mathbb{Z}_{18}^*$



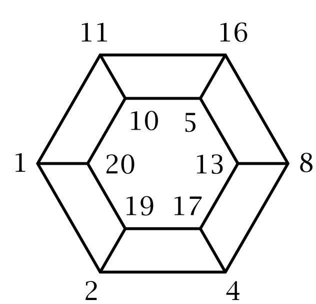
$\mathbb{Z}_{11}^*$

For example in the first figure the group $\mathbb{Z}_7^*$ is cyclic of order 6 generated by 3 with the powers of 3 mod 7 being $3,2,6,4,5,1$. Notice that when $\mathbb{Z}_n^*$ is cyclic, any two opposite vertices are negatives of each other mod n, corresponding to the fact that -1 is the only element of order 2 in $\mathbb{Z}_n^*$ and multiplication by -1 rotates the polygon 180 degrees. Note also that reflecting the polygon across its horizontal axis of symmetry sends each element of $\mathbb{Z}_n^*$ to its multiplicative inverse in $\mathbb{Z}_n^*$.

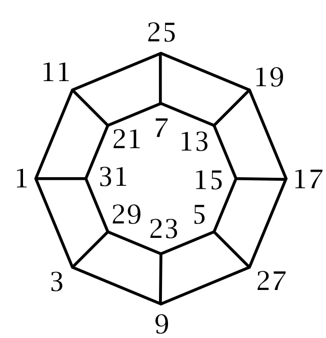
Some cases when $\mathbb{Z}_n^*$ is not cyclic but is the product of a cyclic group of order 2 with a cyclic group are shown in the next three figures.



$\mathbb{Z}_{16}^*$



$\mathbb{Z}_{21}^*$



$\mathbb{Z}_{32}^*$

Here the cyclic factor of order 2 is generated by -1 and multiplication by -1 takes each vertex of the inner polygon to the adjacent vertex of the outer polygon and vice versa. Multiplication by a generator of the other cyclic factor rotates the whole figure. Multiplicative inverses are again given by reflection across the horizontal axis.

Each of these diagrams is known as a *Cayley graph* for the group. The graph has a vertex for each element of the group, and two vertices are joined by an edge whenever one group element is obtained from another by multiplication by one of a chosen set of generators for the group. In the first four examples the group was cyclic so it had a single generator, while in the last three examples the group had two generators, one for each cyclic factor.

The preceding Corollary 7.12 implies that a finite abelian group G of prime order p must be cyclic since any nonidentity element of G must have order p. This holds more generally if the order of G is a product of distinct primes since in a factorization of G as a product of cyclic groups these groups must all have coprime orders so their product will also be cyclic by repeated applications of Proposition 7.8.

By Proposition 7.8, every cyclic group whose order is not a power of a prime can be expressed as a product of two cyclic groups of smaller order. Applying this fact repeatedly, every cyclic group is a product of cyclic groups of prime power order. Hence by Theorem 7.9 every finite abelian group is a product of cyclic groups of prime power order. A cyclic group of prime power order p^k cannot be factored as a product since the factors would have orders p^l for $l < k$ so the elements of the factors would have orders dividing p^{k-1}, hence the same would be true for all elements of the product, contradicting the fact that it is cyclic of order p^k and so contains an element of order p^k.

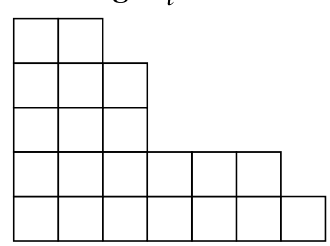
Proposition 7.13. *The factorization of a finite abelian group as a product of cyclic groups of prime power order is unique in the sense that any two such factorizations have the same number of factors of each order.*

For example, if we let C_n denote a cyclic group of order n, then the only two abelian groups of order 4 are C_4 and $C_2 \times C_2$. For order 8 the three possibilities are C_8, $C_4 \times C_2$, and $C_2 \times C_2 \times C_2$. For order 16 there are five possibilities: C_{16}, $C_8 \times C_2$, $C_4 \times C_4$, $C_4 \times C_2 \times C_2$, and $C_2 \times C_2 \times C_2 \times C_2$. These examples illustrate the general fact that the abelian groups of order a prime power p^k correspond exactly to the different partitions of k as a sum of numbers from 1 to k. In the case of $2^4 = 16$ these were the five partitions 4, $3+1$, $2+2$, $2+1+1$, and $1+1+1+1$. (The order of the terms does not matter, so $2+1+1$ is regarded as the same partition as $1+2+1$ and $1+1+2$.)

For groups whose order is a product of powers of different primes one just combines the various groups of each prime power independently. Thus for order $144 = 9 \cdot 16$ there are ten possibilities, the products of the five groups of order 16 listed above with either of the two groups C_9 and $C_3 \times C_3$ of order 9. The only time there is only one group of order n is when n is a product of distinct primes, so the group is a product of cyclic groups of distinct prime orders, making the whole group cyclic.

Proof of Proposition 7.13: The idea will be to characterize the number of cyclic factors of each prime power order in an intrinsic way that does not depend on a particular choice of factorization. For a prime p dividing the order of a finite abelian group G let $G(p)$ be the set of elements in G whose order is a power of p, including the identity element 1 of order p^0. Note that an element g has order a power of p exactly when $g^{p^n} = 1$ for some n. Given a factorization of G as a product $G_1 \times \cdots \times G_k$ of cyclic groups of prime power order, an element $g = (g_1, \cdots, g_k)$ of G has order a power of p exactly when each coordinate g_i has order a power of p since if $g^{p^n} = 1$ then $g_i^{p^n} = 1$ for each i and conversely if $g_i^{p^{n_i}} = 1$ for each i then $g^{p^n} = 1$ for n the largest n_i. For the factors G_i whose order is a power of a prime different from p the only way to have $g_i^{p^n} = 1$ is when $g_i = 1$. We can therefore regard $G(p)$ as the product of the factors G_i whose order is a power of p. This gives a characterization of the product of the factors G_i of order a power of p that does not depend on the choice of the factorization of G.

Thus the problem reduces to the case that $G = G(p)$, i.e., G has order p^n for some n, so we assume this from now on. It remains to give an intrinsic characterization of the number of cyclic factors of order p^r for each r. Suppose we are given a factorization of G as a product $G_1 \times \cdots \times G_k$ of cyclic groups of order $p^{n_1}, \cdots, p^{n_k}$ with each $n_i \geq 1$. We can assume the exponent sequence $n_1, \cdots, n_k$ is in decreasing order, so $n_1 \geq n_2 \geq \cdots \geq n_k$. Such a sequence can be pictured as an arrangement of boxes into rows and columns, with the i^{th} column containing n_i boxes. The figure at the right shows the box diagram for the sequence $5, 5, 4, 2, 2, 2, 1$. If the order of G is p^n then the total number of boxes is n since the cyclic factor G_i of order p^{n_i} corresponds to the i^{th} column with n_i boxes, so the order of G is $p^n = p^{n_1} \cdots p^{n_k} = p^{n_1 + \cdots + n_k}$, hence $n = n_1 + \cdots + n_k$.



Consider the subset $G^{(p)}$ of G consisting of all elements that are p^{th} powers g^p of elements g in G. The p^{th} power of an element $g = (g_1, \cdots, g_k)$ of $G_1 \times \cdots \times G_k$ is $g^p = (g_1^p, \cdots, g_k^p)$ so an element of G is a p^{th} power exactly when each of its coordinates is a p^{th} power. Thus $G^{(p)} = G_1^{(p)} \times \cdots \times G_k^{(p)}$ where $G_i^{(p)}$ consists of the powers g_i^p of elements g_i in G_i. If g_i is a generator of G_i then the elements of $G_i^{(p)}$ are $g_i^p, g_i^{2p}, g_i^{3p}, \cdots, g^{(p^{n_i - 1})p} = 1$ so $G_i^{(p)}$ is a cyclic group of order $p^{n_i - 1}$. Thus the box diagram for $G^{(p)} = G_1^{(p)} \times \cdots \times G_k^{(p)}$ is obtained from the box diagram for $G = G_1 \times \cdots \times G_k$ by deleting the bottom row.

Repeating this process, the subset of $G^{(p)}$ consisting of elements that are p^{th} powers $(g^p)^p$ of elements g^p of $G^{(p)}$ is the subset $G^{(p^2)}$ of G consisting of the powers g^{p^2} of elements g in G. The box diagram for $G^{(p^2)} = G_1^{(p^2)} \times \cdots \times G_k^{(p^2)}$ is obtained by deleting the bottom two rows of the diagram for G. Similarly, the powers g^{p^m} of elements g in G form a subset $G^{(p^m)}$ whose box diagram is obtained by deleting the bottom m rows of the diagram for G.

We noted before that if G has order p^n then the total number of boxes in the box diagram for G is n. In the same way the order of $G^{(p)}$ determines the number of boxes above the bottom row. Thus the orders of G and $G^{(p)}$ determine the number of boxes in the bottom row. Similarly, the number of boxes in the m^{th} row up from the bottom is determined by the orders of $G^{(p^{m-1})}$ and $G^{(p^m)}$. The diagram is completely determined by the numbers of boxes in each row, so the diagram is determined by the intrinsic structure of the group G using the intrinsically defined groups $G^{(p^m)}$ that are contained in G. Since the diagram determines the factorization of G as a product of cyclic groups of order a power of p, this finishes the proof. $\square$

The factorization of a finite abelian group as a product of cyclic groups of prime power order is the unique factorization with the largest number of factors since any other factorization with at least as many factors could be factored further into a product with prime power cyclic factors, contradicting the uniqueness statement in the preceding proposition.

On the other hand there can be different factorizations into cyclic factors with the smallest number of factors. For example, if p and q are distinct primes then $C_{p^2q^2} \times C_{pq}$ and $C_{p^2q} \times C_{pq^2}$ are both the group $C_{p^2} \times C_p \times C_{q^2} \times C_q$. A natural way to factor a group G as a product $G_1 \times \cdots \times G_k$ of cyclic groups with the minimum number of factors is by the following procedure. First factor G as a product of cyclic groups of prime power order. Place these groups in the boxes of a box diagram of the type considered in the previous proof, with one group in each box, so that each column consists of the groups of order a power of a fixed prime, arranged in order of decreasing size as one moves upward in the column. Let G_i be the product of the groups in the i^{th} row of the diagram, numbering the rows from the bottom to the top. Each G_i is a cyclic group since it is the product of cyclic groups of coprime orders. We have $G = G_1 \times \cdots \times G_k$ where k is the number of rows, which is the maximum number of prime power cyclic factors of G for any prime.

For the group $C_{p^2} \times C_p \times C_{q^2} \times C_q$ considered above this procedure yields the factorization $C_{p^2q^2} \times C_{pq}$. For a general finite abelian group G the procedure yields a factorization $G = C_{n_1} \times \cdots \times C_{n_k}$ with each n_i divisible by n_{i+1}. This is the same factorization of G as the one obtained in the proof of Theorem 7.9 since it is uniquely determined by the condition that each n_i is divisible by n_{i+1}.

Two Constructions with Squares

To conclude this section we describe two ways of using the squaring operation in a finite abelian group to build another somewhat simpler group. These constructions will be applied to class groups in the next two sections.

For the first construction we consider elements of a finite abelian group G whose square is the identity. These are the elements of order 1 or 2. These elements form a ***subgroup*** of G, that is, a subset which is a group in its own right. For a subset H of a group G to be a subgroup amounts to H satisfying three properties:

(1) The product of two elements of H is again in H, so within H there is a multiplication operation defined, the same multiplication as in G. The multiplication in H is automatically associative since multiplication in G is associative.

(2) H contains the identity element of G.

(3) The inverse of each element of H is in H.

These properties hold when H consists of the elements of order 1 or 2 in an abelian group G since property (1) means that if $g_1^2 = 1$ and $g_2^2 = 1$ for elements g_1 and g_2 of G then $(g_1g_2)^2 = 1$, which is true since $(g_1g_2)^2 = g_1^2g_2^2$ when G is abelian, while property (2) holds since the identity element of G has order 1 and (3) holds since a group element and its inverse always have the same order.

Proposition 7.14. *In a finite abelian group G the elements whose order is 1 or 2 form a subgroup of order 2^e where e is the number of factors of even order in any factorization of G as a product of cyclic groups. This subgroup is a product of e cyclic groups of order 2, and the order of G is a multiple of 2^e.*

In general, when a finite abelian group G is factored as a product of cyclic groups of prime power order, the number of factors of order a power of the prime p is called the p*-rank* of G. The number e in the proposition is thus the 2-rank of G. The proposition easily generalizes to the statement that the number of elements of G of order 1 or p is p^r where r is the p-rank of G.

Proof: Let $G = G_1 \times \cdots \times G_k$ be a factorization of G as a product of cyclic groups. An element $(g_1, \cdots, g_k)$ of the product has order 1 or 2 exactly when each g_i has order 1 or 2. A cyclic group C_{2n} of even order generated by an element g has just one element of order 2, the element g^n, since a power g^k with $0 < k < n$ has $g^{2k} \neq 1$ and the inverses of these elements are the powers g^k with $n < k < 2n$ so these too do not have order 2. A cyclic group of odd order has no elements of order 2 since the order of an element always divides the order of the group. Thus if e is the number of factors G_i of even order, there are e coordinates g_i of $(g_1, \cdots, g_k)$ where we have a choice of two elements of G_i of order 1 or 2 and in the other coordinates we must have $g_i = 1$. The elements of order 1 or 2 thus form a product of e cyclic groups of order 2. The last statement of the proposition is then obvious. □

Now we turn to the second construction. For any abelian group G we can form another group denoted G/G^2 whose elements are congruence classes of elements of G mod squares, so $g_1 \equiv g_2$ if $g_2 = g_1 g^2$ for some $g \in G$. This is analogous to taking congruence classes of integers mod 2 except now the group operation is multiplication rather than addition. The multiplication in G/G^2 comes from multiplication in

G, so if we denote the congruence class of $g \in G$ by $[g]$ then $[g_1][g_2]$ is defined to be $[g_1g_2]$. This is unambiguous since if $g_1 \equiv g_1'$ and $g_2 \equiv g_2'$, so $g_1' = g_1h_1^2$ and $g_2' = g_2h_2^2$ for some $h_1, h_2 \in G$, then $g_1g_2 \equiv g_1'g_2'$ since $g_1'g_2' = g_1g_2(h_1h_2)^2$. The identity element of G/G^2 is $[1]$ where 1 is the identity of G, and $[g]^{-1} = [g^{-1}]$. Associativity in G/G^2 follows from associativity in G, so G/G^2 is a group, which is abelian since G is abelian.

Proposition 7.15. *For a finite abelian group G factored as a product $G_1 \times \cdots \times G_k$ of cyclic groups G_i the group G/G^2 is a product of cyclic groups of order 2 with one factor for each factor G_i of even order.*

Thus G/G^2 factors as a product of cyclic groups in exactly the same way as the subgroup of G consisting of elements whose square is the identity, even though the constructions of these two groups are quite different.

Proof: For $G = G_1 \times \cdots \times G_k$ the square of an element $(g_1, \cdots, g_k)$ is $(g_1^2, \cdots, g_k^2)$, so the group G/G^2 is the product of the groups G_i/G_i^2. Thus the proposition reduces to the special case that G is a cyclic group. If G is cyclic of even order $2n$ with generator g then the squares in G are the even powers $g^2, g^4, \cdots, g^{2n}, g^{2n+2} = g^2, g^{2n+4} = g^4, \cdots$ which are all congruent to 1. The odd powers $g, g^3, \cdots, g^{2n-1}, g^{2n+1} = g, g^{2n+3} = g^3, \cdots$ are all congruent to each other but not to any even power of g so G/G^2 is cyclic of order 2. If G is cyclic of odd order $2n+1$ then the squares $g^2, g^4, \cdots, g^{2n}, g^{2n+2} = g, g^{2n+4} = g^3, \cdots$ form all of G so G/G^2 has order 1. □

Exercises

1. Show the converse of Proposition 7.8: If a product $G_1 \times G_2$ of finite abelian groups is cyclic then G_1 and G_2 are cyclic of coprime orders.

2. Show that if a prime p divides the order of a finite abelian group G then G contains an element of order p. For which nonprimes is this also true?

3. For each abelian group of order 4, 8, or 16 determine the number of elements of each possible order.

4. Determine the maximum order of elements of a finite abelian group G in terms of the factorization of G as a product of cyclic groups of prime power order, and show that the orders of elements of G are exactly all the divisors of this maximal order.

5. (a) State and prove the analogue of Proposition 7.14 with 2 replaced by an odd prime p.
(b) Do the same for Proposition 7.15.

6. This problem concerns the question of when the group $\mathbb{Z}_n^*$ of congruence classes mod n of integers coprime to n is cyclic.

(a) Show that $\mathbb{Z}_2^*$ and $\mathbb{Z}_4^*$ are cyclic but $\mathbb{Z}_8^*$ is not cyclic and deduce that $\mathbb{Z}_{2^r}^*$ is also not cyclic when $r > 3$.
(b) Show that if $\mathbb{Z}_n^*$ is cyclic then $n = 2, 4, p^r$, or $2p^r$ for some odd prime p. *Hint*: $\mathbb{Z}_n^*$ has even order if $n > 2$.
(c) The group $\mathbb{Z}_{p^r}^*$ is known to be cyclic when p is an odd prime. Show that this implies that $\mathbb{Z}_{2p^r}^*$ is cyclic.

7. Describe each of the following groups $\mathbb{Z}_n^*$ as a product of cyclic groups and draw a Cayley graph: $\mathbb{Z}_{10}^*$, $\mathbb{Z}_{13}^*$, $\mathbb{Z}_{15}^*$, $\mathbb{Z}_{24}^*$, and $\mathbb{Z}_{60}^*$.

7.4 Symmetry and the Class Group

We have defined the symmetric class number h_Δ^s for discriminant Δ to be the number of equivalence classes of primitive forms of discriminant Δ whose topographs have mirror symmetry. Thus h_Δ^s is the number of elements in the class group $CG(\Delta)$ whose order is 1 or 2 since mirror symmetric forms correspond to elements of $CG(\Delta)$ satisfying $Q = Q^{-1}$, which is the same as saying $Q^2 = 1$. (For symmetric forms there is no distinction between equivalence and proper equivalence.) As we saw in the discussion before Proposition 7.14, these elements form a subgroup of $CG(\Delta)$ which could be called the ***symmetric class group*** with the notation $SCG(\Delta)$. Its order is h_Δ^s, and it is a product of cyclic groups of order 2 since each element has order 1 or 2.

From Proposition 7.14 we can immediately deduce the following result:

Proposition 7.16. (a) *The symmetric class number h_Δ^s is equal to 2^r where r is the 2-rank of $CG(\Delta)$, the number of cyclic factors of $CG(\Delta)$ of order a power of 2 when $CG(\Delta)$ is expressed as a product of cyclic groups of prime-power order.*
(b) *The ordinary class number h_Δ is always a multiple of h_Δ^s, with $h_\Delta = h_\Delta^s$ exactly when $CG(\Delta)$ is a product of cyclic groups of order 2, and $h_\Delta^s = 1$ exactly when h_Δ is odd.* □

Applying Theorem 5.9 which computed h_Δ^s in terms of the prime factorization of Δ we conclude:

Corollary 7.17. *If the number of distinct prime divisors of Δ is k then the 2-rank of $CG(\Delta)$ is $k-1$ except when $\Delta = 4(4m+1)$ when the 2-rank is $k-2$, and when $\Delta = 32m$ when the 2-rank is k. In particular the 2-rank is $k-1$ when Δ is a fundamental discriminant.* □

From this corollary we can deduce another:

Corollary 7.18. *If $|\Delta|$ is prime then the class number h_Δ is odd.* □

We know that $CG(\Delta)$ is cyclic if the class number is prime or a product of distinct primes, but there are other cases when the structure of $CG(\Delta)$ as a product of cyclic groups is completely determined if one knows the class number as well as the prime factorization of Δ, using the fact that the latter determines the 2-rank of $CG(\Delta)$ as in Corollary 7.17. For example if the class number is 4 then $CG(\Delta)$ is either C_4 or $C_2 \times C_2$ and these two cases are distinguished by their 2-ranks. We saw this distinction between C_4 and $C_2 \times C_2$ for the fundamental discriminants -56 and -84 both of which have class number 4, but -56 has two distinct prime divisors so its class group is C_4 while -84 has three distinct prime divisors so its class group is $C_2 \times C_2$.

A similar thing works for class number 8 where the group is either C_8, $C_4 \times C_2$, or $C_2 \times C_2 \times C_2$, with different 2-ranks. On the other hand, for class number 16 there is an ambiguity between $C_8 \times C_2$ and $C_4 \times C_4$. The first negative discriminant with class number 16 is $\Delta = -399 = -3 \cdot 7 \cdot 19$, a fundamental discriminant. Since there are three distinct prime factors of Δ the 2-rank of $CG(\Delta)$ is 2 so the ambiguity between $C_8 \times C_2$ and $C_4 \times C_4$ arises here. It is easy to compute that there are ten reduced forms of discriminant -399:

$$\begin{array}{ccccc} [1,1,100] & [2,1,50] & [4,1,25] & [5,1,20] & [10,1,10] \\ [3,3,34] & [6,3,17] & [7,7,16] & [8,7,14] & [10,9,12] \end{array}$$

Labeling these as $Q_1, \cdots, Q_5$ in the first row and $Q_6, \cdots, Q_{10}$ in the second row, we see that there are four forms with mirror symmetry, Q_1, Q_5, Q_6, Q_8, the forms with two of their coefficients equal. This is in agreement with the 2-rank being 2. The six without symmetry count double in the class number which is therefore 16. To determine whether the class group is $C_8 \times C_2$ or $C_4 \times C_4$ it suffices to look for elements of order greater than 4. This happens to be very easy in this case if we look at which forms represent powers of 2. In the list above we see that Q_2 represents 2, Q_3 represents 4, Q_9 represents 8, and Q_8 represents 16. Since powers of primes not dividing the discriminant are always represented by unique equivalence classes of forms, it follows that $Q_2^2 = Q_3^{\pm 1}$, $Q_2^3 = Q_9^{\pm 1}$, and $Q_2^4 = Q_8$, with no sign ambiguity in the last case since Q_8 has mirror symmetry. In particular we see that Q_2 must have order greater than 4, so $CG(\Delta)$ is not $C_4 \times C_4$ and hence it must be $C_8 \times C_2$.

The order of Q_2 is 8 since there are no elements of order 16 in $C_8 \times C_2$. (This also follows from the fact that Q_2^4 has mirror symmetry hence must have order 2.) As in the proof of Theorem 7.9 we can choose Q_2 as a generator of the C_8 factor of $CG(\Delta)$, and a generator of the C_2 factor can be chosen to be either Q_5 or Q_6, the two forms with mirror symmetry that are not a power of Q_2. Additional work would be needed to compute the remaining products Q_iQ_j such as whether Q_2^2 is Q_3 or Q_3^{-1}. However some products can be determined without calculation, for example the fact that the product of any two of the symmetric forms Q_5, Q_6, Q_8 equals the third since the product of two elements of order 2 must have order 1 or 2, but for example Q_5Q_6 cannot be the identity element Q_1 nor can it be Q_5 or Q_6 so it must be Q_8. Thus the elements Q_1, Q_5, Q_6, and Q_8 form a subgroup $C_2 \times C_2$. This is just the symmetric class group $SCG(\Delta)$.

A similar but even simpler sort of ambiguity occurs for class numbers p^2 with p an odd prime, where the choice is between the groups C_{p^2} and $C_p \times C_p$. The first example of this sort among negative discriminants occurs when $\Delta = -199$. The reduced forms are $Q_1 = [1,1,50]$, $Q_2 = [2,1,25]$, $Q_3 = [5,1,10]$, $Q_4 = [4,3,13]$, and $Q_5 = [7,5,8]$. Only Q_1 has mirror symmetry so the other four forms count twice in the class number which is therefore 9. To decide whether $CG(\Delta)$ is C_9 or $C_3 \times C_3$ we observe that Q_2 represents 2, Q_4 represents 2^2, and Q_5 represents 2^3, so Q_2 must have order greater than 3 in $CG(\Delta)$. Since the order of Q_2 must divide the order of $CG(\Delta)$ we see that Q_2 has order 9 and so $CG(\Delta)$ is C_9 rather than $C_3 \times C_3$.

The order of the class group can be made arbitrarily large by taking Δ to have a large number of distinct prime factors, using a product of distinct odd primes if one wants a fundamental discriminant. It is also possible for individual elements of the class group to have large order:

Proposition 7.19. *For arbitrary integers $a > 1$ and $n > 1$ the form $[a,1,a^{n-1}]$ has order n in $CG(\Delta)$ for $\Delta = 1 - 4a^n$.*

Proof: The form $[a,1,a^{n-1}]$ is concordant to itself if $n > 1$ and we can use this fact to compute its powers inductively as in the proof of Theorem 7.7, with the result that $[a,1,a^{n-1}]^k = [a^k,1,a^{n-k}]$. When $k = n$ the latter form is $[a^n,1,1]$ which represents 1 so it is the identity element in the class group. Thus the order of $[a,1,a^{n-1}]$ is a divisor of n. The discriminant $1 - 4a^n$ is negative and the forms $[a^k,1,a^{n-k}]$ are reduced if $k \le n-k$, or in other words if $k \le n/2$. None of these reduced forms is the principal form if $a > 1$ so none is the identity in $CG(\Delta)$. Thus the order of $[a,1,a^{n-1}]$ is greater than $n/2$ so it must be n. $\square$

In general it is a hard question to determine which finite abelian groups occur as class groups. An interesting special case is to determine the values of n such that the product of n cyclic groups of order 2 is a class group $CG(\Delta)$ for some Δ. By Proposition 7.16 this is equivalent to having $h_\Delta = h^s_\Delta$, and we have mentioned that there is a list, probably complete, of 101 negative discriminants Δ with this property. In these 101 cases the number of C_2 factors of $CG(\Delta)$ ranges from 0 to 4, so the class number is 1, 2, 4, 8, or 16. Thus it appears that a product of five or more copies of C_2 cannot occur as a class group $CG(\Delta)$ with $\Delta < 0$. For $\Delta > 0$ less seems to be known.

Here is a table listing the smallest discriminants having class group a given abelian group of order up to 12:

$CG(\Delta)$	C_1	C_2	C_3	C_4	$C_2 \times C_2$	C_5	C_6	C_7
$\Delta < 0$	-3	-15	-23	-39	-84	-47	-87	-71
$\Delta > 0$	5	12	148	136	60	401	316	577

C_8	$C_4 \times C_2$	$C_2 \times C_2 \times C_2$	C_9	$C_3 \times C_3$	C_{10}	C_{11}	C_{12}	$C_6 \times C_2$
-95	-224	-420	-199	-4027	-119	-167	-279	-231
505	396	480	1129	32009	817	1297	1345	940

As one can see, for positive discriminants one usually needs to go farther than for negative discriminants to realize a given group.

Positive Discriminants

While positive discriminants are more difficult both computationally and theoretically, they have an extra piece of structure that adds to their interest, the operation that sends a form Q to its negative $-Q$. This gives a well-defined operation on $CG(\Delta)$ since if two forms Q_1 and Q_2 are properly equivalent then so are $-Q_1$ and $-Q_2$ because an orientation-preserving linear fractional transformation taking the topograph of Q_1 to the topograph of Q_2 takes the topograph of $-Q_1$ to the topograph of $-Q_2$. Also, if Q is primitive then obviously so is $-Q$.

In $CG(\Delta)$ the operation sending Q to $-Q$ is generally different from the operation which sends Q to its mirror image form Q^{-1} in $CG(\Delta)$. For example when $\Delta = 12$ the group $CG(\Delta)$ is cyclic of order 2 consisting of the principal form $Q = x^2 - 3y^2$ and its negative $-Q = -x^2 + 3y^2$ which is equivalent to $3x^2 - y^2$. Thus Q and $-Q$ are distinct elements of $CG(\Delta)$, but $Q = Q^{-1}$ and $-Q = -Q^{-1}$ since Q and $-Q$ have mirror symmetry. Note that there is never any ambiguity about whether $-Q^{-1}$ is $-(Q^{-1})$, the negative of the mirror image of Q, or $(-Q)^{-1}$, the mirror image of the negative of Q, since these are obviously the same.

Proposition 7.20. *Inverses and negatives are related to symmetries and skew symmetries in the following ways:*

(a) $Q = Q^{-1}$ *in* $CG(\Delta)$ *if and only if the topograph of* Q *has a mirror symmetry.*

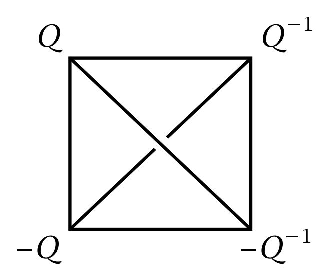
(b) $Q = -Q$ *in* $CG(\Delta)$ *if and only if the topograph of* Q *has a* 180 *degree rotational skew symmetry.*

(c) $Q = -Q^{-1}$ *in* $CG(\Delta)$ *if and only if the topograph of* Q *has a glide reflection skew symmetry.*

Proof: We have already seen that (a) holds. Statements (b) and (c) apply only to hyperbolic forms, in which case we can focus on what is happening along the separator lines in their topographs. We take separator lines to be drawn in the usual way as horizontal lines with positive values above and negative values below. We can assume that the edges leading off the separator line occur at unit intervals.

For (b), the separator line for the negative of a form Q is obtained by first changing the sign of all the labels along the separator line for Q and then rotating the plane by 180 degrees about some point on the separator line to bring the positive labels back above the separator line. If Q is properly equivalent to $-Q$ this means that these two operations of changing signs and rotating produce the same separator line we started with, up to horizontal translation. Thus the composition of a rotation and a translation gives a skew symmetry of the separator line of Q. The two ends of the line are interchanged by this skew symmetry so it must fix some point on the line, as we saw in the discussion of symmetries of hyperbolic forms in Section 5.4. Hence the skew symmetry must be a rotation about this point of the separator line. Thus if $Q = -Q$ in $CG(\Delta)$, the topograph of Q has a 180 degree rotational skew symmetry. The converse is obviously true as well.

For (c), we can transform the separator line of a form Q to the separator line of $-Q^{-1}$ by first changing the signs of the labels and rotating by 180 degrees to get the separator line for $-Q$, then reflecting across a vertical line to convert this to the separator line for $-Q^{-1}$. The composition of the rotation and the reflection is a glide reflection along the separator line. Thus the separator line for Q is transformed into the separator line for $-Q^{-1}$ by a glide reflection and changing the sign of the labels. Hence if Q is properly equivalent to $-Q^{-1}$, the separator line for Q has a skew symmetry obtained by combining a glide reflection with a translation. This combination is again a glide reflection. □



We can picture the relationships between inverses and negatives by the diagram at the right which can be viewed as a picture of a regular tetrahedron. The tetrahedron has three 180 degree rotational symmetries about the three axes passing through midpoints of opposite edges of the tetrahedron. One of these rotations sends each form to its inverse, another sends each form to its negative, and the third sends each form to the negative of its inverse. These rotational symmetries of the tetrahedron are related to symmetries and skew symmetries of forms in the following ways:

- If Q has mirror symmetry then so does $-Q$ so the top two forms are equal in $CG(\Delta)$ and so are the bottom two. The first of the three rotational symmetries of the tetrahedron realizes these equalities in $CG(\Delta)$.
- If Q has a rotational skew symmetry then so does Q^{-1} so the two forms on the left are equal in $CG(\Delta)$ and so are the two on the right. These equalities are realized by the second rotation of the tetrahedron.
- If Q has a glide reflection skew symmetry then so does $-Q$ so the two forms in each diagonal pair are equal in $CG(\Delta)$, and the third rotation of the tetrahedron gives these equalities.

When Q has two of the three types of symmetries and skew symmetries, it has the third type as well, so all four forms are equal in $CG(\Delta)$. In this case we will say that Q is *fully symmetric*. For example the principal form always has mirror symmetry and represents 1 so it is fully symmetric exactly when it represents -1 since Proposition 6.16 says this is equivalent to its having a skew symmetry.

Now let us see how negation of forms relates to multiplication in $CG(\Delta)$. One might guess that $(-Q_1)Q_2 = -(Q_1Q_2)$ as with numbers, but this turns out to be not quite right as the following lemma shows:

Lemma 7.21. *In $CG(\Delta)$ the formula $(-Q_1)Q_2 = -(Q_1Q_2^{-1})$ holds for all Q_1 and Q_2. In particular, when $Q_1 = Q_2$ we have $(-Q_1)Q_1 = -Q_0$ where Q_0 is the principal form.*

Proof: The forms Q_1 and Q_2 are properly equivalent to a pair of concordant forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$. The form $[-a_1, -b, -a_2c]$ is then concordant to the form $[a_2, -b, (-a_1)(-c)] = [a_2, -b, a_1c]$. Taking the product of this pair of concordant forms gives $[-a_1, -b, -a_2c][a_2, -b, a_1c] = [-a_1a_2, -b, -c]$. This says that $(-Q_1)(Q_2^{-1}) = -(Q_1Q_2)$. Replacing Q_2 by Q_2^{-1} then gives the claimed formula $(-Q_1)Q_2 = -(Q_1Q_2^{-1})$. □

Proposition 7.22. *If one element of $CG(\Delta)$ has a glide reflection skew symmetry then so do all elements of $CG(\Delta)$. This occurs exactly for those discriminants for which the principal form represents -1.*

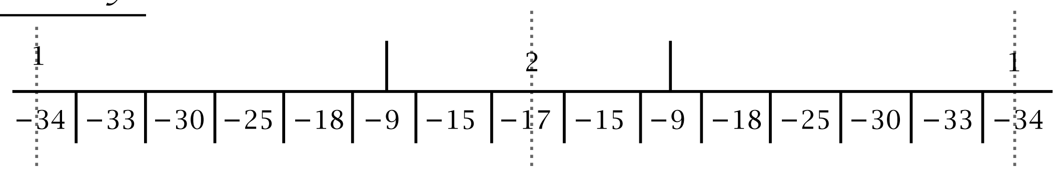
Proof: Suppose that Q is a form with a glide reflection skew symmetry, so $Q = -Q^{-1}$ or equivalently $-Q = Q^{-1}$. Then if Q_0 is the principal form, we have $Q_0 = Q^{-1}Q = (-Q)Q$ and this equals $-Q_0$ by the previous lemma. Thus $Q_0 = -Q_0$ if a single form has a glide reflection skew symmetry. Once one has $Q_0 = -Q_0$, then for arbitrary Q the formula $(-Q)Q = -Q_0$ says that Q is the inverse of $-Q$, so $Q = -Q^{-1}$ which means that Q has a glide reflection skew symmetry. This proves the first statement of the proposition. The second statement then follows since the principal form has a glide reflection skew symmetry exactly when it represents -1. □

Corollary 7.23. *If the class number h_Δ is odd then all forms in $CG(\Delta)$ have a glide reflection skew symmetry but only the principal form has a rotational skew symmetry.*

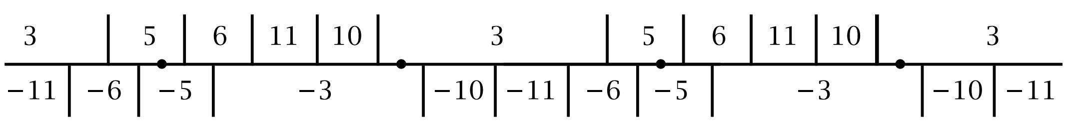
Proof: The principal form Q_0 has mirror symmetry and therefore so does $-Q_0$. Thus $(-Q_0)^2 = Q_0$. If $CG(\Delta)$ has odd order then it has no elements of order 2 so we must have $-Q_0 = Q_0$. Thus Q_0 has a rotational skew symmetry so it must also have a glide reflection skew symmetry. By the preceding proposition all forms in $CG(\Delta)$ then have a glide reflection skew symmetry. Any form which had a rotational skew symmetry would therefore also have a mirror symmetry and hence be of order 1 or 2 in $CG(\Delta)$, so it would have to be Q_0. □

One might ask whether the "one implies all" property in Proposition 7.22 also holds for the other two types of symmetries and skew symmetries. For mirror symmetries the only time all elements of $CG(\Delta)$ have mirror symmetry is when $CG(\Delta)$ is a product of cyclic groups of order 2, a rather rare occurrence that we have discussed before. For rotational skew symmetries it can happen that some forms have rotational skew symmetry while others do not. We just saw that when $CG(\Delta)$ has odd order only the principal form has rotational skew symmetry. An example where another form has rotational skew symmetry but the principal form does not is $\Delta = 136$. Here it is not hard to compute that there are three equivalence classes of forms: $Q_0 = [1, 0, -34]$, $-Q_0 = [-1, 0, 34]$, and $Q_1 = [3, 2, -11]$. Here are the topographs of Q_0 and Q_1:

$Q_0 = x^2 - 34y^2$



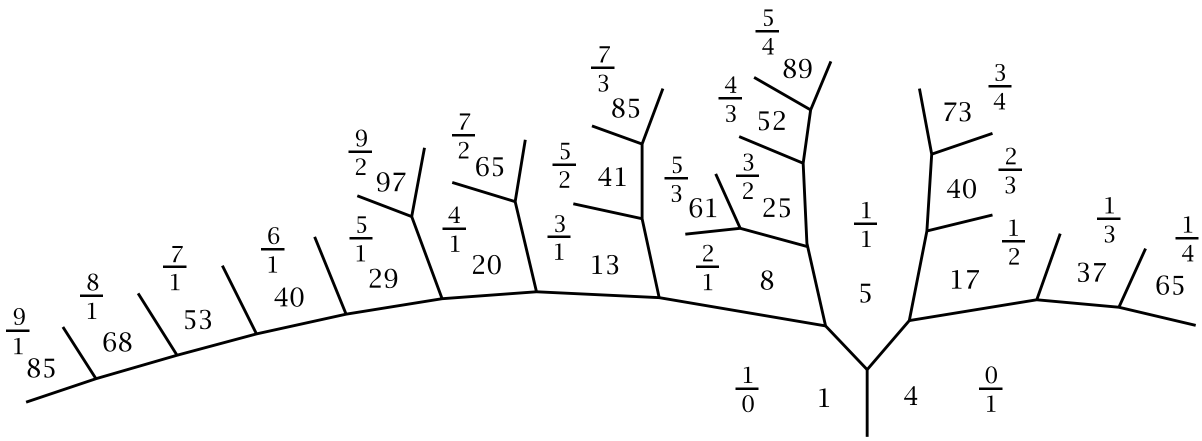
$Q_1 = 3x^2 + 2xy - 11y^2$



Since Q_0 and $-Q_0$ have mirror symmetry while Q_1 does not, the class number is 4. The group $CG(\Delta)$ must be C_4 rather than $C_2 \times C_2$ since it contains a form Q_1 without mirror symmetry, so this form has order 4 rather than 2. Thus Q_1^2 has order 2 so it must be the form $-Q_0$, as is confirmed by the fact that Q_1 represents 3 and $-Q_0$ represents 9. The topographs show that only Q_1 and Q_1^{-1} have a rotational skew symmetry.

When do all primitive forms of discriminant Δ have a rotational skew symmetry? If this happens then in particular the principal form has a rotational skew symmetry, as well as a mirror symmetry, so it also has a glide reflection skew symmetry. The previous proposition then says that all primitive forms have a glide reflection skew symmetry, in addition to the assumed rotational skew symmetry, so they have mirror symmetry as well. Thus the class group is a product of cyclic groups of order 2 and the principal form represents -1. Conversely, these two conditions imply that all principal forms have mirror symmetry and glide reflection skew symmetry, hence also rotational skew symmetry.

Another question one could ask is which discriminants have at least one primitive form with rotational skew symmetry. This turns out to have a very pleasing answer. As we observed near the end of Section 5.4, the pivot points of rotational skew symmetries lie at the midpoints of edges of the separator line where the labels of the adjacent regions in the topograph are a and $-a$. If the edge itself is labeled b then the associated form is $[a, b, -a]$, and all such forms occur this way at pivot points of rotational skew symmetries. The discriminant of the form $[a, b, -a]$ is $b^2 + 4a^2$ so we are looking for solutions of $x^2 + 4y^2 = \Delta$. For $[a, b, -a]$ to be primitive means that the pair (a, b) is primitive, so the question reduces just to finding the numbers represented by the form $x^2 + 4y^2$, excluding squares since we want the resulting forms $[a, b, -a]$ to be hyperbolic. (Squares correspond to 0-hyperbolic forms with rotational skew symmetry.) Here is a portion of the topograph of $x^2 + 4y^2$ showing also the labels $\frac{x}{y} = \frac{b}{a}$ that determine the associated forms $[a, b, -a]$:



The form $x^2 + 4y^2$ has discriminant -16 with class number 1. From Theorems 6.11 and 7.7 we can deduce that the numbers represented by $x^2 + 4y^2$ are the numbers $2^m p_1 \cdots p_k$ where m is 0, 2, or 3 and each p_i is a prime congruent to 1 mod 4. This tells us which discriminants have at least one primitive form with rotational skew symmetry.

A more refined question is how many different elements of $CG(\Delta)$ have rotational skew symmetries. Solutions of $b^2 + 4a^2 = \Delta$ come in groups of four obtained by varying the signs of a and b. If we restrict attention just to the solutions with a positive, the primitive solutions (a, b) correspond exactly to regions in the topograph of $x^2 + 4y^2$ labeled Δ, and these regions come in pairs, one in the upper half of the topograph with $b > 0$ and one in the lower half with $b < 0$. The sign of the label b on an edge of a topograph with a pivot point can be specified by orienting all edges of the separator line so that the regions on the left of the separator line have positive labels. Taking the mirror image topograph then corresponds to changing the sign of b. This might or might not give the same element of $CG(\Delta)$ depending on whether the topograph has mirror symmetry.

The topograph of a form with rotational skew symmetry has two pivot points on the separator line in each period. Thus the number of proper equivalence classes of primitive forms of discriminant Δ with rotational skew symmetry is half the number of regions labeled Δ in the topograph of $x^2 + 4y^2$, and is therefore equal to the number of such regions in the upper half of the topograph. In other words the number of elements of $CG(\Delta)$ with rotational skew symmetry equals the number of times that

Δ appears in the upper half of the topograph of $x^2 + 4y^2$. For example, a prime can appear only once in the upper half of the topograph by Proposition 6.16 so prime discriminants have only one element of $CG(\Delta)$ with rotational skew symmetry, and this element must have mirror symmetry.

In general the number of rotationally skew symmetric forms in $CG(\Delta)$ can be computed from the prime factorization of Δ using methods from the next chapter. The result is that if $\Delta = 2^m p_1^{e_1} \cdots p_k^{e_k}$ for distinct primes $p_i \equiv 1 \bmod 4$ with each $e_i > 0$ then the number of forms in $CG(\Delta)$ with rotational skew symmetry is 2^{k-1} when $m = 0$ or 2, and 2^k when $m = 3$.

Exercises

1. For discriminant $\Delta = -95$ first compute the class number by finding all the reduced forms, then determine the structure of the class group in two different ways, first by applying Corollary 7.17 and then by seeing which forms represent powers of 2 up to 2^4.

2. For discriminant $\Delta = -164$ determine the structure of the class group and find the orders of all its elements.

3. Do the same for discriminant $\Delta = -224$.

4. For discriminant $\Delta = 148$ determine the class group and also the symmetries and skew-symmetries of the forms of that discriminant.

5. Do the same for $\Delta = 145$.

6. (a) Show that the form $[2, 1, m]$ has order at least n in its class group if $2m > 2^n$.
(b) Show that the discriminant $1 - 8m$ in part (a) can be chosen to be a fundamental discriminant.
(c) Do the analogues of (a) and (b) using the form $[3, 2, m]$ of even discriminant.

7. Show that if a form Q of discriminant Δ represents a prime p coprime to Δ then p^k is represented by Q if and only if the order of Q in the class group divides $k - 1$ or $k + 1$.

7.5 Genus and the Class Group

The various genera of forms of discriminant Δ are determined by the characters χ associated to primes p dividing Δ, where χ assigns a value $\chi(n) = \pm 1$ to each integer n not divisible by p. Since each character has a constant value on all numbers in a topograph not divisible by p, we can regard characters as functions from $CG(\Delta)$ to $\{\pm 1\}$. A key property of characters is that they are multiplicative, so

$\chi(n_1n_2) = \chi(n_1)\chi(n_2)$. This implies that characters are also multiplicative as functions on $CG(\Delta)$, meaning that $\chi(Q_1Q_2) = \chi(Q_1)\chi(Q_2)$ for forms Q_1 and Q_2 defining elements of $CG(\Delta)$. This is because the topographs of Q_1 and Q_2 contain numbers n_1 and n_2 not divisible by p and coprime to each other by Proposition 6.26, and then the topograph of Q_1Q_2 contains n_1n_2. Thus $\chi(Q_1Q_2) = \chi(n_1n_2) = \chi(n_1)\chi(n_2) = \chi(Q_1)\chi(Q_2)$.

Since the values of characters are ± 1 this implies that $\chi(Q^2) = +1$ for each primitive form Q. Therefore $\chi(Q_1Q^2) = \chi(Q_1)\chi(Q^2) = \chi(Q_1)$ for all Q_1 and Q. This means that characters define functions on the group $CG(\Delta)/CG(\Delta)^2$ of congruence classes of forms modulo squares.

Let $G(\Delta)$ be the set of genera in discriminant Δ. Since forms that are congruent modulo squares have the same genus, there is a well-defined function Φ from $CG(\Delta)/CG(\Delta)^2$ to $G(\Delta)$ sending each congruence class of forms to the genus of these forms.

Theorem 7.24. *The function Φ from $CG(\Delta)/CG(\Delta)^2$ to $G(\Delta)$ is a one-to-one correspondence. Thus two primitive forms Q_1 and Q_2 of discriminant Δ belong to the same genus if and only if when we regard them as elements of $CG(\Delta)$ we have $Q_2 = Q_1Q^2$ for some primitive form Q of discriminant Δ.*

In particular, taking Q_1 to be the principal form, we see that a form is in the genus of the principal form if and only if it is the square of another form. This fact is sometimes called the *Gauss Duplication Theorem*, "duplication" referring to squaring. This special case in fact implies the general case since if Q_1 and Q_2 are of the same genus then all characters have the same values for Q_1 and Q_2, so all characters have the value $+1$ on $Q_2Q_1^{-1}$ which means that $Q_2Q_1^{-1}$ lies in the genus of the principal form, making $Q_2Q_1^{-1}$ a square Q^2 and hence $Q_2 = Q_1Q^2$.

We will give two proofs of the theorem. The first proof relies on Dirichlet's Theorem on primes in arithmetic progressions which we have not proved in this book, and which we have previously used only at the end of Section 6.3 in the proofs of Theorem 6.27 and Corollaries 6.28 and 6.29. The second proof will use only results proved in this book, notably Legendre's Theorem on solutions of $ax^2 + by^2 = cz^2$ from Section 2.3, but this proof has the disadvantage of applying only for fundamental discriminants. We will also be able to deduce a proof of Theorem 6.27 and its corollaries that does not use Dirichlet's Theorem, again just for fundamental discriminants.

First proof: By the definition of genus, every genus contains at least one form, so Φ is onto. Since a function between two finite sets with the same number of elements is one-to-one if and only if it is onto, it will suffice to show that $CG(\Delta)/CG(\Delta)^2$ and $G(\Delta)$ have the same number of elements. By Corollary 6.28 the number of genera is equal to the number of elements of $CG(\Delta)$ corresponding to forms with mirror symmetry, or in other words the elements of $CG(\Delta)$ of order 1 or 2. By Propositions 7.14 and 7.15 this equals the number of elements of $CG(\Delta)/CG(\Delta)^2$. □

For the second proof of Theorem 7.24 the main step will be the following:

Lemma 7.25. *If a primitive form belongs to the genus of the principal form then it represents a nonzero square.*

Proof: A primitive form $ax^2+bxy+cy^2$ of discriminant Δ represents some positive number coprime to 2Δ so after a change of variables we may assume a is this number. Thus a is positive, odd, and coprime to Δ. If the form belongs to the genus of the principal form, we wish to find an integer solution of $ax^2+bxy+cy^2 = z^2$ with $z \neq 0$. This is equivalent to finding a rational solution with $z \neq 0$ since a rational solution yields an integer solution by multiplying x, y, and z by a common denominator. Having an integer solution (x,y,z) means that the form $ax^2+bxy+cy^2$ represents a square since any common divisor of x and y will divide z and can be canceled from the equation.

After multiplying the equation $ax^2+bxy+cy^2 = z^2$ by $4a$ it becomes:

$$4a(ax^2+bxy+cy^2) = (2ax+by)^2 + (4ac-b^2)y^2 = 4az^2$$

If we let $w = 2ax+by$ this can be written as $w^2 - \Delta y^2 = 4az^2$ or $\Delta y^2 + 4az^2 = w^2$, and a rational solution of this equation will give a rational solution of the original equation $ax^2+bxy+cy^2 = z^2$ with $x = {}^{w-by}/_{2a}$. If we write Δ and $4a$ as squares times squarefree numbers Δ' and a' then the equation $\Delta y^2 + 4az^2 = w^2$ can be replaced by $\Delta' y^2 + a' z^2 = w^2$ by absorbing the square factors of Δ and $4a$ into y^2 and z^2. Since Δ and a were coprime, so are Δ' and a'.

We would like to apply Legendre's Theorem to the equation $\Delta' y^2 + a' z^2 = w^2$. The sign condition in the theorem is satisfied since a is positive, hence so is a'. The condition that the coefficients of y^2, z^2, and w^2 are coprime is satisfied since we chose a to be coprime to Δ. The remaining congruence conditions reduce to Δ' being a square mod a' and a' being a square mod Δ'. For the first of these two conditions we know that Δ is a square mod a since $\Delta = b^2 - 4ac$, hence Δ is a square mod each prime dividing a. From the multiplicative property of Legendre symbols it follows that Δ' is also a square mod these primes and in particular a square mod each prime dividing a'. These primes are odd since a is odd, so Δ' is a square mod a' by Lemma 6.4 since a' is a product of distinct primes.

Now consider the condition that a' is a square mod Δ'. This is equivalent to a' being a square mod each prime p dividing Δ' since Δ' is squarefree. For $p = 2$ this holds automatically. For odd p this means the Legendre symbols $\left(\frac{a'}{p}\right) = \left(\frac{a}{p}\right)$ have value $+1$, which they do if the form $ax^2+bxy+cy^2$ is in the genus of the principal form since this form represents a.

Thus Legendre's Theorem applies and there is a nontrivial integer solution of $\Delta' y^2 + a' z^2 = w^2$. This must have z nonzero, otherwise the equation would become $\Delta' y^2 = w^2$, forcing Δ' to be a square, hence also Δ, contrary to the standing hypothesis for this chapter. □

The lemma can be interpreted as a statement about rational points on quadratic curves: If the form $ax^2 + bxy + cy^2$ lies in the principal genus and its discriminant $b^2 - 4ac$ is not a square, then the curve $ax^2 + bxy + cy^2 = 1$ contains a rational point. As in Chapter 0 it then follows that the curve has a dense set of rational points, and hence the form $ax^2 + bxy + cy^2$ represents infinitely many squares.

The restriction to nonsquare discriminants is not actually necessary. In the proof of the lemma, if Δ is a nonzero square then we would have $\Delta' = 1$ so we would be looking for solutions of $y^2 + a' z^2 = w^2$. This has an obvious solution with $z = 0$ so the curve $y^2 + a' z^2 = 1$ has one rational point $(y, z) = (1, 0)$, hence it has a dense set of rational points with $z \neq 0$ and so the curve $ax^2 + bxy + cy^2 = 1$ also has a dense set of rational points. The case $\Delta = 0$ is rather trivial since the original primitive form $ax^2 + bxy + cy^2$ is then equivalent to the form $\pm x^2$. The negative parabolic form $-x^2$ must obviously be excluded, just as negative elliptic forms are excluded.

Second proof of Theorem 7.24, for fundamental discriminants: If a form Q is in the genus of the principal form then the lemma says it represents a nonzero square n^2. Let the prime factorization of n^2 be $p_1^{2r_1} \cdots p_k^{2r_k}$ for distinct primes p_i. If the discriminant is a fundamental discriminant then Theorem 7.7 says that Q has a corresponding factorization $Q = Q_1^{2r_1} \cdots Q_k^{2r_k}$ in $CG(\Delta)$. Hence Q is the square of $Q_1^{r_1} \cdots Q_k^{r_k}$. □

For nonfundamental discriminants it is not always true that a primitive form that represents a square must be the square of another form. For example for $\Delta = -32$ the form $3x^2 + 2xy + 3y^2$ represents 4 when $(x, y) = (1, -1)$ but this form is not a square since the character χ_8 is defined for $\Delta = -32$ and has the value -1 on this form. However, if a primitive form represents a square coprime to the conductor then Theorem 7.7 does imply that the form is a square.

The first proof of Theorem 7.24 used the fact that the number of genera equals the symmetric class number, Corollary 6.28, whose proof depended on Dirichlet's Theorem. In the case of fundamental discriminants we can now give a different proof of Corollary 6.28 that does not use Dirichlet's Theorem but instead uses Legendre's Theorem, as follows. The symmetric class number is the number of elements of $CG(\Delta)$ whose square is the identity. By Propositions 7.14 and 7.15 this is the same as the number of elements of the group $CG(\Delta)/CG(\Delta)^2$. Using the second proof of Theorem 7.24, this is just the number of genera, so we obtain Corollary 6.28. Since Corollary 6.28 implies Theorem 6.27 using the calculation of the symmetric class number in Theorem 5.9, we also obtain a proof of Theorem 6.27 that does not use Dirichlet's Theorem, in the case of fundamental discriminants.

Let us illustrate the correspondence between elements of $CG(\Delta)/CG(\Delta)^2$ and genera by the example of discriminant $\Delta = -104$. We have already looked at this example in some detail earlier in the chapter where we saw that $CG(\Delta)$ is a cyclic group of order 6 generated by the form $Q_4 = [5,4,6]$. We have $\left(\frac{-104}{p}\right) = \left(\frac{-26}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{13}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{p}{13}\right)$. The product $\left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)$ is $+1$ for $p \equiv 1,3 \bmod 8$ and -1 for $p \equiv 5,7 \bmod 8$ so this is the character we called χ_8' in Section 6.3, while $\left(\frac{p}{13}\right)$ is χ_{13}, with the value $+1$ for $p \equiv 1,3,4,9,10,12 \bmod 13$ and -1 for $p \equiv 2,5,6,7,8,11 \bmod 13$. These are the two characters for $\Delta = -104$. Evaluating these characters on numbers not divisible by 2 or 13 in the topographs shown in Section 7.1, we see that Q_1 and $Q_3^{\pm 1}$ belong to one genus where the character values are $+1,+1$, while Q_2 and $Q_4^{\pm 1}$ make up the other genus with character values $-1,-1$. Expressing the forms as powers of the generator Q_4 we see that the even powers $Q_4^2 = Q_3^{-1}$, $Q_4^4 = Q_3$, and $Q_4^6 = Q_1$ form one genus and the odd powers Q_4, $Q_4^3 = Q_2$, and $Q_4^5 = Q_4^{-1}$ form the other genus. Thus two forms belong to the same genus exactly when one is a square times the other since the squares are the even powers of Q_4.

From Theorem 7.24 we can deduce the following interesting consequence of having a group structure in $CG(\Delta)$:

Corollary 7.27. *Each genus of forms of a given discriminant contains the same number of proper equivalence classes of forms.*

Proof: Let $Q_1, \cdots, Q_k$ be the distinct elements of $CG(\Delta)$ in the genus of the principal form. By Theorem 7.24 these are exactly the elements of $CG(\Delta)$ that are squares. The genus of an arbitrary element Q of $CG(\Delta)$ then consists of $QQ_1, \cdots, QQ_k$ since these are all the elements of $CG(\Delta)$ obtained by multiplying Q by squares. These multiples of Q are all distinct since if $QQ_i = QQ_j$ then after multiplying by Q^{-1} we have $Q_i = Q_j$ so $i = j$. Thus each genus consists of k elements of $CG(\Delta)$. □

For a fixed discriminant Δ the class number is the product of the number of genera times the number of classes in each genus. There are two extreme situations that can occur when one or the other of these two factors is 1:

(1) The number of genera is 1, so the primitive forms of discriminant Δ all have the same genus. Equivalent ways of stating this condition are:

- The only primitive forms with mirror symmetry are the forms equivalent to the principal form.
- $CG(\Delta)$ contains no elements of order 2.
- $CG(\Delta)$ contains no elements of even order.
- The class number is odd.

(2) Each genus consists of a single equivalence class of forms. Again there are equivalent statements:

- The number of genera equals the class number.
- Every form has mirror symmetry.
- Every element of $CG(\Delta)$ has order 2.
- $CG(\Delta)$ is a product of cyclic groups of order 2.
- The representation problem of determining which numbers are represented by each primitive form has a solution just in terms of congruence classes modulo the discriminant.

Discriminants where (1) or (2) occurs are rather rare. For (1), Corollary 5.10 says exactly when this happens in terms of the prime factorization of Δ. For (2) there is no such simple characterization.

An Exact Sequence

The relationships between the class group, genus, and symmetry can be expressed concisely in a sequence of groups and functions between them:

$$SCG(\Delta) \longrightarrow CG(\Delta) \xrightarrow{Sq} CG(\Delta) \xrightarrow{Ch} TS(\Delta) \xrightarrow{Pr} \{\pm 1\}$$

Here $SCG(\Delta)$ is the symmetric class group, the subgroup of $CG(\Delta)$ consisting of symmetric forms, and the function $SCG(\Delta) \to CG(\Delta)$ is just the inclusion of this subgroup into $CG(\Delta)$. The function Sq is squaring, sending a form Q to Q^2. The group $TS(\Delta)$ is the set of “total symbols” $(\pm 1, \cdots, \pm 1)$ with one coordinate for each character defined for discriminant Δ. The group structure in $TS(\Delta)$ is multiplication in each coordinate separately. The function Ch is the “total character” sending each form to the values of the various characters on this form. The last function Pr is the product of the coordinates of $TS(\Delta)$ corresponding to the characters in the product X_Δ defined in Proposition 6.22 that measures whether a prime not dividing Δ is represented in discriminant Δ. For fundamental discriminants this is all the characters and Pr is just the product of all the coordinates in $TS(\Delta)$.

The compositions of two successive functions in the five-term sequence above have a special property: For each pair of adjacent functions $A \xrightarrow{f} B \xrightarrow{g} C$ an element b in the middle group B is sent by g to the identity element of C exactly when b is equal to the image $f(a)$ of some element a in A. A sequence of functions with this property is called an *exact sequence*. Let us see what this means for each of the three middle groups in the five-term sequence above.

(1) Exactness at the first $CG(\Delta)$ term is the fact that the square Q^2 of a form Q is the identity in $CG(\Delta)$ exactly when Q is symmetric.

(2) Exactness at the second $CG(\Delta)$ term means that a form Q belongs to the genus of the principal form exactly when Q is the square of a form in $CG(\Delta)$. This is the Gauss Duplication Theorem.

(3) Exactness at $TS(\Delta)$ means that Pr has the value $+1$ on a total symbol exactly when this is the total symbol given by the character values of some primitive form. This is what we showed to prove Theorem 6.27.

In each case the easier half of the assertion is the statement obtained by omitting the word “exactly”.

Representations by Nonequivalent Forms

Now let us consider the relationship between genus and the simultaneous representation of numbers by forms of the same discriminant that are not equivalent.

Proposition 7.28. *If two primitive forms of the same discriminant represent the same number coprime to the conductor then the two forms are in the same genus.*

For numbers coprime to the discriminant this is a simple consequence of the definition of genus, but the result is less obvious in the more general situation, and indeed often fails to hold for numbers not coprime to the conductor. An example is discriminant -32 with conductor 2 where the two forms $[1,0,8]$ and $[3,2,3]$ both represent 8 but have different genus since the character χ_4 is defined when $\Delta = -32$ and has the value $+1$ on $[1,0,8]$ and -1 on $[3,2,3]$.

Proof: According to Theorem 7.7 we obtain the various primitive forms representing a number n coprime to the conductor as products $Q_1^{\pm e_1} \cdots Q_k^{\pm e_k}$ where the prime factorization of n is $n = p_1^{e_1} \cdots p_k^{e_k}$ and Q_i represents p_i. Changing the exponent of Q_i from $+e_i$ to $-e_i$ amounts to multiplying $Q_i^{e_i}$ by a square $Q_i^{-2e_i}$, and similarly for changing the exponent from $-e_i$ to $+e_i$. As we noted earlier, multiplying a form by the square of another form does not change its genus. So any two primitive forms representing n have the same genus. □

Proposition 7.29. *If two primitive forms are of the same genus then there exist numbers that are represented by both forms, and in fact there are infinitely many such numbers.*

Proof: If the primitive forms Q_1 and Q_2 of discriminant Δ have the same genus then there is a form Q such that $Q_2 = Q_1Q^2$ in $CG(\Delta)$. Choose a number k represented by Q_1. We can then choose a number m represented by Q and coprime to k, and after this a number n represented by Q and coprime to km, so all three of k, m, and n are coprime. Then kmn is represented by both $Q_1Q^2 = Q_2$ and $Q_1QQ^{-1} = Q_1$. There are infinitely many choices possible for n since new choices can always be made coprime to all previously chosen numbers. □

Exercises

1. Find all the instances in the large table in Section 6.2 where two primitive forms of the same discriminant but different genus represent the same power of the conductor.

2. For discriminant $\Delta = -260$ the equivalence classes of forms were worked out in Section 5.2. Show that $CG(\Delta)$ is $C_4 \times C_2$, partition the forms into genera, and determine the order of each element of $CG(\Delta)$. Which elements are squares of other elements?

3. For discriminant $\Delta = -119 = -7 \cdot 17$ show that $CG(\Delta)$ is cyclic, determine its order, and find forms giving all the elements. Then partition these elements according to their genus and determine the order of each element. (All this can be done without actually computing any products using concordant pairs of forms.)

7.6 Rational Equivalence and Rational Forms

Legendre's Theorem shows that determining when quadratic curves contain rational points is much easier than determining when they contain integer points. Pursuing this idea, our goal in this section will be to see how the general theory of quadratic forms becomes much simpler when rational numbers are used in place of integers, and in fact reduces largely to genus theory.

As an illustration consider the two forms $Q_1(x,y) = x^2 + 14y^2$ and $Q_2(x,y) = 2x^2 + 7y^2$ of discriminant -56 that we considered in Section 6.1. These forms have the same genus since the two characters for this discriminant are χ_7 and χ_8 which both take the value $+1$ on the two forms. We could also deduce this from Proposition 7.28 since both forms represent 15. However, the two forms are not equivalent. This means that there is no matrix $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ with integer entries and determinant ± 1 such that $Q_1(px + qy, rx + sy) = Q_2(x,y)$. But if we broaden our perspective to allow rational numbers as entries then there is such a matrix, namely the matrix $\frac{1}{3}\left(\begin{smallmatrix} 2 & 7 \\ -1 & 1 \end{smallmatrix}\right)$ of determinant $+1$, since a simple calculation shows that $Q_1(2x/3 + 7y/3, -x/3 + y/3) = Q_2(x,y)$. There are other matrices that could be used instead of $\frac{1}{3}\left(\begin{smallmatrix} 2 & 7 \\ -1 & 1 \end{smallmatrix}\right)$, for example $\frac{1}{5}\left(\begin{smallmatrix} 6 & 7 \\ -1 & 3 \end{smallmatrix}\right)$.

This example leads us to define two forms Q_1 and Q_2 to be ***rationally equivalent*** if there exists a matrix $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ with rational entries and nonzero determinant such that $Q_1(px+qy, rx+sy) = Q_2(x,y)$. The determinant condition ensures that the matrix has an inverse, also with rational entries, so the change of variables is reversible. In the example the determinant was $+1$, and in this case the forms are said to be *properly* rationally equivalent, or more briefly, properly $\mathbb{Q}$-equivalent.

Having allowed rational coefficients when we change variables, we can go a step further and consider ***rational forms*** $ax^2 + bxy + cy^2$ where the coefficients and variables are all allowed to be rational numbers. Rational equivalence of rational forms is defined just as it was for integral forms in the previous paragraph.

To see the effect of a rational change of variables on the discriminant of a rational form we can use the matrix notation $\left(\begin{smallmatrix} a & \underline{b} \\ \underline{b} & c \end{smallmatrix}\right)$ for a form $[a, b, c]$ from Section 7.1, where $\underline{b} = b/2$. The discriminant $b^2 - 4ac$ is -4 times the determinant of this matrix. When we change variables via a rational matrix $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ the new form corresponds to the matrix $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)\left(\begin{smallmatrix} a & \underline{b} \\ \underline{b} & c \end{smallmatrix}\right)\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ so the discriminant $b^2 - 4ac$ is multiplied by the square of the determinant of the change-of-variables matrix since $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ have the same determinant. In particular this means that properly $\mathbb{Q}$-equivalent forms have the same discriminant.

Using rational numbers gives added flexibility to prove certain statements that do not hold when only integers are allowed. Here are some instances of this:

Proposition 7.30. (a) *If a rational form takes on the nonzero value a then it is properly $\mathbb{Q}$-equivalent to a form $[a, 0, c]$. In particular every rational form is properly $\mathbb{Q}$-equivalent to a form $[a, 0, c]$.*
(b) *If two rational forms of the same discriminant take on the same nonzero value then they are properly $\mathbb{Q}$-equivalent.*

Since the discriminant of a form $[a, 0, c]$ is $-4ac$ it follows that c is determined by a and the discriminant, namely $c = -\Delta/4a$. For example the two forms $[1, 0, 14]$ and $[2, 0, 7]$ of discriminant -56 both take the value 15 so by part (a) of the proposition they are both properly $\mathbb{Q}$-equivalent to $[15, 0, 14/15]$ and hence to each other. As another example the principal form $x^2 + xy + ky^2$ of discriminant $1 - 4k$ takes the value 1 so it is properly $\mathbb{Q}$-equivalent to $x^2 + \frac{4k-1}{4}y^2$ and this form is rationally equivalent to $x^2 + (4k - 1)y^2$, the principal form of discriminant $4(1 - 4k)$.

Proof: Let Q be a rational form taking the nonzero value a when $(x, y) = (p, q)$ for rational numbers p and q, not both zero. The numbers p and q form the first column of a matrix $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)$ of determinant 1 since the equation $ps - qr = 1$ always has a solution with rational numbers r and s. For example, if $p \neq 0$ we can choose $r = 0$ and $s = 1/p$ and if $q \neq 0$ we can choose $s = 0$ and $r = -1/q$. We use the matrix $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)$ to change variables to get a new form $Q(px + ry, qx + sy)$ properly $\mathbb{Q}$-equivalent to Q whose value at $(x, y) = (1, 0)$ is $Q(p, q) = a$. Thus Q is properly $\mathbb{Q}$-equivalent to a form $[a, b, c]$ for some rational numbers b and c. This form can be rewritten as:

$$ax^2 + bxy + cy^2 = a\left(x + \frac{b}{2a}y\right)^2 + \left(c - \frac{b^2}{4a}\right)y^2$$

Thus if we change variables to $X = x + b/2a\, y$ and $Y = y$ the form $[a, b, c]$ becomes $[a, 0, c']$ for $c' = c - b^2/4a$. The matrix for this change of variables has determinant 1 so the form $[a, 0, c']$ is properly $\mathbb{Q}$-equivalent to $[a, b, c]$ and hence also to the original form Q. This proves statement (a).

Statement (b) follows from (a) since the coefficient c in a form $[a, 0, c]$ is determined by a and the discriminant when $a \neq 0$. □

For the next proposition we return to forms with integer coefficients.

Proposition 7.31. *Primitive forms of the same genus are properly $\mathbb{Q}$-equivalent. For fundamental discriminants the converse is also true: Properly $\mathbb{Q}$-equivalent forms have the same genus.*

An example showing the necessity of the extra hypothesis in the converse is provided by the forms $[1, 0, 8]$ and $[3, 2, 3]$ of discriminant -32 which have different genus but are properly $\mathbb{Q}$-equivalent since they both represent 8.

Proof: For the first statement, two primitive forms of the same genus represent the same number by Proposition 7.29, and then the previous proposition says they are properly $\mathbb{Q}$-equivalent.

Conversely, suppose Q and Q' are primitive forms of discriminant Δ that are properly $\mathbb{Q}$-equivalent. Let k be a number represented by Q. If k is divisible by p^2 for some prime p, say $k = p^2 m$, then if Δ is a fundamental discriminant Theorem 7.7 implies that Q is equivalent to the product of a form representing m and the square of a form representing p. The form representing m is then in the same genus as Q and thus also properly $\mathbb{Q}$-equivalent to Q, so for proving the converse we can replace Q by this form. After repetitions of this step we can then assume that the number k represented by Q is squarefree.

Since Q and Q' are properly $\mathbb{Q}$-equivalent they take on the same rational values as the variables range over all rational numbers. Thus there exist integers x, y, z such that $Q'({}^x\!/_z, {}^y\!/_z) = k$ and hence $Q'(x, y) = kz^2$. We would like to say that Q' represents kz^2, and this will be the case if x and y are coprime. Suppose on the contrary that x and y are both divisible by some prime p. We can assume p does not divide z, otherwise the fractions ${}^x\!/_z$ and ${}^y\!/_z$ could be reduced. Since p divides x and y it follows that p^2 divides $Q'(x, y) = kz^2$ and hence p^2 divides k. This contradicts the fact that k is squarefree, so we deduce that Q' represents kz^2. Using Theorem 7.7 again and the assumption that Δ is a fundamental discriminant we conclude that Q' is the product of a form Q'' representing k and the square of some form representing z, so Q' and Q'' are in the same genus. Since Q and Q'' both represent k they have the same genus by Proposition 7.28. Hence Q and Q' have the same genus. □

In the remainder of this section we will describe the classification of rational forms up to rational equivalence. The first difference from the classification of integer forms up to integer equivalence as in Chapter 5 involves the discriminant. As we have seen, a change of variables by a matrix of nonzero rational determinant r multiplies the discriminant by r^2. For example the change of variables replacing (x, y) by (rx, y) has this effect. This leads us to consider nonzero rational numbers modulo squares, so two nonzero rational numbers are regarded as equivalent modulo squares if one is obtained from the other by multiplying by the square of a nonzero rational number. Every nonzero rational number is equivalent modulo squares to an integer since we can multiply by the square of its denominator. Thus p/q becomes pq, turning division into multiplication. After this, any square integer factor of the resulting integer can be eliminated by multiplying by the reciprocal of this square factor. In this way every equivalence class of nonzero rational numbers modulo rational squares contains a squarefree nonzero integer, and this integer is obviously unique.

In particular every nonzero discriminant is equivalent modulo squares to a unique nonzero squarefree integer discriminant which we call a *reduced discriminant*. When we speak of the reduced discriminant of a form we will mean the unique squarefree integer that is equivalent to its discriminant modulo squares. For example for a nonzero squarefree integer d the forms $x^2 - \frac{d}{4}y^2$ and $4x^2 - dy^2$ both have reduced discriminant d. Thus all squarefree nonzero integers occur as reduced discriminants. A reduced discriminant is a fundamental discriminant if it is congruent to $1 \bmod 4$, and otherwise four times the reduced discriminant is a fundamental discriminant.

A form Q and a nonzero rational multiple rQ have the same reduced discriminant. However Q and rQ may not be rationally equivalent. An example is provided by the forms $x^2 + y^2$ and $3x^2 + 3y^2$ with reduced discriminant -1, as we will soon see. On the other hand Q and r^2Q are rationally equivalent since $r^2Q(x, y) = Q(rx, ry)$. It follows that every rational form is rationally equivalent to an integer form, so it will suffice to classify integer forms up to rational equivalence.

Proposition 7.32. *If two rational forms of the same reduced discriminant take on the same nonzero value then they are rationally equivalent.*

Proof: Let the two forms be Q and Q'. Since they have the same reduced discriminant there is a rational number r such that the discriminant of Q' is r^2 times the discriminant of Q. The form $Q''(x, y) = Q(rx, y)$ has the same discriminant as Q' and is rationally equivalent to Q, hence has the same values as Q. Thus we may assume from the start that Q and Q' have the same discriminant. Proposition 7.30 then gives the result. □

For a fixed reduced discriminant δ all rational numbers r occur as values of rational forms of reduced discriminant δ since if Q_0 is the principal form for the associated fundamental discriminant then rQ_0 has the same reduced discriminant as Q_0 and takes the value r. Proposition 7.32 then says that the sets of nonzero values of forms of reduced discriminant δ give a partition of the set of all nonzero rational numbers into disjoint subsets, and these subsets correspond exactly to the rational equivalence classes of forms of reduced discriminant δ.

As a very special case, for reduced discriminant 1 there is the form xy and this takes on all rational values, so all rational forms of reduced discriminant 1 are rationally equivalent. This includes all 0-hyperbolic integer forms since the discriminants of these forms are nonzero squares.

To deal with the general case the following result will be useful:

Proposition 7.33. *The values taken on by a rational form $Q(x,y)$ as x and y range over all rational numbers are exactly the values $r^2Q(x,y)$ as x and y range over all integers and r ranges over all rational numbers.*

Proof: For each integer pair (x,y) and each rational number r we have $r^2Q(x,y) = Q(rx,ry)$ so rational squares times values at integer pairs are values at rational pairs. Conversely, if (x,y) is a rational pair there is a nonzero integer d such that (dx,dy) is an integer pair, and then $Q(x,y) = d^{-2}Q(dx,dy)$ so every value at a rational pair is a rational square times a value at an integer pair. □

Multiplying a form by a nonzero square does not affect the signs of its values, so the basic distinction between elliptic, hyperbolic, 0-hyperbolic, and parabolic forms still holds for rational forms. We have seen that all 0-hyperbolic forms are rationally equivalent. The classification of parabolic forms up to rational equivalence is easy and will be left as an exercise. This leaves hyperbolic and elliptic forms. For elliptic forms we can restrict attention to those taking positive values as we did for integer forms.

As a first example let us work out the classification of forms of reduced discriminant -1 up to rational equivalence. The associated fundamental discriminant is -4, with class number 1 so all integer forms of discriminant -4 are equivalent to x^2+y^2. The values of this form for integers x and y are all the positive numbers whose prime factorization contains primes $p \equiv 3 \bmod 4$ only to even powers. The values for rational x and y are then all such products where negative as well as positive exponents on primes are allowed.

Consider next the form $3x^2 + 3y^2$ which also has reduced discriminant -1. The values this form takes on for rational x and y can be described in the same way as for $x^2 + y^2$ except that now the exponent on the prime 3 must be odd rather than even. Thus this form is not rationally equivalent to x^2+y^2. More generally, for any finite set of primes $p_1, \cdots, p_k$ congruent to 3 mod 4 the values of the form $p_1 \cdots p_k(x^2+y^2)$ are the products in which each p_i has odd exponent. Different sets of primes $p_i \equiv 3$ mod 4, including the empty set for the form $x^2 + y^2$, give forms taking on disjoint sets of values, so all these sets of primes give different rational equivalence classes of forms. Every rational equivalence class is realized in this way since one can take any form in this class and any nonzero value r this form takes on, then take the set of primes $p_i \equiv 3 \bmod 4$ that occur to an odd power in the prime factorization of r. Thus we have determined all of the infinitely many rational equivalence classes of forms of reduced discriminant -1.

Other fundamental discriminants of class number 1 work in the same way. For example for discriminant -3 we have the form x^2+xy+y^2 whose values are products of primes in which primes $p \equiv 2 \bmod 3$ occur only to even powers. The rational equivalence classes then correspond to multiples of x^2+xy+y^2 by finite products of distinct primes $p \equiv 2 \bmod 3$. Instead of x^2+xy+y^2 we could use x^2+3y^2 which has the same reduced discriminant and is rationally equivalent to x^2+xy+y^2 since both forms take the value 3.

In the general case the rational classification of forms of a given reduced discriminant δ involves the different genera of forms of the associated fundamental discriminant Δ. By Proposition 7.31 each of these genera corresponds to exactly one rational equivalence class of forms. Choose one form Q_i in each of these genera. The values of integer forms of discriminant Δ are the numbers whose prime factorization contains certain primes only to even powers, namely the primes not represented in discriminant Δ, which are the primes in certain congruence classes mod Δ. The rational equivalence classes for reduced discriminant δ then correspond exactly to the forms $p_1 \cdots p_k Q_i$ where $p_1, \cdots, p_k$ are distinct primes not represented in discriminant Δ.

Exercises

1. What is the classification of rational forms of discriminant 0 up to rational equivalence?

2. Show that for each nonzero reduced discriminant δ there is a unique form x^2+by^2 of reduced discriminant δ with b a squarefree integer, and show that every form of reduced discriminant δ is rationally equivalent to a form $a(x^2+by^2)$.

8 Quadratic Fields

Even when one's primary interest is in integers it can sometimes be very helpful to consider more general sorts of numbers. For example, when studying the principal quadratic form $x^2 - Dy^2$ of discriminant $4D$ it can be a great aid to understanding to allow ourselves to factor this form as $(x + y\sqrt{D})(x - y\sqrt{D})$. Here we allow D to be negative as well as positive, in which case we would be moving into the realm of complex numbers.

To illustrate this idea, consider the case $D = -1$, so the form is $x^2 + y^2$ which we are factoring as $(x + yi)(x - yi)$. Writing a number n as a sum $a^2 + b^2$ is then equivalent to factoring it as $(a + bi)(a - bi)$. For example $5 = 2^2 + 1^2 = (2 + i)(2 - i)$, and $13 = 3^2 + 2^2 = (3 + 2i)(3 - 2i)$, so 5 and 13 are no longer prime when we allow factorizations using numbers $a + bi$. Sometimes a nonprime number such as 65 can be written as the sum of two squares in more than one way: $65 = 8^2 + 1^2 = 4^2 + 7^2$, so it has factorizations as $(8 + i)(8 - i)$ and $(4 + 7i)(4 - 7i)$. This becomes more understandable if we factorize 65 as:

$$65 = 5 \cdot 13 = (2 + i)(2 - i)(3 + 2i)(3 - 2i)$$

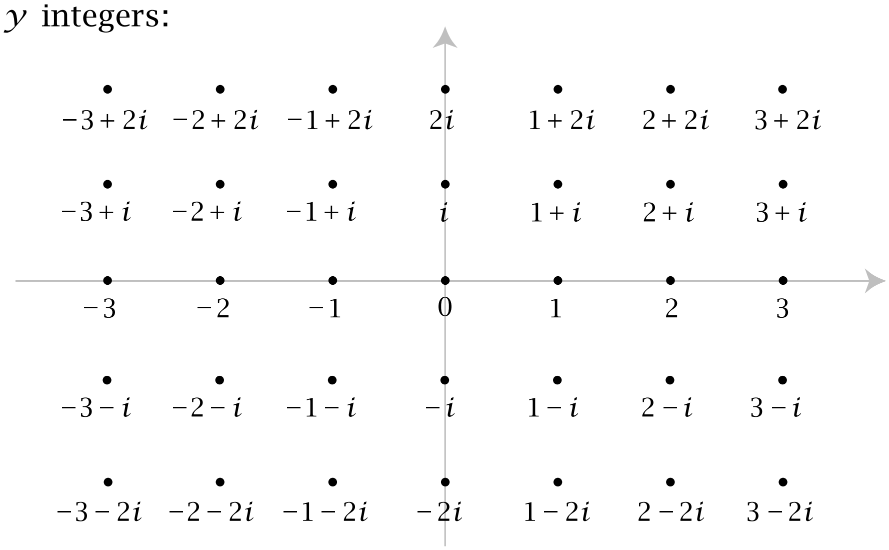
If we combine these four terms as $(2 - i)(3 + 2i) = 8 + i$ and $(2 + i)(3 - 2i) = 8 - i$ we get the representation $65 = 8^2 + 1^2 = (8 + i)(8 - i)$, whereas if we combine them as $(2 + i)(3 + 2i) = 4 + 7i$ and $(2 - i)(3 - 2i) = 4 - 7i$ we get the other representation $65 = 4^2 + 7^2 = (4 + 7i)(4 - 7i)$.

More generally we will consider the set $\mathbb{Z}[\sqrt{D}]$ of all numbers $x + y\sqrt{D}$ with x and y integers. Thus $\mathbb{Z}[\sqrt{D}]$ consists of real numbers if $D > 0$ and complex numbers if $D < 0$. We will always assume the integer D is not a square, so $\mathbb{Z}[\sqrt{D}]$ is not just $\mathbb{Z}$. When $D = -1$ we have $\mathbb{Z}[\sqrt{D}] = \mathbb{Z}[i]$, and numbers $x + yi$ in $\mathbb{Z}[i]$ are known as ***Gaussian integers.***

We will also have occasion to consider numbers $x + y\sqrt{D}$ where x and y are allowed to be rational numbers, not just integers. The set of all such numbers is denoted $\mathbb{Q}(\sqrt{D})$ with round parentheses instead of square brackets to emphasize that $\mathbb{Q}(\sqrt{D})$ is a *field* while $\mathbb{Z}[\sqrt{D}]$ is only a *ring.* In other words, in $\mathbb{Q}(\sqrt{D})$ we can perform all four of the basic arithmetic operations of addition, subtraction, multiplication, and division, whereas in $\mathbb{Z}[\sqrt{D}]$ only the first three operations are possible in general. Division by a nonzero element $x + y\sqrt{D}$ of $\mathbb{Q}(\sqrt{D})$ is possible since it amounts to multiplication by its reciprocal $1/(x + y\sqrt{D}) = (x - y\sqrt{D})/(x^2 - Dy^2)$ which lies in $\mathbb{Q}(\sqrt{D})$ when x and y are rational.

8.1 Prime Factorization

The ring of Gaussian integers $\mathbb{Z}[i]$ can be pictured as a subset of the plane, viewed as complex numbers in the usual way with $x + yi$ corresponding to the point with coordinates (x, y). Thus $\mathbb{Z}[i]$ forms a square grid consisting of the points (x, y) with x and y integers:



Similarly, the ring $\mathbb{Z}[\sqrt{D}]$ with $D < 0$ forms a grid of complex numbers forming rectangles of height $\sqrt{|D|}$ obtained by stretching the previous figure vertically.

When $D > 0$ the numbers in $\mathbb{Z}[\sqrt{D}]$ are real numbers which would normally be regarded as points along the x-axis. However, there is another point of view that will make the case $D > 0$ look just like the case $D < 0$, and this is to regard a number $x + y\sqrt{D}$ in $\mathbb{Z}[\sqrt{D}]$ or more generally $\mathbb{Q}(\sqrt{D})$ as the point $(x, y\sqrt{D})$ in the plane. Thus for example $\mathbb{Z}[\sqrt{2}]$ is exactly the same rectangular grid as $\mathbb{Z}[\sqrt{-2}]$, with rectangles of width 1 and height $\sqrt{2}$. From this point of view the horizontal and vertical axes of the plane, instead of being the real and imaginary axes, are now regarded as the "rational and irrational axes", with the two coordinates x and $y\sqrt{D}$ being the rational and irrational parts of $x + y\sqrt{D}$.

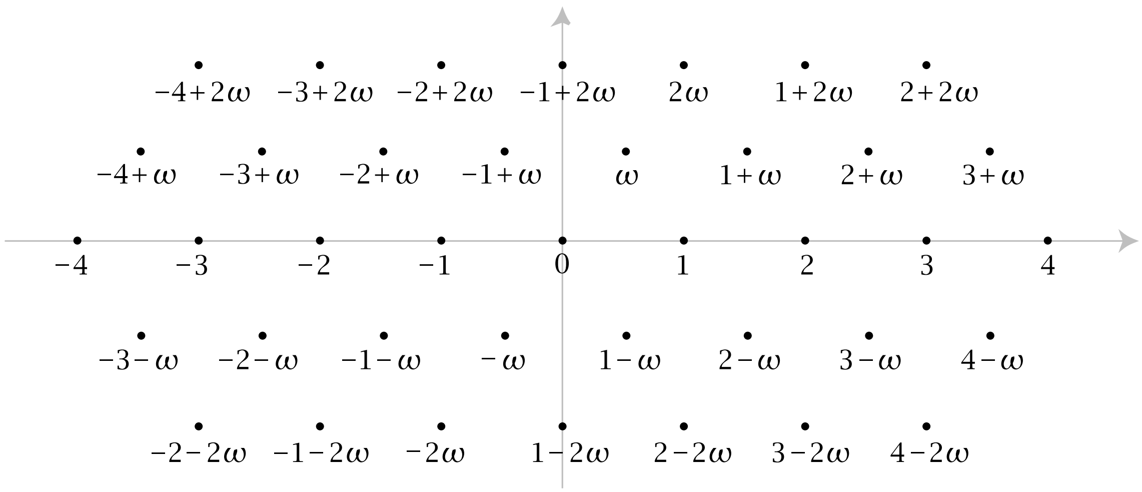
A useful operation with complex numbers is to pass from a number $x + yi$ to its complex conjugate $x - yi$ obtained by reflecting across the x-axis. In $\mathbb{Z}[\sqrt{D}]$ or $\mathbb{Q}(\sqrt{D})$ with $D < 0$ this amounts to replacing $x + y\sqrt{D}$ by its conjugate $x - y\sqrt{D}$. When $D > 0$ we can do exactly the same operation of reflecting $x + y\sqrt{D}$ across the x-axis to the point $x - y\sqrt{D}$, which we again call the conjugate of $x + y\sqrt{D}$.

The ring $\mathbb{Z}[\sqrt{D}]$ is useful for factoring the form $x^2 - Dy^2$ as $(x + y\sqrt{D})(x - y\sqrt{D})$. For this form the discriminant $\Delta = 4D$ is $0 \bmod 4$, and it would be nice to treat also the discriminants $\Delta = 4d + 1 \equiv 1 \bmod 4$, when the principal form is $x^2 + xy - dy^2$. This factors in the following way:

$$x^2 + xy - dy^2 = \Big(x + \frac{1 + \sqrt{1 + 4d}}{2}y\Big)\Big(x + \frac{1 - \sqrt{1 + 4d}}{2}y\Big)$$

To simplify the notation we let $\omega = (1+\sqrt{1+4d})/2$ and $\overline{\omega} = 1-\sqrt{1+4d})/2$, the conjugate of ω, so the factorization becomes $x^2+xy-dy^2 = (x+\omega y)(x+\overline{\omega}y)$. The quadratic equation satisfied by ω is $\omega^2-\omega-d=0$. Thus $\omega^2=\omega+d$ and this allows the product of two numbers of the form $m+n\omega$ to be written in the same form. In other words, the set $\mathbb{Z}[\omega]$ of all numbers $x+y\omega$ with x and y integers is a ring, just like $\mathbb{Z}[\sqrt{D}]$. Note that $\overline{\omega}$ is an element of $\mathbb{Z}[\omega]$ since $\omega+\overline{\omega}=1$, so $\overline{\omega}=1-\omega$

For example, when $d=-1$ we have $\omega=(1+\sqrt{-3})/2$ and the elements of $\mathbb{Z}[\omega]$ form a grid of equilateral triangles in the xy-plane:



The picture for larger negative values of d is similar but stretched in the vertical direction, forming a grid of isosceles triangles. For positive values of d we can do the same thing we did before with $\mathbb{Z}[\sqrt{D}]$ and regard $\mathbb{Z}[\omega]$ as a grid in the plane. For example, for $d=1$ we have $\omega=(1+\sqrt{5})/2$ and $\mathbb{Z}[\omega]$ looks like the picture for $d=-1$ stretched in the vertical direction so that the y-coordinate of ω is $\sqrt{5}/2$ rather than $\sqrt{3}/2$.

Elements of $\mathbb{Z}[\omega]$ can always be written in the form $m+n\omega=(a+b\sqrt{1+4d})/2$ for suitable integers a and b. Here a and b must have the same parity since this is true for $\omega=(1+\sqrt{1+4d})/2$ and hence for any integer multiple $n\omega$, and then adding an arbitrary integer m to $n\omega$ preserves the equal parity condition since it adds an even integer to a. Conversely, if two integers a and b have the same parity then $(a+b\sqrt{1+4d})/2$ lies in $\mathbb{Z}[\omega]$ since by adding or subtracting a suitable even integer from a we can reduce to the case $a=b$ when one has a multiple of ω. Notice that having both a and b even is equivalent to $(a+b\sqrt{1+4d})/2$ lying in $\mathbb{Z}[\sqrt{1+4d}]$, so $\mathbb{Z}[\sqrt{1+4d}]$ is a subset of $\mathbb{Z}[\omega]$. In the figure above we can see that $\mathbb{Z}[\sqrt{1+4d}]$ consists of the even rows, the numbers $m+n\omega$ with n even.

To have a unified notation for both the cases $\mathbb{Z}[\sqrt{D}]$ and $\mathbb{Z}[\omega]$ let us define R_Δ to be $\mathbb{Z}[\sqrt{D}]$ when the discriminant Δ is $4D$ and $\mathbb{Z}[\omega]$ when Δ is $4d+1$. We will often write elements of R_Δ using lower case Greek letters, for example $\alpha=x+y\sqrt{D}$ in $\mathbb{Z}[\sqrt{D}]$ with conjugate $\overline{\alpha}=x-y\sqrt{D}$, or $\alpha=x+y\omega$ in $\mathbb{Z}[\omega]$ with conjugate

$\overline{\alpha} = x + y\overline{\omega} = x + y(1-\omega) = (x+y) - y\omega$.

The main theme of this section and the next will be how elements of R_Δ factor into "primes" within R_Δ. For example, if a prime p in $\mathbb{Z}$ happens to be representable as $p = x^2 - Dy^2$ then this is saying that p is no longer prime in $\mathbb{Z}[\sqrt{D}]$ since it factors as $p = (x+y\sqrt{D})(x-y\sqrt{D}) = \alpha\overline{\alpha}$ for $\alpha = x+y\sqrt{D}$ and $\overline{\alpha} = x - y\sqrt{D}$. Of course, we should say precisely what we mean by a "prime" in $\mathbb{Z}[\sqrt{D}]$ or $\mathbb{Z}[\omega]$. For an ordinary integer $p > 1$, being prime means that p is divisible only by itself and 1. If we allow negative numbers, we can "factor" a prime p as $(-1)(-p)$, but this should not count as a genuine factorization, otherwise there would be no primes at all in $\mathbb{Z}$. In R_Δ things can be a little more complicated because of the existence of ***units*** in R_Δ, the nonzero elements ε in R_Δ whose inverse ε^{-1} also lies in R_Δ. For example, in the Gaussian integers $\mathbb{Z}[i]$ there are four obvious units, ± 1 and $\pm i$, where for $\pm i$ we have $(i)(-i) = 1$ so $i^{-1} = -i$ and $(-i)^{-1} = i$. We will see in a little while that these are the only units in $\mathbb{Z}[i]$. Having four units in $\mathbb{Z}[i]$ instead of just ± 1 complicates the factorization issue slightly, but not excessively so.

For positive values of Δ things are somewhat less tidy because there are always infinitely many units in R_Δ. For example, in $\mathbb{Z}[\sqrt{2}]$ the number $\varepsilon = 3 + 2\sqrt{2}$ is a unit because $(3+2\sqrt{2})(3-2\sqrt{2}) = 1$. All the powers of $3+2\sqrt{2}$ are therefore also units, and there are infinitely many of them since $3 + 2\sqrt{2} > 1$ so the powers $(3+2\sqrt{2})^n$ form an increasing infinite sequence approaching $+\infty$. Their inverses $(3+2\sqrt{2})^{-n} = (3-2\sqrt{2})^n$ are a decreasing infinite sequence approaching 0.

Whenever ε is a unit in R_Δ we can factor any number α in R_Δ as $\alpha = (\alpha\varepsilon)(\varepsilon^{-1})$. If we allowed this as a genuine factorization there would be no primes in R_Δ, so it is best not to consider it as a genuine factorization. This leads us to the following definition: An element α of R_Δ is said to be ***prime*** in R_Δ if it is neither 0 nor a unit, and if whenever we have a factorization of α as $\alpha = \beta\gamma$ with both β and γ in R_Δ, then it must be the case that either β or γ is a unit in R_Δ. Not allowing units as primes is analogous to the standard practice of not considering 1 to be a prime in $\mathbb{Z}$.

If we replace R_Δ by $\mathbb{Z}$ in the definition of primeness above, we get the condition that an integer a in $\mathbb{Z}$ is prime if its only factorizations are the trivial ones $a = (a)(1) = (1)(a)$ and $a = (-a)(-1) = (-1)(-a)$, which is what we would expect. This definition of primeness also means that we are allowing negative primes as the negatives of the positive primes in $\mathbb{Z}$.

A word of caution: An integer p in $\mathbb{Z}$ can be prime in $\mathbb{Z}$ but not prime in R_Δ. For example, in $\mathbb{Z}[i]$ we have the factorization $5 = (2+i)(2-i)$, and as we will be able to verify soon, neither $2+i$ nor $2-i$ is a unit in $\mathbb{Z}[i]$. Hence by our definition 5 is not a prime in $\mathbb{Z}[i]$, even though it is prime in $\mathbb{Z}$. Thus one always has to be careful when speaking about primeness to distinguish "prime in $\mathbb{Z}$" from "prime in R_Δ".

Having defined what we mean by primes in R_Δ it is then natural to ask whether every nonzero element of R_Δ that is not a unit can be factored as a product of primes, and if so, whether this factorization is in any way unique. As we will see, the existence of prime factorizations is fairly easy to prove, but the uniqueness question is much more difficult and subtle. To clarify what the uniqueness question means, notice first that if we have a unit ε in R_Δ we can always modify a factorization $\alpha = \beta\gamma$ to give another factorization $\alpha = (\varepsilon\beta)(\varepsilon^{-1}\gamma)$. This is analogous to writing $6 = (2)(3) = (-2)(-3)$ in $\mathbb{Z}$. This sort of nonuniqueness is unavoidable, but it is also not too serious a problem. So when we speak of factorization in R_Δ being unique, we will always mean unique up to multiplying the factors by units.

A fruitful way to study factorizations in R_Δ is to relate them to factorizations in $\mathbb{Z}$ by associating to each element α in R_Δ the number $N(\alpha) = \alpha\overline{\alpha}$ called the ***norm*** of α. Thus in the two cases $R_\Delta = \mathbb{Z}[\sqrt{D}]$ and $R_\Delta = \mathbb{Z}[\omega]$ we have:

$$N(x + y\sqrt{D}) = (x + y\sqrt{D})(x - y\sqrt{D}) = x^2 - Dy^2$$
$$N(x + y\omega) = (x + y\omega)(x + y\overline{\omega}) = x^2 + xy - dy^2$$

In both cases $N(\alpha)$ is an integer. Notice that when the discriminant is negative, so α is a complex number $a + bi$ for real numbers a and b, the norm of α is just $\alpha\overline{\alpha} = (a + bi)(a - bi) = a^2 + b^2$, the square of the distance from α to the origin in the complex plane. When the discriminant is positive the norm can be negative so it does not have a nice geometric interpretation in terms of distance, but it will be quite useful in spite of this.

The reason the norm is useful for studying factorizations is that it satisfies the following multiplicative property:

Proposition 8.1. $N(\alpha\beta) = N(\alpha)N(\beta)$ *for all* α *and* β *in* R_Δ.

Proof: We will deduce multiplicativity of the norm from multiplicativity of the conjugation operation, the fact that $\overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$. The argument will apply more generally to all elements of $\mathbb{Q}(\sqrt{D})$ for any integer D that is not a square. To verify that $\overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$, write $\alpha = x + y\sqrt{D}$ and $\beta = z + w\sqrt{D}$, so that $\alpha\beta = (xz + ywD) + (xw + yz)\sqrt{D}$. Then:

$$\overline{\alpha\beta} = (xz + ywD) - (xw + yz)\sqrt{D} = (x - y\sqrt{D})(z - w\sqrt{D}) = \overline{\alpha}\,\overline{\beta}$$

For the norm we then have $N(\alpha\beta) = (\alpha\beta)(\overline{\alpha\beta}) = \alpha\beta\overline{\alpha}\,\overline{\beta} = \alpha\overline{\alpha}\beta\overline{\beta} = N(\alpha)N(\beta)$. □

Using the multiplicative property of the norm we can derive a simple criterion for recognizing units:

Proposition 8.2. *An element* $\varepsilon \in R_\Delta$ *is a unit if and only if* $N(\varepsilon) = \pm 1$.

Proof: Suppose ε is a unit, so its inverse ε^{-1} also lies in R_Δ. Then $N(\varepsilon)N(\varepsilon^{-1}) = N(\varepsilon\varepsilon^{-1}) = N(1) = 1$. Since $N(\varepsilon)$ and $N(\varepsilon^{-1})$ are integers this forces $N(\varepsilon)$ to be ± 1. For the converse we use the fact that a nonzero element α in R_Δ has inverse $\alpha^{-1} = \overline{\alpha}/N(\alpha)$ since $\alpha(\overline{\alpha}/N(\alpha)) = 1$. Hence if $N(\varepsilon) = \pm 1$ we have $\varepsilon^{-1} = \pm\overline{\varepsilon}$ which is an element of R_Δ if ε is, so ε is a unit. □

When Δ is negative there are very few units in R_Δ. In the case of $\mathbb{Z}[\sqrt{D}]$ the equation $N(x+y\sqrt{D}) = x^2 - Dy^2 = \pm 1$ has very few integer solutions when $D < 0$, namely, if $D = -1$ there are only the four solutions $(x,y) = (\pm 1, 0)$ and $(0, \pm 1)$ while if $D < -1$ there are only the two solutions $(x,y) = (\pm 1, 0)$. Thus the only units in $\mathbb{Z}[i]$ are ± 1 and $\pm i$, and the only units in $\mathbb{Z}[\sqrt{D}]$ for $D < -1$ are ± 1. Geometrically this is saying that these are the only points in the grid $\mathbb{Z}[\sqrt{D}]$ of distance 1 from the origin, which is obviously true. In the case of $\mathbb{Z}[\omega]$ one can see from the earlier figure of $\mathbb{Z}[\omega]$ that there are just six points of $\mathbb{Z}[\omega]$ of distance 1 from the origin when $d = -1$, and only the two points ± 1 when $d < -1$ and the figure is stretched vertically. When $d = -1$ the six units are ± 1, $\pm\omega$, and $\pm(\omega - 1)$. These are the powers ω^n for $n = 1,2,3,4,5,6$ since the general formula $\omega^2 = \omega + d$ gives $\omega^2 = \omega - 1$ when $d = -1$, and from this it follows that $\omega^3 = -1$, $\omega^4 = -\omega$, $\omega^5 = 1 - \omega$, and $\omega^6 = 1$. When $d < -1$ the only units in $\mathbb{Z}[\omega]$ are ± 1.

The situation for R_Δ with Δ positive is quite different. For $\mathbb{Z}[\sqrt{D}]$ we are looking for solutions of $x^2 - Dy^2 = \pm 1$ with $D > 0$, while for $\mathbb{Z}[\omega]$ the corresponding equation is $x^2 + xy - dy^2 = \pm 1$ with $d > 0$. We know from our study of topographs of hyperbolic forms that these equations have infinitely many integer solutions since the value 1 occurs along the periodic separator line in the topograph of the principal form when $(x,y) = (1,0)$, so it appears infinitely often by periodicity. For some values of D or d the number -1 also appears along the separator line, and then it too appears infinitely often. Thus when $\Delta > 0$ the ring R_Δ has infinitely many units $\varepsilon = x + y\sqrt{D}$ or $x + y\omega$, with arbitrarily large values of x and y.

There is a nice interpretation of units in R_Δ as symmetries of the topograph of the principal form of discriminant Δ. A unit ε in R_Δ defines a transformation T_ε of R_Δ by the formula $T_\varepsilon(\alpha) = \varepsilon\alpha$. In the case of $\mathbb{Z}[\sqrt{D}]$, if $\varepsilon = p + q\sqrt{D}$ then

$$T_\varepsilon(x + y\sqrt{D}) = (p + q\sqrt{D})(x + y\sqrt{D}) = (px + Dqy) + (qx + py)\sqrt{D}$$

while for $\mathbb{Z}[\omega]$, if $\varepsilon = p + q\omega$ we have

$$\begin{aligned} T_\varepsilon(x + y\omega) = (p + q\omega)(x + y\omega) &= (px + qy\omega^2) + (qx + py)\omega \\ &= (px + dqy) + (qx + (p+q)y)\omega \end{aligned}$$

since $\omega^2 = \omega + d$. In both cases we see that T_ε is a linear transformation of x and y, with matrix $\begin{pmatrix} p & Dq \\ q & p \end{pmatrix}$ in the first case and $\begin{pmatrix} p & dq \\ q & p+q \end{pmatrix}$ in the second case. The determinants in the two cases are $p^2 - Dq^2$ and $p^2 + pq - dq^2$ which equal $N(\varepsilon)$ and hence are ± 1 since ε is a unit. Thus T_ε defines a linear fractional transformation giving a symmetry of the Farey diagram. Since $N(\varepsilon\alpha) = N(\varepsilon)N(\alpha)$ we see that T_ε is an orientation-preserving symmetry of the topograph of the norm form when $N(\varepsilon) = +1$ and an orientation-reversing skew symmetry when $N(\varepsilon) = -1$. The symmetry corresponding to the "universal" unit $\varepsilon = -1$ is just the identity since ${}^{-x}/_{-y} = {}^{x}/_{y}$.

When $\Delta < 0$ the only interesting cases are $\Delta = -3$, when T_ε for $\varepsilon = \omega$ is a 120 degree rotation of the topograph, and $\Delta = -4$ when T_ε for $\varepsilon = i$ rotates the topograph by 180 degrees.

When $\Delta > 0$ there is a ***fundamental unit*** ε corresponding to the ± 1 in the topograph of the norm form at the vertex p/q with smallest positive values of p and q. When $N(\varepsilon) = +1$ the transformation T_ε is then the translation giving the periodicity along the separating line since it is an orientation-preserving symmetry. In the opposite case $N(\varepsilon) = -1$ the transformation T_ε is an orientation-reversing skew symmetry so it must be a glide reflection along the separator line by half a period.

Proposition 8.3. *When $\Delta > 0$ the units in R_Δ are exactly the elements $\pm\varepsilon^n$ for $n \in \mathbb{Z}$, where ε is the fundamental unit.*

Proof: The units appear along the separator line at the regions x/y where the norm form takes the value ± 1. From our previous comments, these are the points $T_\varepsilon^n(1/0)$ as n varies over $\mathbb{Z}$. Since T_ε is multiplication by ε, the power T_ε^n is multiplication by ε^n. Thus the values ± 1 occur at the regions labeled x/y for $\varepsilon^n = x + y\sqrt{D}$ or $\varepsilon^n = x + y\omega$. The units are therefore the elements $\pm\varepsilon^n$ where the $\pm$ comes from the fact that the topograph does not distinguish between (x, y) and $(-x, -y)$. $\square$

The conjugation operation in R_Δ sending α to $\overline{\alpha}$ also gives a symmetry of the topograph of the norm form since $N(\alpha) = N(\overline{\alpha})$. Conjugation in $\mathbb{Z}[\sqrt{D}]$ sends an element $x + y\sqrt{D}$ to $x - y\sqrt{D}$ so in the Farey diagram it is reflection across the edge joining $1/0$ and $0/1$. Conjugation in $\mathbb{Z}[\omega]$ sends $x + y\omega$ to $x + y\overline{\omega} = (x + y) - y\omega$ since $\overline{\omega} = 1 - \omega$, so conjugation fixes the vertex $1/0$ and interchanges $0/1$ and $-1/1$ by reflecting across the line perpendicular to the edge from $0/1$ to $-1/1$.

Proposition 8.4. *All symmetries and skew symmetries of the topograph of the norm form are obtainable as combinations of conjugation and the transformations T_ε associated to units ε in R_Δ.*

Proof: It will suffice to reduce an arbitrary symmetry or skew symmetry T to the identity by composing with conjugation and transformations T_ε. If T is a skew symmetry we must have $\Delta > 0$ with -1 appearing along the separator line as well as $+1$. Composing T with a glide reflection T_ε then converts T into a symmetry, so we may assume T is a symmetry from now on. If T reverses orientation of the Farey diagram we may compose it with conjugation to reduce further to the case that T preserves orientation. When $\Delta < 0$ the only possibility for T is then the identity except when $\Delta = -4$ and $T = T_\varepsilon$ for $\varepsilon = i$, or when $\Delta = -3$ and $T = T_\varepsilon$ for $\varepsilon = \omega$ or ω^2. If $\Delta > 0$ the only possibility for T is a translation along the separator line, which is T_ε for some unit ε. $\square$

Now we begin to study primes and prime factorizations in R_Δ. First we have a useful fact:

Proposition 8.5. *If the norm $N(\alpha)$ of an element α in R_Δ is prime in $\mathbb{Z}$ then α is prime in R_Δ.*

For example, when we factor 5 as $(2+i)(2-i)$ in $\mathbb{Z}[i]$, this proposition implies that both factors are prime since the norm of each is 5, which is prime in $\mathbb{Z}$.

Proof: Suppose an element α in R_Δ has a factorization $\alpha = \beta\gamma$, hence $N(\alpha) = N(\beta)N(\gamma)$. If $N(\alpha)$ is prime in $\mathbb{Z}$, this forces one of $N(\beta)$ and $N(\gamma)$ to be ± 1, hence one of β and γ is a unit. This means α is a prime since it cannot be 0 or a unit, as its norm is a prime. □

The converse of this proposition is not generally true. For example the number 3 has norm 9, which is not prime in $\mathbb{Z}$, and yet 3 is prime in $\mathbb{Z}[i]$. This is because if we had a factorization $3 = \alpha\beta$ in $\mathbb{Z}[i]$ with neither α nor β a unit, then the equation $N(\alpha)N(\beta) = N(3) = 9$ would imply that $N(\alpha) = \pm 3 = N(\beta)$, but there are no elements of $\mathbb{Z}[i]$ with norm ± 3 since the equation $x^2 + y^2 = \pm 3$ has no integer solutions.

Now we can prove that prime factorizations always exist:

Proposition 8.6. *Every nonzero element of R_Δ that is not a unit can be factored as a product of primes in R_Δ.*

Proof: We argue by induction on $|N(\alpha)|$. Since we are excluding 0 and units, the induction starts with the case $|N(\alpha)| = 2$. In this case α must itself be a prime by the preceding proposition since 2 is prime in $\mathbb{Z}$. For the induction step, if α is a prime there is nothing to prove. If α is not prime, it factors as $\alpha = \beta\gamma$ with neither β nor γ a unit, so $|N(\beta)| > 1$ and $|N(\gamma)| > 1$. Since $N(\alpha) = N(\beta)N(\gamma)$, it follows that $|N(\beta)| < |N(\alpha)|$ and $|N(\gamma)| < |N(\alpha)|$. By induction, both β and γ are products of primes in R_Δ, hence their product α is also a product of primes in R_Δ. □

Let us investigate how to compute a prime factorization by looking at the case of $\mathbb{Z}[i]$. Assuming that factorizations of Gaussian integers into primes are unique (up to units), which we will prove later, here is a procedure for finding the prime factorization of a Gaussian integer $\alpha = a + bi$:

(1) Factor the integer $N(\alpha) = a^2 + b^2$ into primes p_k in $\mathbb{Z}$.
(2) Determine how each p_k factors into primes in $\mathbb{Z}[i]$.
(3) By the uniqueness of prime factorizations, the primes found in step (2) will be factors of either $a + bi$ or $a - bi$ since they are factors of $(a+bi)(a-bi)$, so all that remains is to test which of the prime factors of each p_k are factors of $a + bi$.

To illustrate this with a simple example, let us see how $3 + i$ factors in $\mathbb{Z}[i]$. We have $N(3+i) = (3+i)(3-i) = 10 = 2 \cdot 5$. These two numbers factor as $2 = (i+i)(1-i)$ and $5 = (2+i)(2-i)$. These are prime factorizations in $\mathbb{Z}[i]$ since $N(1 \pm i) = 2$ and

$N(2 \pm i) = 5$, both of which are primes in $\mathbb{Z}$. Now we test whether for example $1+i$ divides $3+i$ by dividing:

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

Since the quotient $2-i$ is a Gaussian integer, we conclude that $1+i$ is a divisor of $3+i$ and we have the factorization $3+i = (1+i)(2-i)$. This is the prime factorization of $3+i$ since we have already noted that both $1+i$ and $2-i$ are primes in $\mathbb{Z}[i]$.

For a more complicated example consider $244+158i$. For a start, this factors as $2(122+79i)$. Since 122 and 79 have no common factors in $\mathbb{Z}$ we cannot go any farther by factoring out ordinary integers. We know that 2 factors as $(1+i)(1-i)$ and these two factors are prime in $\mathbb{Z}[i]$ since their norm is 2. It remains to factor $122+79i$. This has norm $122^2+79^2 = 21125 = 5^3 \cdot 13^2$. Both 5 and 13 happen to factor in $\mathbb{Z}[i]$, namely $5 = (2+i)(2-i)$ and $13 = (3+2i)(3-2i)$, and these are prime factorizations since the norms of $2 \pm i$ and $3 \pm 2i$ are 5 and 13, primes in $\mathbb{Z}$. Thus we have the prime factorization

$$(122+79i)(122-79i) = 5^3 \cdot 13^2 = (2+i)^3(2-i)^3(3+2i)^2(3-2i)^2$$

Now we look at the factors on the right side of this equation to see which ones are factors of $122+79i$. Suppose for example we test whether $2+i$ divides $122+79i$:

$$\frac{122+79i}{2+i} = \frac{(122+79i)(2-i)}{(2+i)(2-i)} = \frac{323+36i}{5}$$

This is not a Gaussian integer, so $2+i$ does not divide $122+79i$. Let us try $2-i$ instead:

$$\frac{122+79i}{2-i} = \frac{(122+79i)(2+i)}{(2-i)(2+i)} = \frac{165+280i}{5} = 33+56i$$

So $2-i$ does divide $122+79i$. In fact, we can expect that $(2-i)^3$ will divide $122+79i$, and it can be checked that it does. In a similar way one can check whether $3+2i$ or $3-2i$ divides $122+79i$, and one finds that it is $3-2i$ that divides $122+79i$, and in fact $(3-2i)^2$ divides $122+79i$. After these calculations one might expect that $122+79i$ was the product $(2-i)^3(3-2i)^2$, but upon multiplying this product out one finds that it is the negative of $122+79i$. Thus:

$$122+79i = (-1)(2-i)^3(3-2i)^2$$

The factor -1 is a unit, so it could be combined with one of the other factors, for example changing one of the factors $2-i$ to $i-2$. Alternatively, we could replace the factor -1 by i^2 and then multiply each $3-2i$ factor by i to get a neater prime factorization:

$$122+79i = (2-i)^3(2+3i)^2$$

Combining these calculations, we have the prime factorization for $244+158i$:

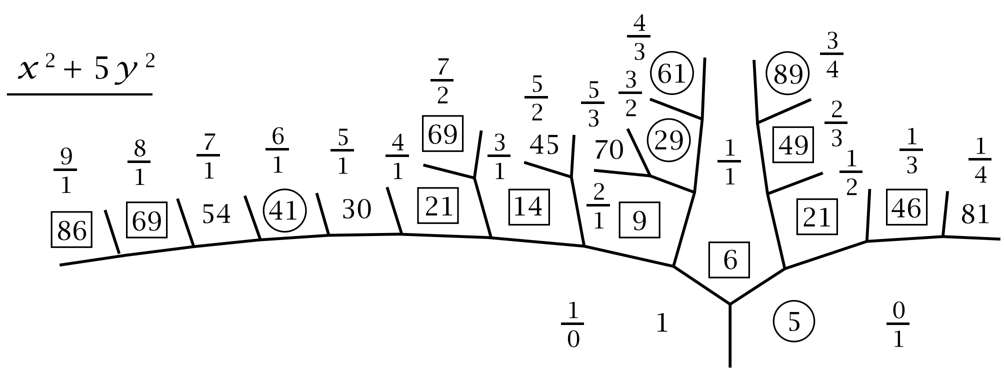
$$244+158i = (1+i)(1-i)(2-i)^3(2+3i)^2$$

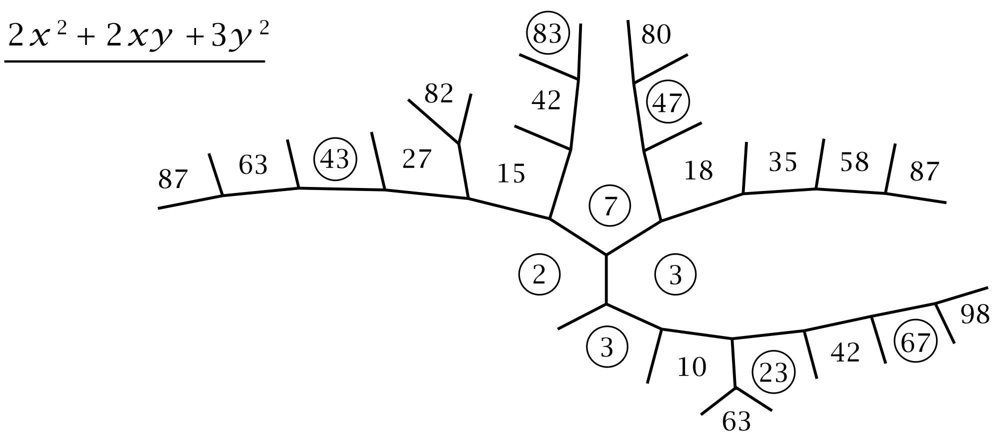
The method in this example for computing prime factorizations in $\mathbb{Z}[i]$ depended on unique factorization. When unique factorization fails, things are more complicated. One of the simplest instances of this is in $\mathbb{Z}[\sqrt{-5}]$ where we have the factorizations:

$$6 = (2)(3) = \big(1+\sqrt{-5}\big)\big(1-\sqrt{-5}\big)$$

The only units in $\mathbb{Z}[\sqrt{-5}]$ are ± 1, so these two factorizations do not differ just by units. We can see that 2, 3, and $1 \pm \sqrt{-5}$ are prime in $\mathbb{Z}[\sqrt{-5}]$ by looking at norms. Using the formula $N(x+y\sqrt{-5}) = x^2+5y^2$ we see that the norms of 2, 3, and $1\pm\sqrt{-5}$ are 4, 9, and 6, so if one of 2, 3, or $1\pm\sqrt{-5}$ was not a prime, it would have a factor of norm 2 or 3 since these are the only numbers that occur in nontrivial factorizations of 4, 9, and 6 in $\mathbb{Z}$. However, the equations $x^2+5y^2 = 2$ and $x^2+5y^2 = 3$ obviously have no integer solutions so there are no elements of $\mathbb{Z}[\sqrt{-5}]$ of norm 2 or 3. Thus in $\mathbb{Z}[\sqrt{-5}]$ the number 6 has two prime factorizations that do not differ just by units.

This example can be explained by the fact that x^2+5y^2 is not the only quadratic form of discriminant -20, up to equivalence. Another form of the same discriminant is $2x^2+2xy+3y^2$, and this form takes on the values 2 and 3 that the form x^2+5y^2 omits, even though x^2+5y^2 does take on the value $6 = 2\cdot 3$. Here are the topographs of these two forms, with prime values circled:





The appearance of 6 in the topograph of $x^2+5y^2 = (x+y\sqrt{-5})(x-y\sqrt{-5})$ when $x/y = 1/1$ gives the factorization $6 = (1+\sqrt{-5})(1-\sqrt{-5})$.

The boxed nonprime numbers in the topograph of x^2+5y^2 give rise to other nonunique prime factorizations. For example $14 = (2)(7) = (3+\sqrt{-5})(3-\sqrt{-5})$

where the second factorization comes from the appearance of 14 in the topograph of $x^2 + 5y^2$ when $x/y = 3/1$. As with the earlier factorizations of 6, the nonappearance of 2 and 7 in the topograph of $x^2 + 5y^2$ implies that 2, 7, and $3 \pm \sqrt{-5}$ are prime in $\mathbb{Z}[\sqrt{-5}]$. Some numbers in the topograph of $x^2 + 5y^2$ occur in boxes twice, leading to three different prime factorizations. Thus 21 factors into primes in $\mathbb{Z}[\sqrt{-5}]$ as $3 \cdot 7$, as $(1 + 2\sqrt{-5})(1 - 2\sqrt{-5})$ and as $(4 + \sqrt{-5})(4 - \sqrt{-5})$. Another example is $69 = 3 \cdot 23 = (7 + 2\sqrt{-5})(7 - 2\sqrt{-5}) = (8 + \sqrt{-5})(8 - \sqrt{-5})$.

The numbers that appear in the topograph of the second form $2x^2 + 2xy + 3y^2$ are not the norms of elements of $\mathbb{Z}[\sqrt{-5}]$ but one might imagine that they are the norms of "ideal numbers" of some sort. Thus 2 might be the norm of an ideal number P, so $2 = P\overline{P}$, and 3 might be the norm of an ideal number Q, so $3 = Q\overline{Q}$. The product PQ would then have norm $(PQ)(\overline{PQ}) = (P\overline{P})(Q\overline{Q}) = 2 \cdot 3 = 6$, so it is possible that $PQ = 1 + \sqrt{-5}$. If this is true, it would explain very nicely the two factorizations of 6 as $2 \cdot 3 = (P\overline{P})(Q\overline{Q})$ and as $(1 + \sqrt{-5})(1 - \sqrt{-5}) = (PQ)(\overline{PQ})$.

One can also see how some numbers might have three different prime factorizations. For example for $21 = 3 \cdot 7$, if we have $3 = P\overline{P}$ and $7 = Q\overline{Q}$ then there are three ways to group these four ideal numbers into pairs, as $(P\overline{P})(Q\overline{Q})$, as $(PQ)(\overline{P}\,\overline{Q})$, and as $(P\overline{Q})(\overline{P}Q)$, and these three groupings could give the three factorizations of 21. The reason there are only two factorizations for $2 \cdot 3$ and $2 \cdot 7$ is that in the factorization $2 = P\overline{P}$ the two factors P and $\overline{P}$ happen to be equal, so there is no difference between $(PQ)(\overline{P}\,\overline{Q})$ and $(P\overline{Q})(\overline{P}Q)$.

Much of this chapter will be devoted to making sense of these "ideal numbers". They will be realized not by actual numbers but by certain sets of numbers in R_Δ called simply "ideals". These ideals behave like actual numbers in some respects. Most importantly they can be multiplied and they have norms which are ordinary integers, behaving much like norms of elements of R_Δ. On the other hand there is no method for adding ideals that behaves like addition of numbers, so ideals are not entirely like numbers. However, this will not matter for studying prime factorizations where multiplication is what one is interested in.

There is a natural notion of what a prime ideal is, and the big theorem about ideals in R_Δ is that they have unique factorizations into prime ideals when Δ is a fundamental discriminant. This gives information about prime factorizations of elements of R_Δ because each element of R_Δ gives rise to a special kind of ideal called a principal ideal. For some discriminants all ideals are principal ideals, and for these discriminants the unique prime factorization of ideals translates into unique prime factorization of elements of R_Δ.

In the remainder of this section and continuing in the next we will go further into prime factorizations of elements of R_Δ before beginning the study of ideals in Section 8.3.

The question of how a prime p in $\mathbb{Z}$ factors in R_Δ can be rephrased in terms of the norm form $x^2 - Dy^2$ or $x^2 + xy - dy^2$, according to the following result:

Proposition 8.7. *Let p be a prime in $\mathbb{Z}$. Then:*

(a) *If either p or $-p$ is represented by the norm form for R_Δ, so $N(\alpha) = \pm p$ for some α in R_Δ, then p factors in R_Δ as $p = \pm\alpha\overline{\alpha}$ and both these factors are prime in R_Δ.*

(b) *If neither p nor $-p$ is represented by the norm form then p remains prime in R_Δ.*

In statement (a) note that when $\Delta < 0$ the norm only takes positive values, so if a positive prime p factors in R_Δ it must factor as $p = \alpha\overline{\alpha}$, never as $-\alpha\overline{\alpha}$. However for $\Delta > 0$ this need not be the case. For example for $\mathbb{Z}[\sqrt{3}]$ the topograph of $x^2 - 3y^2$ shown in Section 4.1 contains the value -2 but not 2, so the prime 2 factors as $-(1 + \sqrt{3})(1 - \sqrt{3})$ in $\mathbb{Z}[\sqrt{3}]$ but not as $\alpha\overline{\alpha}$.

Proof: For part (a), if $p = \pm N(\alpha)$, then p factors in R_Δ as $p = \pm\alpha\overline{\alpha} = \pm N(\alpha)$. The two factors are prime since their norm is $\pm p$ which is prime in $\mathbb{Z}$ by assumption.

For (b), if p is not a prime in R_Δ then it factors in R_Δ as $p = \alpha\beta$ with neither α nor β a unit. Then $N(p) = p^2 = N(\alpha)N(\beta)$ with neither $N(\alpha)$ nor $N(\beta)$ equal to ± 1, hence we must have $N(\alpha) = N(\beta) = \pm p$. The equation $N(\alpha) = \pm p$ says that the norm form represents $\pm p$. Thus if the norm form represents neither p nor $-p$ then p must be prime in R_Δ. □

Proposition 8.8. *If R_Δ has unique factorization into primes then the only primes in R_Δ are the primes described in* (a) *or* (b) *of the preceding proposition, or units times these primes.*

This can be false without unique prime factorization since the primes in R_Δ obtained by factoring a prime integer p have norm dividing $N(p) = p^2$, but we have seen for example that $1 + \sqrt{-5}$ is prime in $\mathbb{Z}[\sqrt{-5}]$ and has norm 6.

Proof: Let α be an arbitrary prime in R_Δ. The norm $n = N(\alpha) = \alpha\overline{\alpha}$ is an integer in $\mathbb{Z}$ so it can be factored as a product $n = p_1 \cdots p_k$ of primes in $\mathbb{Z}$. By Proposition 8.7 each p_i either stays prime in R_Δ or factors as a product $\pm\alpha_i\overline{\alpha}_i$ of two primes in R_Δ. This gives a factorization of n into primes in R_Δ. A second factorization of n into primes in R_Δ can be obtained from the formula $n = \alpha\overline{\alpha}$ by factoring $\overline{\alpha}$ into primes since the first factor α is already prime by assumption. (In fact if α is prime then $\overline{\alpha}$ will also be prime, but we do not need to know this.) If we have unique factorization in R_Δ then the prime factor α of the second prime factorization will have to be one of the prime factors in the first prime factorization of n, or a unit times one of these primes. Thus α will be a unit times a prime of one of the two types described in Proposition 8.7. □

There are two qualitatively different ways in which a prime p in $\mathbb{Z}$ can factor as the product of two primes in R_Δ, depending on whether the two primes in R_Δ differ by just a unit, or equivalently, whether p is a unit times the square of an element of R_Δ. For example in $\mathbb{Z}[i]$ we have $2 = (1+i)(1-i)$ and the two factors $1+i$ and $1-i$ differ only by a unit since $-i(1+i) = 1-i$. Thus $2 = \varepsilon(1+i)^2$ for the unit $\varepsilon = -i$. In fact 2 is the only prime that can be factored in $\mathbb{Z}[i]$ as $p = \varepsilon(a+bi)^2$ for some unit ε. The units in $\mathbb{Z}[i]$ are ± 1 and $\pm i$ so the only possibilities are $p = \pm(a+bi)^2$ and $p = \pm i(a+bi)^2$. In the first case $p = \pm(a+bi)^2 = \pm(a^2-b^2+2abi)$ so $2ab = 0$ hence either a or b is 0, but that would say $p = \pm a^2$ or $p = \pm b^2$ which is impossible since p is prime. The other case is $p = \pm i(a+bi)^2 = \pm((a^2-b^2)i - 2ab)$ hence $p = \pm 2ab$ so a and b are ± 1 and $p = 2$.

Exercises

1. (a) Show that if α and β are elements of $\mathbb{Z}[\sqrt{D}]$ such that α is a unit times β, then $N(\alpha) = \pm N(\beta)$.

(b) Either prove or give a counterexample to the following statement: If α and β are Gaussian integers with $N(\alpha) = N(\beta)$ then α is a unit times β.

2. Show that a Gaussian integer $x + yi$ with both x and y odd is divisible by $1+i$ but not by $(1+i)^2$.

3. There are four different ways to write the number $1105 = 5\cdot 13\cdot 17$ as a sum of two squares. Find these four ways using the factorization of 1105 into primes in $\mathbb{Z}[i]$. Here we are not counting 5^2+2^2 and 2^2+5^2 as different ways of expressing 29 as the sum of two squares. Note that an equation $n = a^2 + b^2$ is equivalent to an equation $n = (a+bi)(a-bi)$.

4. (a) Find four different units in $\mathbb{Z}[\sqrt{3}]$ that are positive real numbers, and find four that are negative.

(b) Do the same for $\mathbb{Z}[\sqrt{11}]$.

5. Make a list of all the Gaussian primes $x + yi$ with $-7 \le x \le 7$ and $-7 \le y \le 7$. (The only actual work here is to figure out the primes $x + yi$ with $0 \le y \le x \le 7$ since the rest are obtainable from these by symmetry properties.)

6. Factor the following Gaussian integers into primes in $\mathbb{Z}[i]$: $3+5i$, $8-i$, $10+i$, $5-12i$, $35i$, $-35+120i$, $253+204i$.

7. (a) Show that if an odd prime p factors in $\mathbb{Z}[\omega]$ for $\omega = (1+\sqrt{-3})/2$ then it factors in $\mathbb{Z}[\sqrt{-3}]$.

(b) Do the same with -3 replaced by -7.

(c) Show that this no longer holds when -3 is replaced by -11.

8. (a) Determine how the number 2 factors into primes in R_Δ for $\Delta = -3, -4, -7, -8, -11, -12, -15$, and -16.
(b) Do the same for $\Delta = 5, 8$, and 12.

9. Show that if an element α in R_Δ is prime then so is $\overline{\alpha}$.

10. (a) Find a number n that has exactly two different factorizations into primes in $\mathbb{Z}[\sqrt{-6}]$, up to multiplication by units, and find another number that has exactly three such factorizations.
(b) Do the same for $\mathbb{Z}[\sqrt{10}]$ where things are slightly more complicated since there are many more units.

11. Show that the factorization of a prime p in $\mathbb{Z}$ into primes in R_Δ is always unique up to units. (See Propositions 6.16 and 8.4.)

8.2 Unique Factorization via the Euclidean Algorithm

The main goal in this section is to show that unique factorization holds for the Gaussian integers $\mathbb{Z}[i]$ and in a few other cases as well. The plan will be to see that Gaussian integers have a Euclidean algorithm much like the Euclidean algorithm in $\mathbb{Z}$, then deduce unique factorization from this Euclidean algorithm.

In order to prove that prime factorizations are unique we will use the following special property that holds in $\mathbb{Z}$ and in some of the rings R_Δ as well:

$(*)$ *If a prime p divides a product ab then p must divide either a or b.*

One way to prove this for $\mathbb{Z}$ would be to consider the prime factorization of ab, which can be obtained by factoring each of a and b into primes separately. Then if the prime p divides ab, it would have to occur in the prime factorization of ab, hence it would occur in the prime factorization of either a or b, which would say that p divides a or b.

This argument assumed implicitly that the prime factorization of ab was unique. Thus the property $(*)$ is a consequence of unique factorization into primes. But the property $(*)$ also implies that prime factorizations are unique. To see why, consider two factorizations of a number $n > 1$ into positive primes:

$$n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$$

We can assume $k \le l$ by interchanging the p_i's and q_i's if necessary. We want to argue that if $(*)$ holds for each p_i, then the q_i's are just a permutation of the p_i's and in particular $k = l$. The argument to prove this goes as follows. Consider first the prime p_1. This divides the product $q_1(q_2 \cdots q_l)$ so by property $(*)$ it divides either q_1 or $q_2 q_3 \cdots q_l$. In the latter case, another application of $(*)$ shows that p_1

divides either q_2 or $q_3q_4\cdots q_l$. Repeating this argument as often as necessary, we conclude that p_1 must divide at least one q_i. After permuting the q_i's we can assume that p_1 divides q_1. We are assuming all the p_i's and q_i's are positive, so the fact that the prime p_1 divides the prime q_1 implies that p_1 equals q_1. We can then cancel p_1 and q_1 from the equation $p_1p_2\cdots p_k = q_1q_2\cdots q_l$ to get $p_2\cdots p_k = q_2\cdots q_l$. Now repeat the argument to show that p_2 equals some remaining q_i which we can assume is q_2 after a permutation. After further repetitions we eventually reach the point that the final p_k is a product of the remaining q_i's. But then since p_k is prime there could only be one remaining q_i, so we would have $k = l$ and $p_k = q_k$, finishing the argument.

If we knew the analogue of property $(*)$ held for primes in R_Δ we could make essentially the same argument to show that unique factorization holds in R_Δ. The only difference in the argument would be that we would have to take units into account. The argument would be exactly the same up to the point where we concluded that p_1 divides q_1. Then the fact that q_1 is prime would not say that p_1 and q_1 were equal, but only that q_1 is a unit times p_1, so we would have an equation $q_1 = ep_1$ with e a unit. Then we would have $p_1p_2\cdots p_k = ep_1q_2\cdots q_l$. Canceling p_1 would then yield $p_2p_3\cdots p_k = eq_2q_3\cdots q_l$. The product eq_2 is prime if q_2 is prime, so if we let $q_2' = eq_2$ we would have $p_2p_3\cdots p_k = q_2'q_3\cdots q_l$. The argument could then be repeated to show eventually that the q_i's are the same as the p_i's up to permutation and multiplication by units, which is what unique factorization means.

Since the property $(*)$ implies unique factorization, it will not hold in R_Δ when R_Δ does not have unique factorization. For a concrete example consider $\mathbb{Z}[\sqrt{-5}]$. Here we had nonunique prime factorizations $6 = 2\cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5})$. The prime 2 thus divides the product $(1+\sqrt{-5})(1-\sqrt{-5})$ but it does not divide either factor $1\pm\sqrt{-5}$ since $(1\pm\sqrt{-5})/2$ is not an element of $\mathbb{Z}[\sqrt{-5}]$.

Our task now is to prove the property $(*)$ without using unique factorization. As we saw in Chapter 2, an equation $ax+by = 1$ always has integer solutions (x,y) whenever a and b are coprime integers. This fact can be used to show that property $(*)$ holds in $\mathbb{Z}$. To see how, suppose that a prime p divides a product ab. It will suffice to show that if p does not divide a then it must divide b. If p does not divide a, then since p is prime, p and a are coprime. This implies that the equation $px + ay = 1$ is solvable with integers x and y. Now multiply this equation by b to get an equation $b = pbx + aby$. The number p divides the right side of this equation since it obviously divides pbx and it divides ab by assumption. Hence p divides b, which is what we wanted to show.

The fact that equations $ax + by = 1$ in $\mathbb{Z}$ are solvable whenever a and b are coprime can be deduced from the Euclidean algorithm in the following way. What the Euclidean algorithm gives is a method for starting with two positive integers α_0 and

α_1 and constructing a sequence of positive integers α_i and β_i satisfying the following equations:

$$\begin{aligned}
\alpha_0 &= \beta_1\alpha_1 + \alpha_2 \\
\alpha_1 &= \beta_2\alpha_2 + \alpha_3 \\
&\vdots \\
\alpha_{n-2} &= \beta_{n-1}\alpha_{n-1} + \alpha_n \\
\alpha_{n-1} &= \beta_n\alpha_n + \alpha_{n+1} \\
\alpha_n &= \beta_{n+1}\alpha_{n+1}
\end{aligned}$$

From these equations we can deduce two consequences:

(1) α_{n+1} divides α_0 and α_1.
(2) The equation $\alpha_{n+1} = \alpha_0 x + \alpha_1 y$ is solvable in $\mathbb{Z}$.

To see why (1) is true, note that the last equation implies that α_{n+1} divides α_n. Then the next-to-last equation implies that α_{n+1} also divides α_{n-1}, and the equation before this then implies that α_{n+1} also divides α_{n-2}, and so on until one deduces that α_{n+1} divides all the α_i's and in particular α_0 and α_1.

To see why (2) is true, observe that each equation before the last one allows an α_i to be expressed as a linear combination of α_{i-1} and α_{i-2}. Then by repeatedly substituting in, one can express each α_i in terms of α_0 and α_1 as a linear combination $x\alpha_0 + y\alpha_1$ with integer coefficients x and y. In particular α_{n+1} can be represented in this way, which says that the equation $\alpha_{n+1} = \alpha_0 x + \alpha_1 y$ is solvable in $\mathbb{Z}$.

Now if we assume that α_0 and α_1 are coprime then α_{n+1} must be 1 by (1), and by (2) we get integers x and y such that $\alpha_0 x + \alpha_1 y = 1$, as we wanted.

Putting all the preceding arguments together, we see that the Euclidean algorithm in $\mathbb{Z}$ implies unique factorization in $\mathbb{Z}$.

A very similar argument works in R_Δ provided that one has a Euclidean algorithm to produce the sequence of equations above starting with any pair of nonzero elements α_0 and α_1 in R_Δ, with all the numbers α_i and β_i now being elements of R_Δ. The statements (1) and (2) again follow from these equations, with the equation in (2) now being solvable with x and y in R_Δ. For the application to unique factorization the coefficients α_0 and α_1 will be coprime in the sense that their only common divisors are units, so α_{n+1} will be a unit. A solution of $\alpha_{n+1} = \alpha_0 x + \alpha_1 y$ can then be modified by multiplying x and y by α_{n+1}^{-1} to get a solution of $1 = \alpha_0 x + \alpha_1 y$. By the argument given before, this is all that is needed to imply unique factorization in R_Δ.

Let us show now that there is a Euclidean algorithm in the Gaussian integers $\mathbb{Z}[i]$. The key step is to be able to find, for each pair of nonzero Gaussian integers α_0 and α_1, two more Gaussian integers β_1 and α_2 such that $\alpha_0 = \beta_1\alpha_1 + \alpha_2$ with α_2 being "smaller" than α_1. We measure "smallness" of complex numbers by computing their distance to the origin in the complex plane. For a complex number $\alpha = x + yi$ this distance is $\sqrt{x^2 + y^2}$. Here $x^2 + y^2$ is just the norm $N(\alpha)$ when x and y are integers, so we could measure the size of a Gaussian integer α by $\sqrt{N(\alpha)}$. However it is simpler just to use $N(\alpha)$ without the square root, so this is what we will do.

Thus our goal is to find an equation $\alpha_0 = \beta_1\alpha_1 + \alpha_2$ with $N(\alpha_2) < N(\alpha_1)$, starting from two given nonzero Gaussian integers α_0 and α_1. If we can always do this, then by repeating the process we can construct a sequence of α_i's and β_i's where the successive α_i's have smaller and smaller norms. Since these norms are nonnegative integers, they cannot keep decreasing infinitely often, so eventually the process will reach an α_i of norm 0, hence this α_i will be 0 and the Euclidean algorithm will end in a finite number of steps, as it should.

The equation $\alpha_0 = \beta_1\alpha_1 + \alpha_2$ is saying that when we divide α_1 into α_0, we obtain a quotient β_1 and a remainder α_2. What we want is for the remainder α_2 to have a smaller norm than α_1. To get an idea how we can do this let us look instead at the equivalent equation

$$\frac{\alpha_0}{\alpha_1} = \beta_1 + \frac{\alpha_2}{\alpha_1}$$

If we were working with ordinary integers, the quotient β_1 would be the integer part of the rational number α_0/α_1 and α_2/α_1 would be the remaining fractional part. For Gaussian integers we do something similar, but instead of taking β_1 to be the "integer part" of α_0/α_1 we take it to be a Gaussian integer of *minimum distance* to α_0/α_1.

As an example let us take $\alpha_0 = 12 + 15i$ and $\alpha_1 = 5 + 2i$. Then:

$$\frac{\alpha_0}{\alpha_1} = \frac{12+15i}{5+2i} = \frac{(12+15i)(5-2i)}{(5+2i)(5-2i)} = \frac{90+51i}{29} = (3+2i) + \frac{3-7i}{29} = \beta_1 + \frac{\alpha_2}{\alpha_1}$$

Here in the last step we choose $3 + 2i$ as β_1 because 3 is the closest integer to $^{90}/_{29}$ and 2 is the closest integer to $^{51}/_{29}$. Having found a likely candidate for β_1, we can use the equation $\alpha_0 = \beta_1\alpha_1 + \alpha_2$ to find α_2:

$$\begin{aligned} 12 + 15i &= (3+2i)(5+2i) + \alpha_2 \\ &= (11+16i) + \alpha_2 \quad \text{hence} \quad \alpha_2 = 1 - i \end{aligned}$$

Since $N(1-i) = 2$ and $N(5+2i) = 29$ we have $N(\alpha_2) < N(\alpha_1)$ as we wanted.

In fact choosing β_1 as a closest Gaussian integer to the "Gaussian rational" α_0/α_1 will always lead to an α_2 with $N(\alpha_2) < N(\alpha_1)$. This is because if we write the quotient α_2/α_1 in the form $x + yi$ for rational numbers x and y (so in the example above we have $x = {}^3/_{29}$ and $y = {}^{-7}/_{29}$) then for β_1 to be a Gaussian integer closest to α_0/α_1 means that $|x| \le {}^1/_2$ and $|y| \le {}^1/_2$, and therefore:

$$N\Big(\frac{\alpha_2}{\alpha_1}\Big) = x^2 + y^2 \le {}^1/_4 + {}^1/_4 < 1$$

$$\text{and hence} \qquad N(\alpha_2) = N\Big(\frac{\alpha_2}{\alpha_1}\cdot\alpha_1\Big) = N\Big(\frac{\alpha_2}{\alpha_1}\Big)N(\alpha_1) < N(\alpha_1)$$

Thus we have $N(\alpha_2) < N(\alpha_1)$. This shows that there is a general Euclidean algorithm in $\mathbb{Z}[i]$, and so $\mathbb{Z}[i]$ has unique factorization.

Just as an exercise let us finish carrying out the Euclidean algorithm for $\alpha_0 = 12 + 15i$ and $\alpha_1 = 5 + 2i$. The next step is to divide $\alpha_2 = 1 - i$ into $\alpha_1 = 5 + 2i$:

$$\frac{5+2i}{1-i} = \frac{(5+2i)(1+i)}{(1-i)(1+i)} = \frac{3+7i}{2} = (1+3i) + \frac{1+i}{2}$$

Notice that the fractions $3/2$ and $7/2$ are exactly halfway between two consecutive integers, so instead of choosing $1 + 3i$ for the closest integer to $3+7i/2$ we could equally well have chosen $2+3i$ or $1+4i$ or $2+4i$. Let us stick with the choice $1+3i$ and use this to calculate the next α_i:

$$5 + 2i = (1+3i)(1-i) + \alpha_3 = (4+2i) + \alpha_3$$

hence $\alpha_3 = 1$. The final step would be simply to write $1 - i = (1-i)1 + 0$. Thus the full Euclidean algorithm gives the following equations:

$$\begin{aligned} 12 + 15i &= (3+2i)(5+2i) + (1-i) \\ 5 + 2i &= (1+3i)(1-i) + 1 \\ 1 - i &= (1-i)(1) + 0 \end{aligned}$$

In particular, since the last nonzero remainder is 1, a unit in $\mathbb{Z}[i]$, we deduce that this is the greatest common divisor of $12 + 15i$ and $5 + 2i$, where "greatest" means "of greatest norm". In other words $12 + 15i$ and $5 + 2i$ have no common divisors other than units. (This also follows from the fact that the norms $N(12 + 15i) = 9 \cdot 41$ and $N(5 + 2i) = 29$ are coprime.)

The equations that display the results of carrying out the Euclidean algorithm can be used to express the last nonzero remainder in terms of the original two numbers:

$$\begin{aligned} 1 &= (5+2i) - (1+3i)(1-i) \\ &= (5+2i) - (1+3i)[(12+15i) - (3+2i)(5+2i)] \\ &= -(1+3i)(12+15i) + (-2+11i)(5+2i) \end{aligned}$$

If it had happened that the last nonzero remainder was a unit other than 1, we could have expressed this unit in terms of the original two Gaussian integers, and then multiplied the equation by the inverse of the unit to get an expression for 1 in terms of the original two Gaussian integers.

Having shown that prime factorizations in $\mathbb{Z}[i]$ are unique, let us see what this implies about the representation problem for the norm form $x^2 + y^2$. The equation $x^2 + y^2 = n$ can be written as $(x+yi)(x-yi) = n$. If the prime factorization of $x+yi$ in $\mathbb{Z}[i]$ is $x + yi = \alpha_1 \cdots \alpha_l$ and the prime factorization of n in $\mathbb{Z}$ is $n = p_1 \cdots p_m$ then the equation $x^2 + y^2 = n$ becomes $\alpha_1\overline{\alpha}_1 \cdots \alpha_l\overline{\alpha}_l = p_1 \cdots p_m$. A prime p in $\mathbb{Z}$ either splits as a product $\alpha\overline{\alpha}$ of two primes in $\mathbb{Z}[i]$ or remains prime in $\mathbb{Z}[i]$. Unique prime factorization means that, up to units, the factorization $n = \alpha_1 \cdots \alpha_l$ is obtained from the factorization $n = p_1 \cdots p_m$ by replacing each p_k that splits by a product $\alpha_j\overline{\alpha}_j$. Each prime p_k that does not split will be equal to some α_j or $\overline{\alpha}_j$, but in this case both factors α_j and $\overline{\alpha}_j$ are integers so they are equal. This means that the two factors α_j and $\overline{\alpha}_j$ give two factors of the product $p_1 \cdots p_m$ that are the same nonsplit prime. Thus nonsplit primes must occur to an even power in n. Conversely, if the nonsplit prime factors of n occur only to even powers then we obtain a factorization $n = \alpha_1\overline{\alpha}_1 \cdots \alpha_l\overline{\alpha}_l$ and hence a solution of $x^2 + y^2 = n$ with $x + yi = \alpha_1 \cdots \alpha_l$.

Thus we see that the equation $x^2 + y^2 = n$ has an integer solution (x, y) exactly when each nonsplit prime factor p of n occurs with an even exponent in n. The split primes are the primes represented by $x^2 + y^2$, so these are 2 and primes $p = 4k + 1$ as we saw in Chapter 6. Hence the numbers expressible as the sum of two squares are the numbers in which each prime factor $p = 4k + 3$ occurs to an even power. This agrees with the answer we got in Chapter 6, but the only results from that chapter we have used here are the fact that all primes $p = 4k + 1$ are represented by $x^2 + y^2$ and the easy facts that 2 is represented but primes $p = 4k + 3$ are not represented.

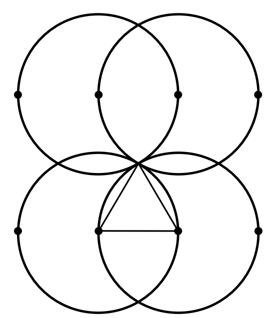
Going further, we can also answer the more subtle question of when the equation $x^2 + y^2 = n$ has a solution with x and y coprime. For x and y not to be coprime means they are both divisible by some prime p, which is the same as saying that $x + yi$ is divisible by p in $\mathbb{Z}[i]$, or we could equally well say $x - yi$ instead of $x + yi$. If a prime factor p of n in $\mathbb{Z}$ does not split in $\mathbb{Z}[i]$ then p will be prime in $\mathbb{Z}[i]$ so in the factorization $n = (x + yi)(x - yi)$ we will have p as a prime factor of $x + yi$ or $x - yi$ in $\mathbb{Z}[i]$, so x and y will not be coprime. Thus n must be a product of primes that split in $\mathbb{Z}[i]$. If one of these primes splits as $p = \alpha\overline{\alpha}$ then we cannot have both α and $\overline{\alpha}$ as two of the factors of $x + yi$, otherwise p would divide $x + yi$. Thus if p appears to the kth power in n, we must have α^k as a factor of $x + yi$ and $\overline{\alpha}^k$ as a factor of $x - yi$ or vice versa, at least when α and $\overline{\alpha}$ do not differ just by a unit. If α and $\overline{\alpha}$ differ just by a unit then we must have $k = 1$, otherwise $x + yi$ would have p as a factor. We noted earlier that 2 is the only prime in $\mathbb{Z}$ that splits as a product of two primes in $\mathbb{Z}[i]$ that differ just by a unit, so the final result is that $x^2 + y^2 = n$ has a solution with x and y coprime exactly when $n = 2^a p_1 \cdots p_k$ with each p_i a prime congruent to 1 mod 4 and $a \le 1$. This too agrees with what we showed in Chapter 6.

An advantage of using Gaussian integers to determine the numbers represented by $x^2 + y^2$ is that this gives a way of computing explicitly all the representations of a given number n. Computing the topograph does this, but the amount of work needed increases rapidly as n gets larger since one is computing the representations of all numbers smaller than n at the same time. To illustrate how Gaussian integers speed things up for specific values of n let us see how to find all the primitive solutions of $x^2 + y^2 = 5^k$. For $k = 1$ we have the solution $(x, y) = (2, 1)$ corresponding to the factorization $5 = (2 + i)(2 - i)$, so a primitive solution for arbitrary k is obtained by expressing $(2 + i)^k$ as $x + yi$. One could use the binomial theorem for this, but this would involve computing binomial coefficients, so let us instead proceed by induction on k using the formula $(x+yi)(2+i) = (2x-y)+(x+2y)i$. This yields the following sequence of pairs (x,y) for $k = 1,2,3,4,5,6,7,8$:

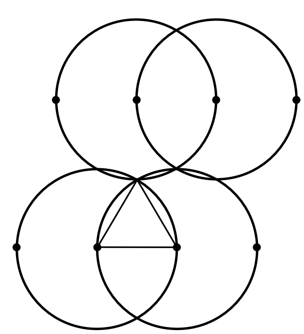
$$(2,1),(3,4),(2,11),(-7,24),(-38,41),(-117,44),(-278,-29),(-527,-336)$$

The signs are irrelevant for solutions of $x^2+y^2 = 5^k$ but they cannot be ignored when computing with the inductive formula. For each k there are exactly eight primitive solutions, corresponding to $(2+i)^k$ and $(2-i)^k$ along with multiples of these by each of the four units $\pm 1, \pm i$. In terms of x and y these are the groups $(\pm x, \pm y)$ and $(\pm y, \pm x)$. In the topograph of x^2+y^2 the value 5^k will appear just once in each quadrant since each pair of solutions (x,y) and $(-x,-y)$ determines the same fraction x/y. This was guaranteed to happen by Proposition 6.16 which states that any two occurrences of the same prime power in a topograph are related by a symmetry of the topograph, for primes not dividing the conductor, and the conductor here is 1.

For negative discriminants it is not difficult to figure out exactly when R_Δ has a Euclidean algorithm. Recall that this means that for each pair of nonzero elements α_0 and α_1 in R_Δ there should exist elements β and α_2 such that $\alpha_0 = \beta\alpha_1 + \alpha_2$ and $N(\alpha_2) < N(\alpha_1)$. Since α_2 is determined by α_0, α_1, and β, this is equivalent to saying that there should exist an element β in R_Δ such that $N(\alpha_0 - \beta\alpha_1) < N(\alpha_1)$. The last inequality can be rewritten as $N({}^{\alpha_0}/_{\alpha_1} - \beta) < 1$. Geometrically this is saying that every point ${}^{\alpha_0}/_{\alpha_1}$ in the plane should be within a distance less than 1 of some point β in the lattice R_Δ. We can check this by seeing whether the interiors of all the circles of radius 1 centered at points of R_Δ completely cover the plane.



For $\mathbb{Z}[\sqrt{D}]$ with $D < 0$ the critical case $D = -3$ is shown in the figure at the right, where the triangle is an equilateral triangle of side length 1. Here the four circles of radius 1 centered at 0, 1, $\sqrt{-3}$, and $1+\sqrt{-3}$ intersect at the point $(1+\sqrt{-3})/2$ so this point is not within distance less than 1 of an element of $\mathbb{Z}[\sqrt{-3}]$ and therefore the Euclidean algorithm fails in $\mathbb{Z}[\sqrt{-3}]$. For $D < -3$ the lattice $\mathbb{Z}[\sqrt{D}]$ is stretched vertically so the Euclidean algorithm fails in these cases too. For $D = -2$ the lattice is compressed vertically so $\mathbb{Z}[\sqrt{-2}]$ does have a Euclidean algorithm.



In the case of $\mathbb{Z}[\omega]$ with $\omega = (1+\sqrt{1+4d})/2$ and $d < 0$ the upper row of disks is at height $\sqrt{|1+4d|}/2$ above the lower row, so from the figure we see that the condition we need is that this height should be less than $1 + \frac{\sqrt{3}}{2}$. Thus we need $\sqrt{|1+4d|} < 2+\sqrt{3}$. Squaring both sides gives $|1+4d| < 7+4\sqrt{3}$ which is satisfied only in the cases $d = -1,-2,-3$.

In summary, we have shown the following result:

Proposition 8.9. *The only negative discriminants Δ for which R_Δ has a Euclidean algorithm are $\Delta = -3,-4,-7,-8,-11$.*

Notice that these are the first five negative discriminants.

For even discriminants $\Delta = 4D$ it is easy to prove that unique factorization fails in $R_\Delta = \mathbb{Z}[\sqrt{D}]$ in all cases when Δ is negative and there is no Euclidean algorithm:

Proposition 8.10. *Unique factorization fails in $\mathbb{Z}[\sqrt{D}]$ whenever $D < -2$, and it also fails when $D > 0$ and $D \equiv 1$ modulo 4.*

Proof: The number $D^2 - D$ factors in $\mathbb{Z}[\sqrt{D}]$ as $(D + \sqrt{D})(D - \sqrt{D})$, and it also factors as $D(D-1)$. The number 2 divides either D or $D-1$ since one of these two consecutive integers must be even. However, 2 does not divide either $D + \sqrt{D}$ or $D - \sqrt{D}$ in $\mathbb{Z}[\sqrt{D}]$ since $(D \pm \sqrt{D})/2$ is not an element of $\mathbb{Z}[\sqrt{D}]$ as the coefficient of $\sqrt{D}$ in this quotient is not an integer. If we knew that 2 was prime in $\mathbb{Z}[\sqrt{D}]$ we would then have two distinct factorizations of $D^2 - D$ into primes in $\mathbb{Z}[\sqrt{D}]$: one obtained by combining prime factorizations of D and $D-1$ in $\mathbb{Z}[\sqrt{D}]$ and the other obtained by combining prime factorizations of $D + \sqrt{D}$ and $D - \sqrt{D}$. The first factorization would contain the prime 2 and the second would not.

It remains to check that 2 is a prime in $\mathbb{Z}[\sqrt{D}]$ in the cases listed. If it is not a prime, then it factors as $2 = \alpha\beta$ with neither α nor β a unit, so we would have $N(\alpha) = N(\beta) = \pm 2$. Thus the equation $x^2 - Dy^2 = \pm 2$ would have an integer solution (x, y). This is clearly impossible if $D = -3$ or any negative integer less than -3. If $D > 0$ and $D \equiv 1$ modulo 4 then if we look at the equation $x^2 - Dy^2 = \pm 2$ modulo 4 it becomes $x^2 - y^2 \equiv 2 \bmod 4$, but this is impossible since x^2 and y^2 are congruent to 0 or 1 modulo 4, so $x^2 - y^2$ is congruent to 0, 1, or -1. □

This proposition says in particular that unique factorization fails in $\mathbb{Z}[\sqrt{-3}]$, $\mathbb{Z}[\sqrt{-7}]$, and $\mathbb{Z}[\sqrt{-11}]$. However, when we enlarge these rings to $\mathbb{Z}[\omega]$ for ω equal to $(1+\sqrt{-3})/2$, $(1+\sqrt{-7})/2$, and $(1+\sqrt{-11})/2$ we do have unique factorization. A similar thing happens when we enlarge $\mathbb{Z}[\sqrt{-8}]$ to $\mathbb{Z}[\sqrt{-2}]$. In all these cases the enlargement replaces a nonfundamental discriminant by one which is fundamental.

One might wonder whether there are other ways to enlarge $\mathbb{Z}[\sqrt{D}]$ to make prime factorization unique when it is not unique in $\mathbb{Z}[\sqrt{D}]$ itself. Without changing things too drastically, suppose we just tried a different choice of ω. In order to do multiplication within the set $\mathbb{Z}[\omega]$ of numbers $x + y\omega$ with x and y integers one must be able to express ω^2 as $m\omega + n$, which means that ω must satisfy a quadratic equation $\omega^2 - m\omega - n = 0$. This has roots $(m \pm \sqrt{m^2 + 4n})/2$, so we see that larger denominators than 2 in the definition of ω will not work. If m is even, say $m = 2k$, then ω becomes $k \pm \sqrt{k^2 + n}$, with no denominators at all and we are back in the situation of a ring $\mathbb{Z}[\sqrt{D}]$. If m is odd, say $m = 2k+1$, then ω becomes $(2k+1 \pm \sqrt{4k^2 + 4k + 1 + 4n})/2$ which equals $k + (1 \pm \sqrt{1 + 4d})/2$ for $d = k^2 + k + n$ so the ring $\mathbb{Z}[\omega]$ in this case would be the same as when we chose $\omega = (1 + \sqrt{1 + 4d})/2$.

It is known that there are only nine negative discriminants for which R_Δ has unique factorization, the discriminants

$$\Delta = -3, -4, -7, -8, -11, -19, -43, -67, -163$$

These are exactly the nine negative discriminants for which all quadratic forms of that discriminant are equivalent. This is not an accident since the usual way one determines whether unique factorization holds is by proving that unique factorization holds precisely when all forms of the given discriminant are equivalent, as we will see later in the chapter. This is for negative discriminants. For positive discriminants the condition is that all forms are equivalent to either the principal form or its negative.

For positive discriminants the norm form is hyperbolic so it takes on both positive and negative values. The Euclidean algorithm is then modified so that in the equations $\alpha_{i-1} = \beta_i \alpha_i + \alpha_{i+1}$ it is required that $|N(\alpha_{i+1})| < |N(\alpha_i)|$. It is known that there are exactly 16 positive fundamental discriminants for which there is a Euclidean algorithm in R_Δ:

$$\Delta = 5, 8, 12, 13, 17, 21, 24, 28, 29, 33, 37, 41, 44, 57, 73, 76$$

The determination of this list is quite a bit more difficult than for negative discriminants since the norm no longer has the nice geometric meaning of the square of the distance to the origin in the plane.

There are many positive fundamental discriminants for which R_Δ has unique factorization even though there is no Euclidean algorithm. The fundamental discriminants less than 100 with this property are $53, 56, 61, 69, 77, 88, 89, 92, 93, 97$.

To conclude this section we give two applications of unique factorization to quadratic forms. The first will be to find all primitive solutions of $x^2 + 7y^2 = 2^k$. This equation came up in Section 6.2 when we were considering which powers of a prime that divides the conductor for a given nonfundamental discriminant are represented by primitive forms of that discriminant. For the form $x^2 + 7y^2$ the discriminant is -28 with class number 1 and conductor 2 so the question was which powers of 2 are represented by $x^2 + 7y^2$. Obviously 2 and 2^2 are not represented, but we showed that all powers 2^k with $k \geq 3$ are represented. However the method there did not produce actual primitive solutions of $x^2 + 7y^2 = 2^k$ so that is what we will find here.

The form $x^2 + 7y^2$ is the norm form in $\mathbb{Z}[\sqrt{-7}]$ so we are looking for elements $x + y\sqrt{-7}$ of $\mathbb{Z}[\sqrt{-7}]$ of norm $x^2 + 7y^2 = 2^k$ with x and y coprime. The ring $\mathbb{Z}[\sqrt{-7}]$ does not have unique factorization, so we will enlarge it to $\mathbb{Z}[\omega]$ for $\omega = (1 + \sqrt{-7})/2$ since $\mathbb{Z}[\omega]$ does have unique factorization. The only units in $\mathbb{Z}[\omega]$ are ± 1 so prime factorizations are unique up to signs.

We have $N(\omega) = \omega\overline{\omega} = 2$ so $N(\omega^k) = 2^k$. The prime factorization of 2^k in $\mathbb{Z}[\omega]$ is $2^k = \omega^k \overline{\omega}^k$ so the elements of $\mathbb{Z}[\omega]$ of norm 2^k are, up to sign, the products $\omega^l \overline{\omega}^m$ with $l + m = k$. We need to determine which of these products lie in $\mathbb{Z}[\sqrt{-7}]$

and are primitive, that is, not an integer multiple of another element of $\mathbb{Z}[\sqrt{-7}]$ unless that integer is ± 1.

Consider first the case $m = 0$. If ω^k is an element $a + b\sqrt{-7}$ of $\mathbb{Z}[\sqrt{-7}]$ then the norm equation $a^2 + 7b^2 = 2^k$ implies that a and b have the same parity. If they are both even then ω^k would be divisible by 2 in $\mathbb{Z}[\sqrt{-7}]$ and hence also divisible by 2 in $\mathbb{Z}[\omega]$, but this is impossible since 2 factors as $\omega\overline{\omega}$ and $\overline{\omega}$ is not one of the prime factors of ω^k since $\overline{\omega} \neq \pm\omega$. If a and b are both odd then ω^k is 2 times an element of $\mathbb{Z}[\omega]$ and we have the same contradiction. Thus we must have $m > 0$, and similarly we must have $l > 0$.

If $m = 1$ then we are considering the product $\omega^{k-1}\overline{\omega}$ which equals $2\omega^{k-2}$. This is twice an element of $\mathbb{Z}[\omega]$ so it lies in $\mathbb{Z}[\sqrt{-7}]$ and can be written as $x + y\sqrt{-7}$ for some integers x and y. If x and y are not coprime, they are divisible by some prime p which must be 2 since odd primes do not divide $2\omega^{k-2}$ in $\mathbb{Z}[\omega]$, as $N(2\omega^{k-2}) = 2^k$. This leaves the possibility that x and y are both even. If this is the case then we can cancel a 2 from both sides of the equation $2\omega^{k-2} = x + y\sqrt{-7}$ to get ω^{k-2} as an element of $\mathbb{Z}[\sqrt{-7}]$, which is impossible if $k \geq 3$ as we saw in the preceding paragraph. Thus we conclude that $x + y\sqrt{-7} = 2\omega^{k-2}$ gives a primitive solution of $x^2 + 7y^2 = 2^k$ when $k \geq 3$. Similarly, if $l = 1$ we would obtain the conjugate solution $x - y\sqrt{-7}$, just changing the sign of y.

There remains the possibility that both l and m are greater than 1. In these cases $\omega^l\overline{\omega}^m$ would be divisible by 4, giving an element $x + y\sqrt{-7}$ of $\mathbb{Z}[\sqrt{-7}]$ with x and y even, so we would not get a primitive solution of $x^2 + 7y^2 = 2^k$.

Thus we have shown that there are exactly four primitive solutions of $x^2 + 7y^2$ for each $k \geq 3$, differing only in the signs of x and y so there is a unique primitive solution with x and y positive. We can compute this solution by computing $2\omega^{k-2}$ as an element $x + y\sqrt{-7}$. This can be done inductively using the formula:

$$(a + b\sqrt{-7})\left(\frac{1 + \sqrt{-7}}{2}\right) = \frac{(a - 7b) + (a + b)\sqrt{-7}}{2}$$

Here is a table of these values for $k \leq 15$:

k	3	4	5	6	7	8	9	10
(a, b)	$(1, 1)$	$(-3, 1)$	$(-5, -1)$	$(1, -3)$	$(11, -1)$	$(9, 5)$	$(-13, 7)$	$(-31, -3)$

11	12	13	14	15
$(-5, -17)$	$(57, -11)$	$(67, 23)$	$(-47, 45)$	$(-181, -1)$

Omitting the minus signs gives the positive primitive solutions. However, if we tried to simplify the calculations by omitting the minus signs at each step, this does not work since for example if we use the solution $(3, 1)$ for $k = 4$ instead of $(-3, 1)$ in the formula $((a - 7b) + (a + b)\sqrt{-7})/2$, this would give the nonprimitive solution $(-2, 2)$ for $k = 5$ instead of $(-5, -1)$.

This problem has some history. In the early 1900s the number theorist Ramanujan observed that the Diophantine equation $x^2 + 7 = 2^k$ has solutions for $k = 3, 4, 5, 7, 15$ and he conjectured that there were no solutions for larger k. In terms of the preceding example this is saying that the only solutions of $x^2 + 7y^2 = 2^k$ with $y = 1$ occur in these five cases, so $x = 1, 3, 5, 11, 181$ as in the table above. (Note that a solution with $y = 1$ must be primitive.) Ramanujan's conjecture was later proved in a paper by Skolem, Chowla, and Lewis published in 1959.

For the other application of unique factorization we consider the forms $x^2 + 18y^2$ and $2x^2 + 9y^2$ of discriminant -72. The class number here is 2 and these forms are in the two classes. The discriminant -72 is not fundamental since $-72 = 3^2(-8)$ with -8 a fundamental discriminant, so the conductor is 3. This leads us to ask which powers of 3 are represented by the two forms. Neither form represents 3 and only the second form represents 9, but both forms represent 27, coincidentally when $(x, y) = (3, 1)$ in both cases.

As in the preceding example we will enlarge the ring $\mathbb{Z}[\sqrt{-18}]$, which is R_Δ for $\Delta = -72$, to the corresponding ring $\mathbb{Z}[\sqrt{-2}]$ which is R_Δ for $\Delta = -8$, in order to take advantage of the fact that $\mathbb{Z}[\sqrt{-2}]$ has unique factorization while $\mathbb{Z}[\sqrt{-18}]$ does not. Note that $\sqrt{-18} = 3\sqrt{-2}$ so $\mathbb{Z}[\sqrt{-18}]$ is contained in $\mathbb{Z}[\sqrt{-2}]$ as the numbers $a + 3b\sqrt{-2}$.

First we consider the form $x^2 + 18y^2 = N(x + 3y\sqrt{-2})$ so we are looking for elements $a + 3b\sqrt{-2}$ of $\mathbb{Z}\sqrt{-18}]$ of norm 3^k with a and b coprime. An element of $\mathbb{Z}[\sqrt{-2}]$ of norm 3 is $1 + \sqrt{-2}$, so $(1 + \sqrt{-2})^k$ has norm 3^k. However $(1 + \sqrt{-2})^k$ does not lie in $\mathbb{Z}[\sqrt{-18}]$, for suppose $(1 + \sqrt{-2})^k = a + 3b\sqrt{-2}$ for some integers a and b. Taking norms, we would then have $3^k = a^2 + 18b^2$. This implies 3 divides a, hence 3 divides $(1 + \sqrt{-2})^k = a + 3b\sqrt{-2}$ in $\mathbb{Z}[\sqrt{-2}]$, but this is impossible since the prime factorization of 3 in $\mathbb{Z}[\sqrt{-2}]$ is $(1 + \sqrt{-2})(1 - \sqrt{-2})$ and $1 - \sqrt{-2}$ is not a prime factor of $(1 + \sqrt{-2})^k$.

To get an element of $\mathbb{Z}[\sqrt{-18}]$ of norm 3^k we now try $3(1 + \sqrt{-2})^{k-2}$ which has this norm and lies in $\mathbb{Z}[\sqrt{-18}]$ since it is 3 times an element of $\mathbb{Z}[\sqrt{-2}]$. Thus we can write $3(1 + \sqrt{-2})^{k-2} = a + b\sqrt{-18}$ for some integers a and b. To check whether a and b are coprime we note first that by taking norms we see that the only prime that could divide a and b is 3. If 3 does divide a and b we can divide the equation $3(1 + \sqrt{-2})^{k-2} = a + b\sqrt{-18}$ by 3 and deduce that $(1 + \sqrt{-2})^{k-2}$ is an element of $\mathbb{Z}[\sqrt{-18}]$, but we saw in the preceding paragraph that this is not the case if $k \geq 3$. Thus we have a solution of $x^2 + 18y^2 = 3^k$ with coprime integers x and y for each $k \geq 3$.

Now we turn to the form $2x^2 + 9y^2$. The starting point here is the observation that if we restrict the form $x^2 + 18y^2$ to pairs (x, y) with x even, then we have $(2x)^2 + 18y^2$ which is just $2(2x^2 + 9y^2)$, or twice the form $2x^2 + 9y^2$. Thus we are looking for elements $2x + y\sqrt{-18}$ of $\mathbb{Z}[\sqrt{-18}]$ of norm $2 \cdot 3^k$ with x and y coprime.

A reasonable guess might be $\sqrt{-2}\cdot 3(1+\sqrt{-2})^{k-2}$ which has norm $2\cdot 3^k$. This lies in $\mathbb{Z}[\sqrt{-18}]$ since it is 3 times an element of $\mathbb{Z}[\sqrt{-2}]$ so we can write it as $a+b\sqrt{-18}$. A prime dividing a and b must divide the norm $2\cdot 3^k$ so it must be 2 or 3. If 2 divided a and b then 4 would divide the norm so this is impossible. If 3 divided a and b then after canceling this 3 we would have $\sqrt{-2}(1+\sqrt{-2})^{k-2}$ being an element of $\mathbb{Z}[\sqrt{-18}]$, but this is impossible by the same argument that showed $(1+\sqrt{-2})^k$ was not in $\mathbb{Z}[\sqrt{-18}]$. Thus a and b are coprime. It remains only to check that a is even, but this is immediate from the norm equation $a^2+18b^2=2\cdot 3^k$.

These arguments show that all the powers 3^k with $k\ge 3$ are represented by both x^2+18y^2 and $2x^2+9y^2$. This sort of behavior, with nonequivalent forms of the same discriminant representing the same prime powers, can only happen for nonfundamental discriminants, and then only for powers of primes dividing the conductor, as we know from Chapter 6.

The trick of realizing $2x^2+9y^2$ as a multiple of the form obtained by restricting the norm form x^2+18y^2 to certain values of x and y in $\mathbb{Z}[\sqrt{-18}]$ is in fact part of a general pattern that will be explored in the next section.

Exercises

1. (a) According to Proposition 8.10, unique factorization fails in $\mathbb{Z}[\sqrt{D}]$ when $D=-3$ since the number $D(D-1)=12$ has two distinct prime factorizations in $\mathbb{Z}[\sqrt{D}]$. On the other hand, if we enlarge $\mathbb{Z}[\sqrt{-3}]$ to $\mathbb{Z}[\omega]$ for $\omega=(1+\sqrt{-3})/2$ then unique factorization is restored. Explain how the two prime factorizations of 12 in $\mathbb{Z}[\sqrt{-3}]$ give rise to the same prime factorization in $\mathbb{Z}[\omega]$ (up to units).

(b) Do the same thing for the case $D=-7$.

2. Show that the number 8 has two different prime factorizations in $\mathbb{Z}[\sqrt{-7}]$, one with three prime factors and the other with two prime factors.

3. In R_Δ for $\Delta=-3$ show that the only primes α for which $\overline{\alpha}$ is a unit times α are $\sqrt{-3}$ and units times $\sqrt{-3}$.

4. In this problem we consider $\mathbb{Z}[\sqrt{-2}]$, so elements of $\mathbb{Z}[\sqrt{-2}]$ are sums $x+y\sqrt{-2}$ for integers x and y, with $N(x+y\sqrt{-2})=(x+y\sqrt{-2})(x-y\sqrt{-2})=x^2+2y^2$.

(a) Draw the topograph of x^2+2y^2 including all values less than 70 (by symmetry, it suffices to draw just the upper half of the topograph). Circle the values that are prime (prime in $\mathbb{Z}$, that is). Also label each region with its x/y fraction.

(b) Which primes in $\mathbb{Z}$ factor in $\mathbb{Z}[\sqrt{-2}]$?

(c) Using the information in part (a), list all primes in $\mathbb{Z}[\sqrt{-2}]$ of norm less than 70.

(d) Draw a diagram in the xy-plane showing all elements $x+y\sqrt{-2}$ in $\mathbb{Z}[\sqrt{-2}]$ of norm less than 70 as small dots, with larger dots or squares for the elements that are prime in $\mathbb{Z}[\sqrt{-2}]$. (There is symmetry, so the primes in the first quadrant determine the primes in the other quadrants.)

(e) Show that the only primes $x + y\sqrt{-2}$ in $\mathbb{Z}[\sqrt{-2}]$ with x even are $\pm\sqrt{-2}$. (Your diagram in part (d) should give some evidence that this is true.)

(f) Factor $4 + \sqrt{-2}$ into primes in $\mathbb{Z}[\sqrt{-2}]$.

(g) Use the unique factorization property in $\mathbb{Z}[\sqrt{-2}]$ to determine which numbers are represented by the form $x^2 + 2y^2$, as was done in the text for $x^2 + y^2$.

5. Following the two examples at the end of this section, find primitive solutions of $x^2 + 18y^2 = 3^k$ and of $2x^2 + 9y^2 = 3^k$ for $k = 3, 4, 5, 6, 7, 8$.

8.3 The Correspondence Between Forms and Ideals

So far in this chapter we have focused on principal forms, and now we begin to extend what we have done to arbitrary forms. For principal forms we began by factoring them as a product of two linear factors whose coefficients involved square roots, for example the factorization $x^2 - Dy^2 = (x + \sqrt{D}y)(x - \sqrt{D}y)$ in the case of discriminant $\Delta = 4D$. For a general form $Q(x,y) = ax^2 + bxy + cy^2$ of discriminant Δ the corresponding factorization is $a(x - \alpha y)(x - \overline{\alpha}y)$ where α is a root of the quadratic equation $ax^2 + bx + c = 0$. Thus we have:

$$ax^2 + bxy + cy^2 = a\Big(x - \frac{-b + \sqrt{\Delta}}{2a}y\Big)\Big(x - \frac{-b - \sqrt{\Delta}}{2a}y\Big)$$

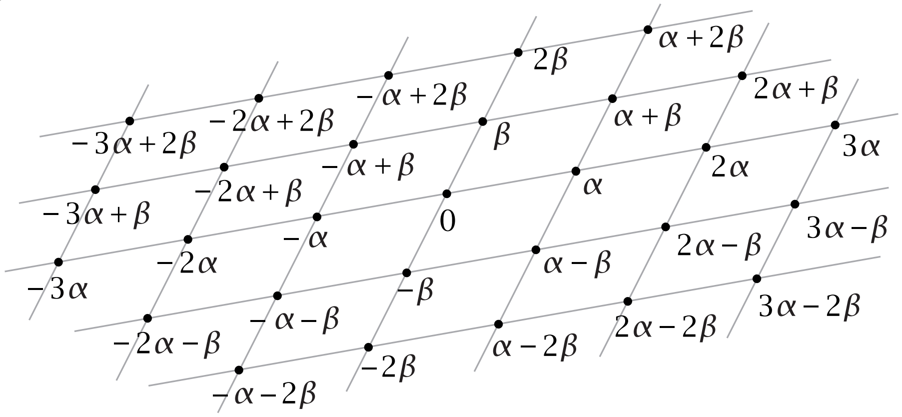
An equivalent equation that will be more convenient for our purposes is obtained by multiplying both sides of the preceding equation by the coefficient a:

$$a(ax^2 + bxy + cy^2) = \Big(ax + \frac{b + \sqrt{\Delta}}{2}y\Big)\Big(ax + \frac{b - \sqrt{\Delta}}{2}y\Big)$$

The advantage of writing the equation this way is that in each of the two linear factors on the right the coefficients of x and y now lie in the ring R_Δ since b must have the same parity as Δ. Thus if $\Delta = 4D$ we can eliminate the denominator 2 in the coefficient of y to obtain an element of $\mathbb{Z}[\sqrt{D}]$ while if $\Delta = 4d + 1$ the fraction lies in $\mathbb{Z}[\omega]$ since b is odd. Another thing to observe is that the right side of the equation is just the norm $N\Big(ax + \frac{b+\sqrt{\Delta}}{2}y\Big)$, so the displayed equation above can be written more concisely as $aQ(x,y) = N\Big(ax + \frac{b+\sqrt{\Delta}}{2}y\Big)$.

For a form $Q(x,y) = ax^2 + bxy + cy^2$ the set of numbers $ax + \frac{b+\sqrt{\Delta}}{2}y$ as x and y range over all integers forms a lattice contained in the larger lattice R_Δ in the plane. Here by a ***lattice*** we mean a set of numbers of the form $\alpha x + \beta y$ for fixed nonzero elements α and β of R_Δ, with x and y varying over $\mathbb{Z}$, and we assume that α and β

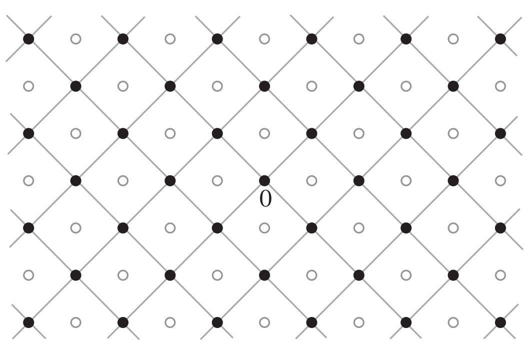
do not lie on the same line through the origin. We denote this lattice by $L(\alpha,\beta)$ and call α and β a ***basis*** for the lattice.



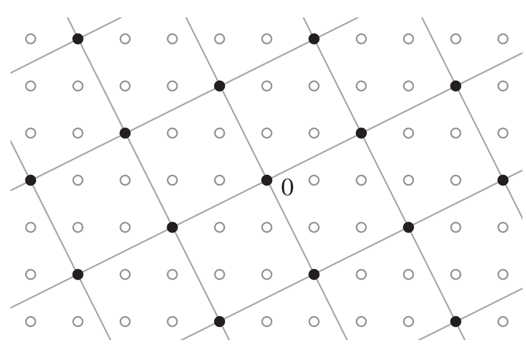
In particular, associated to the form Q we have the lattice $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ consisting of all the numbers $ax + \frac{b+\sqrt{\Delta}}{2}y$ for integers x and y. The earlier equation $aQ(x,y) = N\left(ax + \frac{b+\sqrt{\Delta}}{2}y\right)$ then says that the form Q is obtained from the lattice L_Q by taking the norms of all its elements and multiplying by the constant factor $1/a$, which can be regarded as a sort of normalization constant as we will see in more detail later.

Let us look at some examples to see what L_Q can look like in the case $\Delta = -4$ so $R_\Delta = \mathbb{Z}[i]$, the Gaussian integers. In this case we have $ax + \frac{b+\sqrt{\Delta}}{2}y = ax + (b' + i)y$ where $b' = b/2$ is an integer since b always has the same parity as Δ. For the principal form $x^2 + y^2$ we have $a = 1$ and $b' = 0$ so $L_Q = L(1,i) = \mathbb{Z}[i]$. Four more cases are shown in the figures below.

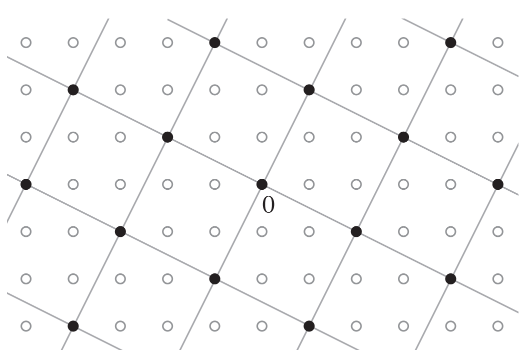
$2x^2 + 2xy + y^2 \longleftrightarrow L(2, 1+i)$



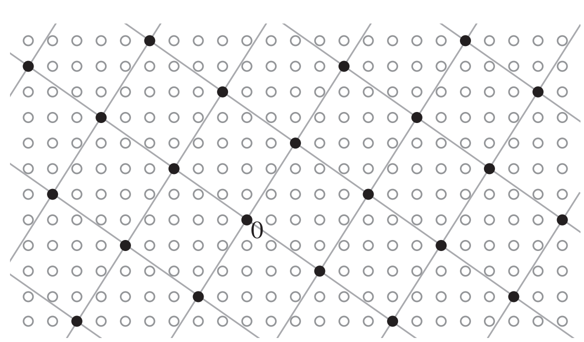
$5x^2 + 4xy + y^2 \longleftrightarrow L(5, 2+i)$



$5x^2 + 6xy + 2y^2 \longleftrightarrow L(5, 3+i)$



$13x^2 + 10xy + 2y^2 \longleftrightarrow L(13, 5+i)$



In each case the lattice forms a grid of squares, rotated and expanded from the square grid formed by $\mathbb{Z}[i]$ itself. Not all lattices in $\mathbb{Z}[i]$ form square grids since for example one could have a lattice of long thin rectangles such as $L(10, i)$.

A 90 degree rotation of the plane about the origin takes a square lattice to itself. Conversely, a lattice L that is taken to itself by a 90 degree rotation about the origin must be a square lattice. To see this, observe first that the 90 degree rotation takes a point α of L that is closest to the origin to another point β of L closest to the origin, with the sum $\alpha + \beta$ giving the fourth vertex of a square of lattice points. The lattice $L(\alpha, \beta)$ is then a square lattice contained in L. In fact we must have $L = L(\alpha, \beta)$, for if there were a point of L in the interior of a square of $L(\alpha, \beta)$ then such a point would be closer to a corner of the square than the length of the side of the square, which is impossible since the minimum distance between any two points in a lattice equals the minimum distance from the origin to a lattice point.

Since 90 degree rotation is the same as multiplication of complex numbers by i, we could also say that square lattices are those that are taken to themselves by multiplication by i. Once a lattice has this property, it follows that multiplication by an arbitrary element of $\mathbb{Z}[i]$ takes the lattice into itself. Namely, if we know that $i\alpha$ is in a lattice L whenever α is in L, then for arbitrary integers m and n it follows that $m\alpha$ and $ni\alpha$ are in L and hence $(m + ni)\alpha$ is in L.

There is a standard term for this concept. A lattice L in R_Δ is called an ***ideal*** if for each element α in L and each β in R_Δ the product $\beta\alpha$ is in L. In other words, L is taken to itself by multiplication by every element of R_Δ. The term “ideal” may seem like an odd name, but it originally arose in a slightly different context where it seems more natural, as we will see later in the chapter. For now we can just imagine that ideals are the best kind of lattices, “ideal lattices”.

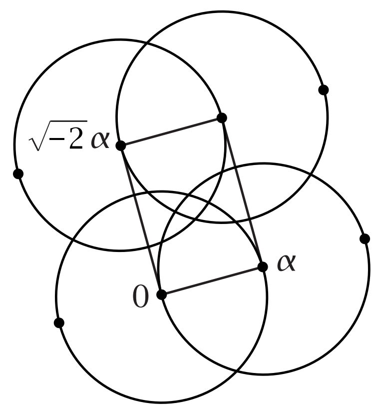
The fact that all lattices L_Q in $\mathbb{Z}[i]$ are square lattices is a special case of the following general fact:

Proposition 8.11. *For each quadratic form $Q = ax^2 + bxy + cy^2$ of discriminant Δ the lattice $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ is an ideal in R_Δ.*

Proof: To cover all discriminants at once we can write R_Δ as $\mathbb{Z}[\tau]$ for $\tau = \frac{e+\sqrt{\Delta}}{2}$ where e is 0 if $\Delta = 4D$ and 1 if $\Delta = 4d + 1$. What we need to check in order to verify that the lattice $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ is an ideal is that both of the products $\tau \cdot a$ and $\tau \cdot \frac{b+\sqrt{\Delta}}{2}$ are elements of L_Q. For the product $\tau \cdot a$ this means we want to solve the equation $\frac{e+\sqrt{\Delta}}{2} \cdot a = ax + \frac{b+\sqrt{\Delta}}{2} \cdot y$ for integers x and y. Comparing the coefficients of $\sqrt{\Delta}$ on both sides of the equation, we get $y = a$, an integer. Substituting $y = a$ into the equation then gives $\frac{ea}{2} = ax + \frac{ba}{2}$ so $x = \frac{e-b}{2}$. This is an integer since both e and b have the same parity as Δ.

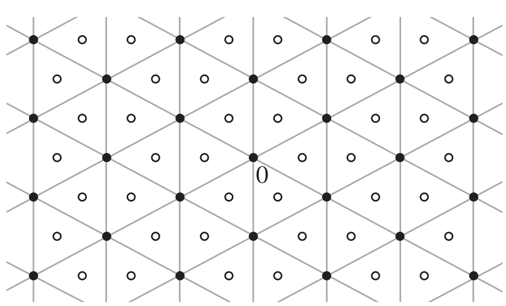
For the other product $\tau \cdot \frac{b+\sqrt{\Delta}}{2}$ we have an equation $\frac{e+\sqrt{\Delta}}{2} \cdot \frac{b+\sqrt{\Delta}}{2} = ax + \frac{b+\sqrt{\Delta}}{2} \cdot y$ which we can rewrite as $\frac{eb+\Delta+(e+b)\sqrt{\Delta}}{4} = ax + \frac{b+\sqrt{\Delta}}{2} y$. From the coefficients of $\sqrt{\Delta}$ we get $y = \frac{e+b}{2}$ which is an integer since e and b have the same parity. Then the equation becomes $\frac{eb+\Delta}{4} = ax + \frac{eb+b^2}{4}$ which simplifies to $\Delta = 4ax + b^2$. Since $\Delta = b^2 - 4ac$ we have the integer solution $x = -c$. $\square$

We saw in the case of $\mathbb{Z}[i]$ that all ideals are square lattices, so they are obtained from $\mathbb{Z}[i]$ by rotation about the origin and expansion. There are a few other negative discriminants where the same thing happens and all ideals differ only by rotation and rescaling, either expansion or contraction. One example is when $\Delta = -8$ so we have $R_\Delta = \mathbb{Z}[\sqrt{-2}]$ which forms a rectangular lattice with rectangles of side lengths 1 and $\sqrt{2}$. For an arbitrary ideal L in $\mathbb{Z}[\sqrt{-2}]$ let α be a nonzero point in L closest to the origin. Since L is an ideal, the product $\sqrt{-2}\,\alpha$ must also be in L. Since multiplication by $\sqrt{-2}$ rotates the plane by 90 degrees and expands it by a factor of $\sqrt{2}$, the set of all linear combinations $\alpha x + \sqrt{-2}\,\alpha y$ for integers x and y forms a rectangular sublattice L' of L obtained from $\mathbb{Z}[\sqrt{-2}]$ by rotation and expansion. Since we chose α as the closest point of L to the origin, say of distance A to the origin, there can be no points of L within a distance less than A of any point of L'. In other words, if one takes the union of the interiors of all disks of radius A centered at points of L', this union intersects L just in L'. However, this union is the whole plane since the ratio of the side lengths of the rectangles of L' is $\sqrt{2}$. Thus L equals the rectangular lattice L'.

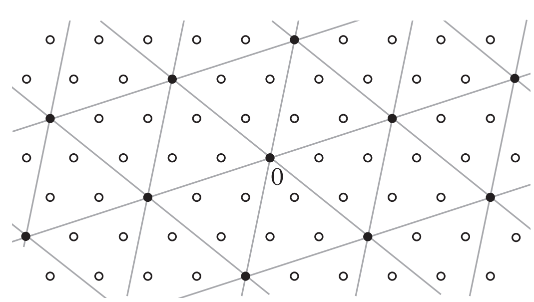


This is essentially the same geometric argument we used to show that $\mathbb{Z}[\sqrt{-2}]$ has a Euclidean algorithm. There were five negative discriminants Δ for which R_Δ has a Euclidean algorithm, $\Delta = -3, -4, -7, -8, -11$. The argument in the preceding paragraph shows that in each of these cases all ideals in R_Δ are equivalent under rotation and rescaling. In the case $\Delta = -3$ the Eisenstein integers $\mathbb{Z}[\omega]$ form a grid of equilateral triangles so all ideals are also grids of equilateral triangles that are taken to themselves by multiplication by ω, rotating the plane by 60 degrees. Two examples are shown below.

$3x^2 + 3xy + y^2 \longleftrightarrow L(3, 1+\omega)$



$7x^2 + 5xy + y^2 \longleftrightarrow L(7, 2+\omega)$



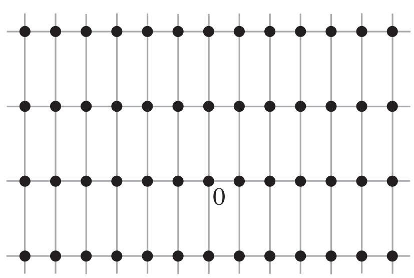
For $\Delta = -7$ and -11 the lattice $R_\Delta = \mathbb{Z}[\omega]$ for $\Delta = -3$ is stretched vertically to form a grid of isosceles triangles and all ideals are also grids of isosceles triangles, rotated and rescaled from the triangles in R_Δ.

We have been using the fact that multiplication by a fixed nonzero complex number α always has the effect of rotating and rescaling the plane, keeping the origin fixed. Since multiplication by α sends 1 to α, the rescaling factor is the distance from α to the origin and the angle of rotation is the angle between the positive x-axis and the ray from the origin to α. Since α can be any nonzero complex number, every rotation and rescaling is realizable as multiplication by a suitably chosen α.

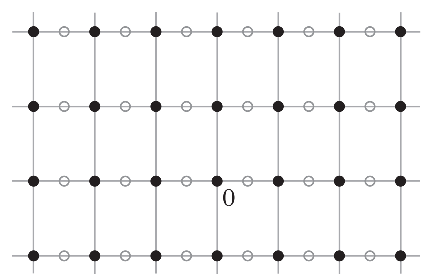
Let us look at some examples of discriminants where not all forms are equivalent to see whether there is more variety in the shapes of the lattices L_Q, so they are not all obtained from R_Δ by rotation and rescaling. The examples will all be for negative discriminants since this is the case that the norm of an element of R_Δ has the geometric interpretation as the square of the distance to the origin, but when we make general statements about lattices these will apply to both positive and negative discriminants.

For a first example consider the lattices L_Q in $\mathbb{Z}[\sqrt{-6}]$ for the two nonequivalent forms x^2+6y^2 and $2x^2+3y^2$ of discriminant -24.

$x^2+6y^2 \longleftrightarrow L(1,\sqrt{-6})$



$2x^2+3y^2 \longleftrightarrow L(2,\sqrt{-6})$

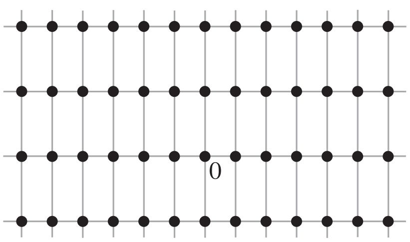


The two lattices do not appear to differ just by rotation and rescaling, and we can verify this by computing the ratio of the distances from the origin to the closest lattice point and to the next-closest lattice point on a different line through the origin. For the lattice $\mathbb{Z}[\sqrt{-6}]$ this ratio is $1/\sqrt{6}$ while for the other lattice it is $2/\sqrt{6}$. If the lattices differed only by rotation and rescaling, the ratios would be the same.

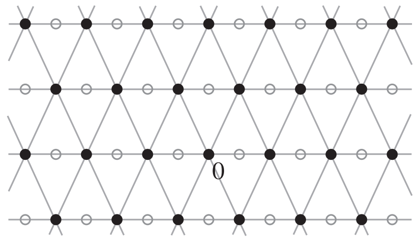
Instead of measuring the distances from the origin to a nearby lattice point we could measure the square of the distance, which is the norm of the lattice point. For the forms shown above we would then get the ratios $1/6$ and $4/6 = 2/3$. It is no accident that these are the ratios between the coefficients of x^2 and y^2 in the two forms since these coefficients give the two smallest values of the forms, which occur on either side of the source edge in their topographs. The norms of points in the lattice are related to the values of the form by the formula $aQ(x,y) = N(ax + \frac{b+\sqrt{\Delta}}{2}y)$, so the smallest norms correspond to the smallest values of the form, with the scaling factor a in the left side of the formula accounting for the fact that the fraction $4/6$ reduces to $2/3$ by dividing numerator and denominator by $a=2$.

As another example, consider the lattices L_Q in $\mathbb{Z}[\sqrt{-5}]$ for the nonequivalent forms x^2+5y^2 and $2x^2+2xy+3y^2$ of discriminant -20.

$x^2+5y^2 \longleftrightarrow L(1,\sqrt{-5})$



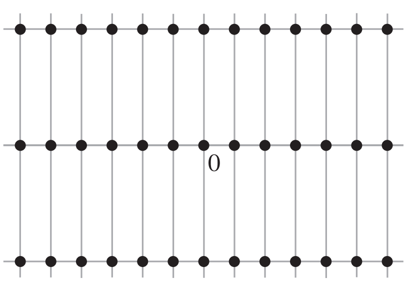
$2x^2+2xy+3y^2 \longleftrightarrow L(2,1+\sqrt{-5})$



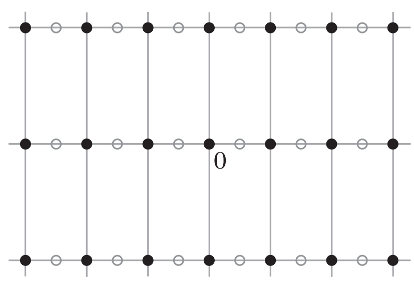
It is clear visually that the two lattices are not related just by rotation and rescaling since the first lattice is rectangular while the second is not, and we can verify this by computing the ratios of the norms of the two closest lattice points to the origin lying on different lines through the origin. For the first lattice the ratio is $^1/_5$ corresponding to the topograph having a source edge with adjacent labels 1 and 5, as in the preceding example. For the second lattice the points closest to the origin are ± 2 and $\pm 1 \pm \sqrt{-5}$ with norms 4 and 6, giving a ratio $^4/_6$ which reduces to $^2/_3$ via the rescaling factor $a = 2$. The topograph of the second form has a source vertex surrounded by the labels $2,3,3$ for $^x/_y = {}^1/_0$, $^0/_1$, and $^{-1}/_1$. The two 3's correspond to the two equal sides of the isosceles triangles in the figure, of norm 6 which rescales to 3

A slightly more complicated example is $\mathbb{Z}[\sqrt{-14}]$ with $\Delta = -56$ where there are four proper equivalence classes of forms:

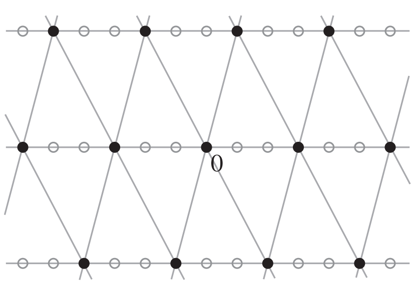
$x^2+14y^2 \longleftrightarrow L(1,\sqrt{-14})$



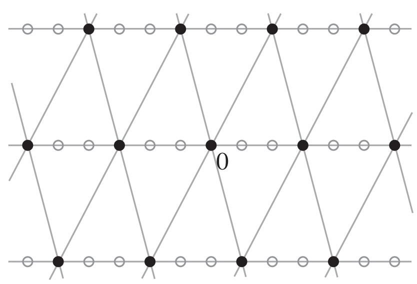
$2x^2+7y^2 \longleftrightarrow L(2,\sqrt{-14})$



$3x^2+2xy+5y^2 \longleftrightarrow L(3,1+\sqrt{-14})$



$3x^2-2xy+5y^2 \longleftrightarrow L(3,-1+\sqrt{-14})$



For the first two forms the ratios of smallest norms are $^1/_{14}$ and $^4/_{14} = {}^2/_7$. For the second two forms the norms of the three sides of the triangles are 9, 15, and 18 so the ratio for the smaller two norms is $^9/_{15} = {}^3/_5$. The second two forms are equivalent but not properly equivalent since their topographs have a source vertex surrounded by the three distinct numbers 3, 5, and 6, the rescalings of the norms 9, 15, and 18. The topographs of these two forms are mirror images obtained by changing the sign of x or y, thus changing the sign of the coefficient of the middle term xy in the form. The corresponding lattices are also mirror images obtained by reflecting across either the x-axis or the y-axis, which also amounts to changing the sign of x or y. These two lattices are not equivalent under rotation and rescaling, so none of the four lattices in this example are equivalent by rotation and rescaling.

Recall that the three values of an elliptic form surrounding a source vertex satisfy the triangle inequalities, so each value is less than or equal to the sum of the other two. This means that for the triangles in the lattices the square of each side length is less than or equal to the sum of the squares of the other two side lengths. Comparing these inequalities with the Pythagorean theorem, this is just saying that the triangles are acute triangles, unless the square of one side is actually equal to the sum of the squares of the other two sides in which case it is a right triangle. This only happens when there is a source edge instead of a source vertex. In this case the grid is rectangular, with each rectangle subdivided into two right triangles by either of its diagonals, but there is no reason to choose one diagonal rather than the other so it seems best to ignore the diagonals and just draw the rectangles.

As we noted above, the two lattices $L(3, 1+\sqrt{-14})$ and $L(3, -1+\sqrt{-14})$ in $\mathbb{Z}[\sqrt{-14}]$ are mirror images of each other under reflection across either the x-axis or the y-axis. Reflecting a lattice across the y-axis gives the same result as reflecting across the x-axis since lattices always have 180 degree rotational symmetry about the origin. Reflecting a lattice across the x-axis amounts to taking the conjugates of all elements of the lattice, so the reflection of a lattice $L = L(\alpha, \beta)$ is the lattice $\overline{L} = L(\overline{\alpha}, \overline{\beta})$ called the ***conjugate lattice.*** If L is an ideal it is easy to check that $\overline{L}$ is also an ideal, so in this case $\overline{L}$ is the ***conjugate ideal*** of L. For lattices coming from forms, the conjugate of $L(a, \frac{b+\sqrt{\Delta}}{2})$ is $L(a, \frac{b-\sqrt{\Delta}}{2})$ which is the same as $L(a, \frac{-b+\sqrt{\Delta}}{2})$.

A lattice is equal to its conjugate exactly when it is symmetric with respect to reflection across the coordinate axes. In the example of lattices in $\mathbb{Z}[\sqrt{-14}]$ the first two lattices have this symmetry property while the second two do not.

Proposition 8.12. *A lattice $L(a, \frac{b+\sqrt{\Delta}}{2})$ is equal to its conjugate if and only if $b \equiv 0$ mod a. These are the rectangular lattices $L(a, \frac{\sqrt{\Delta}}{2})$ with $b = 0$ and the isosceles triangle lattices $L(a, \frac{a+\sqrt{\Delta}}{2})$ with $b = a$.*

Proof: Consider the points of a lattice $L(a, \frac{b+\sqrt{\Delta}}{2})$ that are in the same horizontal row as $\frac{b+\sqrt{\Delta}}{2}$. These points are equally spaced along this row at distance $|a|$ apart. The lattice equals its conjugate exactly when reflection across the y-axis takes this set of points to itself, so the only possibilities are that the set contains the point $\frac{\sqrt{\Delta}}{2}$ or it contains $\frac{a+\sqrt{\Delta}}{2}$. Hence the lattice is either the rectangular lattice $L(a, \frac{\sqrt{\Delta}}{2})$ or the isosceles triangle lattice $L(a, \frac{a+\sqrt{\Delta}}{2})$. In both cases these are lattices $L(a, \frac{b+\sqrt{\Delta}}{2})$ with $b \equiv 0 \bmod a$, and conversely any lattice $L(a, \frac{b+\sqrt{\Delta}}{2})$ with $b \equiv 0 \bmod a$ is equal to one of these two lattices since $L(a, \frac{b+\sqrt{\Delta}}{2})$ is unchanged when multiples of a are added to $\frac{b+\sqrt{\Delta}}{2}$, thus adding multiples of $2a$ to b. □

The two types of self-conjugate lattices $L(a, \frac{\sqrt{\Delta}}{2})$ and $L(a, \frac{a+\sqrt{\Delta}}{2})$ correspond to the forms $ax^2 + cy^2$ and $ax^2 + axy + cy^2$ whose topographs have mirror symmetry. As we saw in Proposition 5.6, all forms with mirror symmetric topographs are equivalent to forms of these two types.

In general, most ideals $L(a, \frac{b+\sqrt{\Delta}}{2})$ are not self-conjugate. For example in the Gaussian integers $\mathbb{Z}[i]$ all ideals are square lattices rotated and expanded from the full lattice $\mathbb{Z}[i]$, but the only ones that are vertically and horizontally symmetric are the ones where the angle of rotation is a multiple of 45 degrees, so these are the lattices $\mathbb{Z}[i]$ and $L(2, 1+i)$ or rescalings of these.

The examples we have seen so far lead one to ask how exact a correspondence there is between proper equivalence classes of forms of a given discriminant Δ and the shapes of lattices that are ideals in R_Δ, where two lattices that differ only by rotation and rescaling are regarded as having the same shape. The main theorem in this section will be that this is an exact one-to-one correspondence for negative discriminants, while for positive discriminants there is an analogous one-to-one correspondence using a more algebraic analogue of "shape" for lattices that works for both positive and negative discriminants.

Before getting to the main theorem we will first explain a few general facts about lattices in R_Δ. Let us write R_Δ as $\mathbb{Z}[\tau]$ for $\tau = \sqrt{D}$ when $\Delta = 4D$ and $\tau = \frac{1+\sqrt{\Delta}}{2}$ when $\Delta = 4d + 1$. Let L be a lattice in $\mathbb{Z}[\tau]$. Since L is not entirely contained in the x-axis there exist elements $m + n\tau$ in L with $n > 0$. Choose such an element $\alpha = m + n\tau$ with minimum positive n, so α lies in the n^{th} row of $\mathbb{Z}[\tau]$ and there are no elements of L in any row between the 0^{th} and the n^{th} rows. Since L is a lattice all elements of L must then lie in rows numbered an integer multiple of n. In particular the element $k\alpha$ lies in the kn^{th} row for each integer k. These elements $k\alpha$ lie on a line through the origin, and L must also contain elements not on this line, so some kn^{th} row must contain another element β of L besides $k\alpha$. The difference $\beta - k\alpha$ then lies in the x-axis and is a nonzero integer in L. Choosing a minimal positive integer p in L, the lattice property of L implies that the integers in L are precisely the integer multiples of p. It follows that L contains the lattice $L(p, \alpha) = L(p, m + n\tau)$, and in fact L is equal to $L(p, m + n\tau)$ otherwise either p or n would not be minimal. We are free to change m by adding or subtracting any integer multiple of p without affecting the lattice, so we may assume $0 \le m < p$.

Thus we see that every lattice L in $\mathbb{Z}[\tau]$ has a basis of the special type $p, m + n\tau$ for p and n positive integers and m an integer in the range $0 \le m < p$. Such a basis is called a ***reduced basis***. A reduced basis for a lattice L is unique since p is the smallest positive integer in L and the first row of L above the x-axis is in the n^{th} row of $\mathbb{Z}[\tau]$, with the elements of L in this row equally spaced p units apart so there is a unique such element $m + n\tau$ with $0 \le m < n$. Thus one can tell whether two lattices in $\mathbb{Z}[\tau]$ are equal by finding a reduced basis for each lattice and seeing whether these reduced bases are equal.

Let us describe how to compute a reduced basis for a lattice $L(\alpha_1, \alpha_2)$ where α_1, α_2 is an arbitrary given basis. There are three simple ways to change from one basis to another basis for the same lattice:

(1) Replace one α_i with $\alpha_i + k\alpha_j$, adding an integer k times the other basis element α_j to α_i. Geometrically this changes the parallelogram with vertices $0, \alpha_1, \alpha_2, \alpha_1 + \alpha_2$ to a parallelogram with one side the same, the side from 0 to α_j, but the opposite side with ends α_i and $\alpha_i + \alpha_j$ is translated along the line containing it.

(2) Replace one α_i by $-\alpha_i$.

(3) Interchange α_1 and α_2.

These operations on bases can be interpreted as operations on matrices if we let $\alpha_1 = a_1 + b_1\tau$ and $\alpha_2 = a_2 + b_2\tau$ and then consider the matrix $A = \left(\begin{smallmatrix} a_1 & a_2 \\ b_1 & b_2 \end{smallmatrix}\right)$. The operation (1) changes one column of A by adding k times the other column to it. Operation (2) multiplies one column by -1, and operation (3) interchanges the two columns. The goal is to use these three operations to change the given matrix A to a matrix of the special form $\left(\begin{smallmatrix} a & b \\ 0 & c \end{smallmatrix}\right)$ with a and c positive and with $0 \le b < a$, so this will be the matrix of a reduced basis.

First we focus on the second row of A. This must have a nonzero entry since α_1 and α_2 are not both contained in the x-axis. The nonzero entries in the second row can be made positive by type (2) operations. If both b_i entries are positive choose a column with smallest positive entry b_i. By subtracting a suitable multiple of this column from the other column we can make the other column have its entry b_j satisfy $0 \le b_j < b_i$. This process can be repeated using columns with successively smaller second entries until only one nonzero b_i remains. Switching this column with the first column if necessary, we can then assume that $b_1 = 0$ and $b_2 > 0$. Then a_1 must be nonzero, and if it is negative we can make it positive by multiplying the first column by -1. Finally, we can make a_2 satisfy $0 \le a_2 < a_1$ by adding or subtracting a multiple of the first column to the second column to finish the process.

An important quantity associated to a lattice L in $\mathbb{Z}[\tau]$ is the number of parallel translates of L, including L itself, that are needed to completely cover all points of the larger lattice $\mathbb{Z}[\tau]$. For example if $a, b + c\tau$ is a reduced basis for L one can first translate L horizontally by the numbers $0, 1, \cdots, a-1$ to cover all of the x-axis and all rows of $\mathbb{Z}[\tau]$ containing points of L. Then c translates of these rows in the direction of τ will cover $\mathbb{Z}[\tau]$ for a total of ac translates of L to cover $\mathbb{Z}[\tau]$.

For a lattice L in $\mathbb{Z}[\tau]$ the number of translates of L needed to cover all of $\mathbb{Z}[\tau]$ is called the ***norm*** of L and written $N(L)$. Any two translates of L are either disjoint or coincide exactly, so there is a unique set of translates of L covering $\mathbb{Z}[\tau]$. Thus there is no ambiguity in the value of $N(L)$. As the reader can see by looking at the various lattices we have pictured earlier in this section, the norm of a lattice measures how "large" or "spread out" the lattice is compared with $\mathbb{Z}[\tau]$.

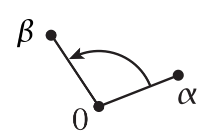
Another way to interpret the norm is in terms of areas. For a basis α,β for a lattice L consider the parallelogram $P_{\alpha,\beta}$ with vertices 0, α, β, and $\alpha+\beta$.

Proposition 8.13. *For a lattice L in $\mathbb{Z}[\tau]$ with basis α,β the area of the parallelogram $P_{\alpha,\beta}$ is independent of the choice of the basis α,β. The ratio of this area to the corresponding area for any basis parallelogram for the full lattice $\mathbb{Z}[\tau]$ is equal to the norm $N(L)$.*

Proof: The operations (1)–(3) on bases do not change the area of basis parallelograms, so every basis parallelogram for L has area equal to the area of $P_{a,b+c\tau}$ for the reduced basis $a, b+c\tau$ for L. To prove the statement about the ratio of areas, note that the area of $P_{a,b+c\tau}$ does not depend on b so we can assume that $b=0$. The parallelogram $P_{a,c\tau}$ decomposes as ac nonoverlapping copies of the parallelogram $P_{1,\tau}$ for $\mathbb{Z}[\tau]$, so the ratio of the areas is ac, which is the norm of the lattice $L=L(a,b+c\tau)$. □

There is also a more algebraic description of the norm of a lattice $L(\alpha,\beta)$ in terms of determinants. If we write $\alpha=a+b\tau$ and $\beta=c+d\tau$ then we have the associated matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. An operation of type (1) adding a multiple of one column to the other does not change the determinant of the matrix, while operations (2) and (3) only change the sign of the determinant. Since the absolute value of the determinant is unchanged by all three types of operations, it can be computed from a reduced basis $a, b+c\tau$ where it is ac, the norm of the lattice. Thus for a lattice L with basis $a+b\tau, c+d\tau$ we have $N(L)=|ad-bc|$.

The sign of the determinant $ad-bc$ has a geometric interpretation as well. We will say the basis α,β is *positively ordered* if the angle from the ray from 0 through α to the ray from 0 through β is between 0 and π, and if the angle is between 0 and $-\pi$ then we say the basis is *negatively ordered*. Reversing the order of two basis elements thus changes the positive ordering to the negative ordering and vice versa. The statement is then that α,β is positively or negatively ordered exactly according to whether $ad-bc$ is positive or negative. To verify this we again use the operations (1)–(3). Operation (1) does not change whether a basis is positively ordered or negatively ordered, while operations (2) and (3) take a positively ordered basis to a negatively ordered basis and vice versa. The sign of the determinant behaves in exactly the same way, so if we go backwards



through the sequence of operations converting α, β into a reduced basis, which is obviously positively ordered with positive determinant, we see that at each step the assertion continues to be true.

Given a lattice $L(\alpha, \beta)$ and a nonzero element γ of $\mathbb{Z}[\tau]$ we can multiply all elements of L by γ to form a new lattice $\gamma L = L(\gamma\alpha, \gamma\beta)$. To check that this is indeed a lattice we should check that $\gamma\alpha$ and $\gamma\beta$ do not lie on the same line through the origin, but if they did then we would have $\gamma\alpha = t\gamma\beta$ for some real number t, and then after canceling γ from this equation we would have $\alpha = t\beta$ which would mean that α and β were on the same line through the origin, so $L(\alpha, \beta)$ would not be a lattice.

When $\Delta < 0$ the lattice γL is a rotation and rescaling of L, but for $\Delta > 0$ the geometric relation between the two lattices is not as simple. There is however a simple formula relating the norms of L and γL, valid for both positive and negative discriminants:

Proposition 8.14. $N(\gamma L) = |N(\gamma)| N(L)$.

The absolute value is needed when $\Delta > 0$ since norms of lattices are always positive but $N(\gamma)$ can be negative when $\Delta > 0$. When $\Delta < 0$ the formula is just $N(\gamma L) = N(\gamma)N(L)$ and can be seen geometrically since multiplication by γ rescales by the distance from γ to the origin, which is $\sqrt{N(\gamma)}$, so the areas of parallelograms are multiplied by $N(\gamma)$, the square of the rescaling factor.

Proof: This is a calculation with determinants that will be easier if we regard $\mathbb{Z}[\tau]$ as a subset of $\mathbb{Q}(\sqrt{\Delta})$. Let $\gamma = p + q\sqrt{\Delta}$ and let $\alpha = a + b\sqrt{\Delta}$ and $\beta = c + d\sqrt{\Delta}$ where p, q, a, b, c, d are rational numbers. Multiplication by γ is a linear transformation of $\mathbb{Q}(\sqrt{\Delta})$:

$$(p + q\sqrt{\Delta})(x + y\sqrt{\Delta}) = (px + q\Delta y) + (qx + py)\sqrt{\Delta}$$

The matrix of this transformation is $\begin{pmatrix} p & q\Delta \\ q & p \end{pmatrix}$. Thus $\gamma\alpha$ and $\gamma\beta$ correspond to the columns of the product $\begin{pmatrix} p & q\Delta \\ q & p \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. The absolute value of the determinant of this product is therefore $N(\gamma L)$. This equals the product of the absolute values of the determinants of the two individual matrices, which is $|N(\gamma)|N(L)$ since the determinant of the first matrix in the product is $p^2 - \Delta q^2 = N(\gamma)$ and the absolute value of the determinant of the second matrix in the product is $N(L)$. □

When L is an ideal in $R_\Delta = \mathbb{Z}[\tau]$ then so is γL. This is because if α is in L and β is in R_Δ then $\beta(\gamma\alpha)$ is in γL since it equals $\gamma(\beta\alpha)$ and this is in γL since $\beta\alpha$ is in L if L is an ideal.

An important special case is when $L = R_\Delta$ so γR_Δ is the ideal consisting of all multiples of γ by elements of R_Δ. This is called the ***principal ideal*** generated by γ. The usual notation for this ideal is simply (γ), although this notation can sometimes be a little confusing since parentheses are also used in formulas for multiplication of elements. For example, in the previous paragraph we had an equality $\beta(\gamma\alpha) = \gamma(\beta\alpha)$ in which these were just elements of R_Δ, not ideals. However, this equation remains valid when $(\gamma\alpha)$ and $(\beta\alpha)$ are regarded as ideals since it is always true for principal ideals that $\delta(\varepsilon) = (\delta\varepsilon)$ so the equation of ideals $\beta(\gamma\alpha) = \gamma(\beta\alpha)$ can be written as $(\beta\gamma\alpha) = (\gamma\beta\alpha)$ which holds since $\beta\gamma = \gamma\beta$.

Since $N(R_\Delta) = 1$ the preceding proposition gives a simple relationship between the norm of an element and the norm of the ideal it generates:

Corollary 8.15. $N((\alpha)) = |N(\alpha)|$ *for each nonzero element* α *in* R_Δ. □

For negative discriminants, principal ideals $(\alpha) = \alpha R_\Delta$ have the same shape as the full lattice R_Δ. Conversely, if an ideal L in R_Δ has the same shape as R_Δ this means that $L = \alpha R_\Delta$ for some complex number α, and α has to lie in R_Δ and in fact in L since α is the element $\alpha \cdot 1$ in $\alpha R_\Delta = L$. As the examples earlier in this section show, for some negative discriminants such as $-3, -4, -7, -8$, and -11 all ideals have the same shape and hence all ideals are principal ideals, while for other negative discriminants there can exist nonprincipal ideals since not all ideals have the same shape as the principal ideals.

We have been focusing on the ideals $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ associated to quadratic forms $Q(x,y) = ax^2 + bxy + cy^2$ of discriminant Δ, and it is natural to ask whether every ideal in R_Δ is equal to L_Q for some form Q of discriminant Δ. One way to see that this is not true is to observe that the lattices $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ have the special property that they contain an element $\frac{b+\sqrt{\Delta}}{2}$ lying in the first row of the lattice R_Δ above the x-axis, but this is not the case for all ideals since we can expand an ideal L_Q by a positive integer factor n to get a new ideal nL_Q which has no elements in the first row of R_Δ above the x-axis if $n > 1$. However, nothing more complicated than this can happen:

Proposition 8.16. *Every ideal in* R_Δ *is equal to* nL_Q *for some positive integer* n *and some form* $Q(x,y) = ax^2 + bxy + cy^2$ *of discriminant* Δ *with* $a > 0$.

Since an ideal L_Q has an element in the first row of R_Δ above the x-axis it cannot be a multiple nL of any other ideal L with $n > 1$. We call an ideal with this property a ***primitive ideal***, in analogy with the definition of a primitive form. The proposition says that all ideals are positive integer multiples of primitive ideals, and the primitive ideals are just the ideals L_Q coming from forms.

Proof: We write R_Δ as $\mathbb{Z}[\tau]$ as before. Let L be an ideal in $\mathbb{Z}[\tau]$. Since L is a lattice it has a reduced basis $p, m + n\tau$. Then $p\tau$ lies in L since p does. Since $p\tau$ is in the p^{th} row of $\mathbb{Z}[\tau]$ we must have $p = an$ for some positive integer a. For $\alpha = m + n\tau$ the product $\alpha\tau$ must also lie in L. In the case $\Delta = 4D$ we have $\tau = \sqrt{D}$ so $\alpha\tau = m\tau + n\tau^2 = m\tau + nD$. This is in the m^{th} row of $\mathbb{Z}[\tau]$ so n must divide m, say $m = nq$. In the case $\Delta = 4d + 1$ we have $\tau^2 = \tau + d$ so $\alpha\tau = (m + n)\tau + nd$.

This is in the $(m+n)^{th}$ row of $\mathbb{Z}[\tau]$ so n divides $m+n$ and hence also m so we can again write $m = nq$. Thus $L = L(p, m+n\tau) = L(na, nq+n\tau) = nL(a, q+\tau)$. Here $L(a, q+\tau)$ is an ideal since $nL(a, q+\tau)$ is an ideal.

To finish the proof we would like to find integers b and c such that $q+\tau = \frac{b+\sqrt{\Delta}}{2}$ and $\Delta = b^2 - 4ac$ since $L(a, q+\tau)$ will then be L_Q for $Q = ax^2 + bxy + cy^2$ with discriminant Δ. Consider first the case $\Delta = 4D$ so $q+\tau = q+\sqrt{D}$. This is an element of the ideal $L(a, q+\sqrt{D})$ so if we multiply it by its conjugate $q+\overline{\tau} = q-\sqrt{D}$ we get an integer lying in $L(a, q+\sqrt{D})$. This integer must be a multiple of a, the smallest positive integer in $L(a, q+\sqrt{D})$, so we have $(q+\tau)(q+\overline{\tau}) = (q+\sqrt{D})(q-\sqrt{D}) = q^2 - D = ac$ for some integer c. Hence $(2q)^2 - 4D = 4ac$, and since $4D = \Delta$ this can be rewritten as $\Delta = b^2 - 4ac$ for $b = 2q$. We also have $q+\tau = q+\sqrt{D} = \frac{b+\sqrt{\Delta}}{2}$ so the case $\Delta = 4D$ is finished.

In the other case $\Delta = 4d+1$ we again look at the product $(q+\tau)(q+\overline{\tau})$. By the same reasoning as in the first case this must be a multiple of a, so $(q+\tau)(q+\overline{\tau}) = ac$ for some integer c. Writing this out, we have $(q+\frac{1+\sqrt{\Delta}}{2})(q+\frac{1-\sqrt{\Delta}}{2}) = ac$. Multiplying this equation by 4 gives $(2q+1+\sqrt{\Delta})(2q+1-\sqrt{\Delta}) = 4ac$ which simplifies to $(2q+1)^2 - \Delta = 4ac$. Thus if we take $b = 2q+1$ we have $\Delta = b^2 - 4ac$ and $q+\tau = q+\frac{1+\sqrt{\Delta}}{2} = \frac{b+\sqrt{\Delta}}{2}$. This finishes the case $\Delta = 4d+1$. □

The preceding proposition allows us to relate norms of ideals to the representation problem for forms. As we know, the numbers represented by the principal form of discriminant Δ are just the norms of primitive elements of R_Δ. If we now consider all forms, not just the principal form, then there is an analogous statement for norms of ideals in R_Δ:

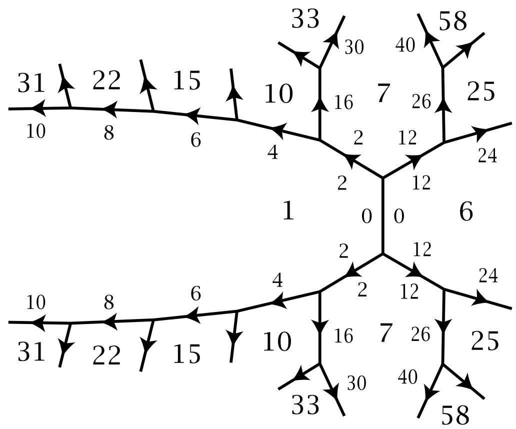
Proposition 8.17. *The positive numbers represented by forms of discriminant Δ are exactly the norms of primitive ideals in R_Δ. More specifically, the positive numbers represented by a form Q are exactly the norms of ideals $L_{Q'}$ associated to forms Q' equivalent to Q.*

Since the norms of arbitrary ideals are just squares times the norms of primitive ideals, it follows that the norms of all ideals are just the positive values of all forms of the given discriminant.

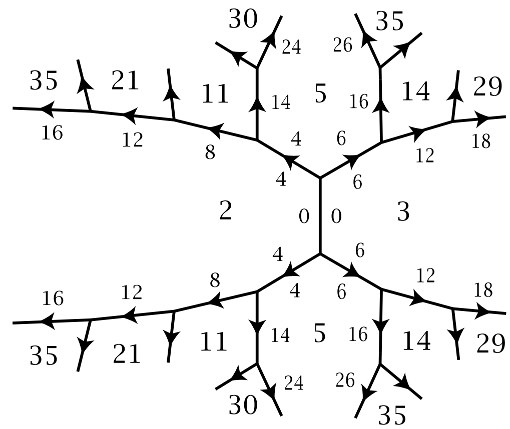
Proof: If a positive number a is represented by a form of discriminant Δ then this form is equivalent to a form $ax^2 + bxy + cy^2$. The associated ideal $L(a, \frac{b+\sqrt{\Delta}}{2})$ has norm a and is primitive. Thus all positive represented numbers are norms of primitive ideals. Conversely, by Proposition 8.16 every primitive ideal can be written as the ideal $L(a, \frac{b+\sqrt{\Delta}}{2})$ associated to a form $ax^2+bxy+cy^2$ of discriminant Δ with $a > 0$. This form represents a and the ideal $L(a, \frac{b+\sqrt{\Delta}}{2})$ has norm a, so all norms of primitive ideals are represented by forms. □

Let us look at an example, the case $\Delta = -24$ with $R_\Delta = \mathbb{Z}[\sqrt{-6}]$. Here the class number is 2 corresponding to the forms $x^2 + 6y^2$ and $2x^2 + 3y^2$.

$x^2 + 6y^2$



$2x^2 + 3y^2$



To each form $ax^2 + bxy + cy^2$ of discriminant -24 we have the associated primitive ideal $L(a, \frac{b+\sqrt{-24}}{2}) = L(a, \frac{b}{2} + \sqrt{-6})$ of norm a. This corresponds to a region labeled a in one of the two topographs, with b the label on one of the edges bordering this region. The sign of b depends on the orientation of this edge, and in the topographs shown above we have oriented the edges to make all edge labels positive. We could instead orient the edges surrounding the a region so that their labels form an arithmetic progression with increment $2a$ when traversed in the clockwise direction around the border of the a region. Then there is a unique edge such that $0 \le b < 2a$, or equivalently $0 \le \frac{b}{2} < a$, which is exactly the condition for the basis $a, \frac{b}{2} + \sqrt{-6}$ to be a reduced basis for the ideal $L(a, \frac{b}{2} + \sqrt{-6})$. Thus there is an exact one-to-one correspondence between primitive ideals and regions in the two topographs since any two regions with the same a and b labels must be related by an orientation-preserving symmetry of the topograph, but these topographs have only mirror symmetry.

For example, ideals of norm 5 correspond to regions labeled 5 in the two topographs, and there are just two of these, both in the second topograph, with the upper region corresponding to $L(5, 2 + \sqrt{-6})$ (from the edge labeled 4) and the lower region corresponding to $L(5, 3 + \sqrt{-6})$ (from the edge labeled 6). Thus these are the only two ideals of norm 5. These two ideals are conjugate since the conjugate of $L(5, 2 + \sqrt{-6})$ is $L(5, 2 - \sqrt{-6}) = L(5, -2 + \sqrt{-6}) = L(5, 3 + \sqrt{-6})$. This happens generally for all regions in the topographs, as conjugate ideals are obtained by reflecting across the horizontal line of symmetry of the topographs. The two regions in each topograph that intersect the symmetry line correspond to ideals that equal their conjugate, namely $L(1, \sqrt{-6}) = \mathbb{Z}[\sqrt{-6}]$ and $L(6, \sqrt{-6}) = (\sqrt{-6})$ for the first topograph, and $L(2, \sqrt{-6})$ and $L(3, \sqrt{-6})$ for the second topograph.

Nonprimes can appear more than twice in the topographs, as happens for 35 which appears four times. From these regions we can read off the four ideals of norm 35. In the upper half of the second topograph the two regions labeled 35 give the ideals $L(35, 8 + \sqrt{-6})$ and $L(35, 13 + \sqrt{-6})$ and in the lower half of the topograph we have their conjugates $L(35, 27 + \sqrt{-6})$ and $L(35, 22 + \sqrt{-6})$.

The ideals corresponding to regions in the first topograph are principal ideals since the form here is the norm form $N(x + y\sqrt{-6}) = x^2 + 6y^2$. For example the label 25 in the upper right is the norm of the ideal $L(25, 13 + \sqrt{-6})$, from the edge labeled 26, and similarly the label 25 in the lower right is the norm of the ideal $L(25, 12 + \sqrt{-6})$. These two regions correspond to the fractions $x/y = \pm 1/2$ so 25 is the norm of $1+2\sqrt{-6}$ and $1-2\sqrt{-6}$, hence also of the principal ideals $(1+2\sqrt{-6})$ and $(1 + 2\sqrt{-6})$. The principal ideal (5) has norm 25 as well but is not a primitive ideal. The ideal $L(25, 13 + \sqrt{-6})$, being primitive, must therefore be either $(1 + 2\sqrt{-6})$ or $(1 - 2\sqrt{-6})$. To decide which, we need to determine which of the two principal ideals contains 25 and $13 + \sqrt{-6}$. They both contain 25 since $25 = (1 + 2\sqrt{-6})(1 - 2\sqrt{-6})$ so we need to determine whether $13 + \sqrt{-6}$ is a multiple of $1 + 2\sqrt{-6}$ or of $1 - 2\sqrt{-6}$ by an element of $\mathbb{Z}[\sqrt{-6}]$. This is done by computing the relevant quotients:

$$\frac{13 + \sqrt{-6}}{1 + 2\sqrt{-6}} = \frac{13 + \sqrt{-6}}{1 + 2\sqrt{-6}} \cdot \frac{1 - 2\sqrt{-6}}{1 - 2\sqrt{-6}} = \frac{25 - 25\sqrt{-6}}{25} = 1 - \sqrt{-6}$$
$$\frac{13 + \sqrt{-6}}{1 - 2\sqrt{-6}} = \frac{13 + \sqrt{-6}}{1 - 2\sqrt{-6}} \cdot \frac{1 + 2\sqrt{-6}}{1 + 2\sqrt{-6}} = \frac{1 + 27\sqrt{-6}}{25}$$

This last quotient is not in $\mathbb{Z}[\sqrt{-6}]$ so we conclude that $L(25, 13+\sqrt{-6})$ is the principal ideal $(1 + 2\sqrt{-6})$. Taking conjugates gives $L(25, 12 + \sqrt{-6}) = (1 - 2\sqrt{-6})$.

For most negative discriminants the same one-to-one correspondence holds between primitive ideals and regions in the topographs for that discriminant, where for topographs without mirror symmetry we should take both the topograph itself and its mirror image topograph. The only exceptional negative discriminants are $\Delta = -3$ and $\Delta = -4$, the two cases when the topographs have orientation-preserving symmetries. In these cases the regions that correspond to each other under orientation-preserving symmetries correspond to a single primitive ideal. For positive discriminants the situation is very similar, the only differences being that one only considers regions in the topographs with positive labels, and then the primitive ideals correspond to regions within one period of the periodic topograph since the orientation-preserving symmetries are just the translations along the periodic separator line.

As we saw in Chapter 6, a key part of the problem of determining which numbers are represented by forms of a given discriminant is determining which primes are represented. The corresponding problem for ideals is to determine which primes p are norms of ideals in R_Δ. These ideals must be primitive, the ideals $L(p, \frac{b+\sqrt{\Delta}}{2})$ for Δ a square mod $4p$, namely $\Delta \equiv b^2 \bmod 4p$, coming from the equation $\Delta = b^2 - 4ac$ with $a = p$.

In Proposition 6.15 we saw that if a prime p is represented by a form of discriminant Δ then this form is unique up to equivalence. Furthermore, by Proposition 6.16

all the appearances of p in a topograph are images of each other under symmetries of the topograph. This means that there are at most two ideals in R_Δ of norm p, the ideal $L(p, \frac{b+\sqrt{\Delta}}{2})$ and its conjugate $L(p, \frac{b-\sqrt{\Delta}}{2}) = L(p, \frac{-b+\sqrt{\Delta}}{2})$. When the ideal and its conjugate are equal there is only one ideal of norm p.

Proposition 8.18. (a) *The ideals in R_Δ of prime norm p with p odd are:*

- *For $\Delta = 4d$, the ideal $L(p, B+\sqrt{d})$ and its conjugate $L(p, -B+\sqrt{d})$, where $d \equiv B^2$ mod p.*
- *For Δ odd, the ideal $L(p, B+\frac{1+\sqrt{\Delta}}{2})$ and its conjugate $L(p, -B-1+\frac{1+\sqrt{\Delta}}{2})$, where $\Delta \equiv (2B+1)^2$ mod p.*

(b) *The ideals in R_Δ of norm 2 are:*

- *For $\Delta = 4d$ with d even, the ideal $L(2, \sqrt{d})$.*
- *For $\Delta = 4d$ with d odd, the ideal $L(2, 1+\sqrt{d})$.*
- *For $\Delta = 8k+1$, the ideal $L(2, \frac{1+\sqrt{\Delta}}{2})$ and its conjugate $L(2, 1+\frac{1+\sqrt{\Delta}}{2})$.*

(c) *An ideal of prime norm p equals its conjugate if and only if p divides Δ.*

Proof: The condition for p to be the norm of an ideal in R_Δ is that $\Delta \equiv b^2$ mod $4p$ for some integer b, and the ideal is then $L(p, \frac{b+\sqrt{\Delta}}{2})$. If $\Delta = 4d$ then b must be even so $b = 2B$ for some integer B. The congruence $\Delta \equiv b^2$ mod $4p$ is then equivalent to $d \equiv B^2$ mod p. The ideal in this case is $L(p, B+\sqrt{d})$. If Δ is odd then so is b and we can write $b = 2B+1$. The congruence $\Delta \equiv b^2$ mod $4p$ is then $\Delta \equiv (2B+1)^2$ mod $4p$. This implies $\Delta \equiv (2B+1)^2$ mod p and the converse is also true since $\Delta \equiv (2B+1)^2$ mod 4 when Δ is odd, both sides of this congruence being 1 mod 4. The ideal $L(p, \frac{b+\sqrt{\Delta}}{2})$ is then $L(p, B+\frac{1+\sqrt{\Delta}}{2})$. This finishes part (a).

When $p = 2$ the congruence $\Delta \equiv b^2$ mod $4p$ becomes $\Delta \equiv b^2$ mod 8 which is solvable just when $\Delta \equiv 0, 1, 4$ mod 8, with solutions $b = 0, 1, 2$. This gives the ideals in part (b). (The first two ideals equal their conjugates so there is no need to include their conjugates.)

For part (c) the condition for $L(p, \frac{b+\sqrt{\Delta}}{2})$ to equal its conjugate is that p divides b, by Proposition 8.12. When p is prime this is equivalent to p dividing Δ since $\Delta = b^2 - 4pc$. ☐

We have seen how to go from a quadratic form Q to an ideal L_Q, and it will be useful to go in the opposite direction as well, from an ideal L in R_Δ to a quadratic form Q_L of discriminant Δ. As motivation we can start with the earlier formula $aQ(x,y) = N(ax + \frac{b+\sqrt{\Delta}}{2}y)$ which says that, up to the constant factor a, the form $Q(x,y) = ax^2 + bxy + cy^2$ can be obtained by restricting the usual norm in R_Δ to the elements $ax + \frac{b+\sqrt{\Delta}}{2}y$ in the ideal L_Q. We can try the same thing for any lattice $L = L(\alpha, \beta)$ in R_Δ, defining a quadratic form by:

$$Q(x,y) = N(\alpha x + \beta y) = (\alpha x + \beta y)(\overline{\alpha}x + \overline{\beta}y) = \alpha\overline{\alpha}x^2 + (\alpha\overline{\beta} + \overline{\alpha}\beta)xy + \beta\overline{\beta}y^2$$

Here the coefficients of x^2, xy, and y^2 are integers since they are equal to their conjugates. The form Q depends on the choice of the basis α, β for L. Another basis α', β' can be expressed as linear combinations $\alpha' = p\alpha + q\beta$, $\beta' = r\alpha + s\beta$ with integer coefficients. Since the change of basis can be reversed, going from α', β' back to α, β, the 2×2 matrix $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ has determinant ± 1, and conversely any such matrix gives a valid change of basis for L. Changing the basis also produces a change of variables in the form $Q(x, y)$ since $N(\alpha' x + \beta' y) = N((p\alpha + q\beta)x + (r\alpha + s\beta)y) = N(\alpha(px+ry)+(\beta(qx+sy)) = Q(px+ry, qx+sy)$. Here the matrix is the transpose $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)$, with the same determinant ± 1. Thus changing the basis for L produces an equivalent form, and every equivalent form can be realized by some change of basis for L.

The form $N(\alpha x + \beta y)$ depends on the ordering for the two basis elements α and β since reversing their order interchanges x and y, which gives a mirror image topograph. We can eliminate this ambiguity by always using the positive ordering for α and β. If we only use positively ordered bases, then the change of basis matrices have determinant $+1$ since a change of basis transformation takes a positively ordered basis to a positively ordered basis if and only if its matrix has positive determinant. This is because changing a basis amounts to replacing its matrix $\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right)$ by a product $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right)$ with $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ the matrix of the change of basis. Thus if we always use positively ordered bases, the lattice L gives rise to a proper equivalence class of quadratic forms.

The norm form $N(\alpha x + \beta y)$ associated to a lattice $L = L(\alpha, \beta)$ in R_Δ might not have discriminant Δ. For example, if we replace L by $nL = L(n\alpha, n\beta)$ this multiplies the norm form by n^2 and so the discriminant is multiplied by n^4. We can always rescale a form to have any discriminant we want just by multiplying it by a suitable positive constant, but this may lead to forms with noninteger coefficients. To illustrate this potential difficulty, suppose we take $\Delta = -4$ so $R_\Delta = \mathbb{Z}[i]$. The lattice $L(2, i)$ in $\mathbb{Z}[i]$ yields the form $N(2x + iy) = 4x^2 + y^2$ of discriminant -16, but to rescale this to have discriminant -4 we would have to take the form $2x^2 + \frac{1}{2}y^2$.

Fortunately this problem does not occur if we consider only lattices that are ideals. By Proposition 8.16 each ideal L in R_Δ is equal to a multiple $nL_Q = L(na, n\frac{b+\sqrt{\Delta}}{2})$ for some form $Q(x, y) = ax^2 + bxy + cy^2$ of discriminant Δ with $a > 0$. We have $aQ(x, y) = N(ax + \frac{b+\sqrt{\Delta}}{2}y)$, hence $n^2 aQ(x, y) = N(nax + n\frac{b+\sqrt{\Delta}}{2}y)$ which is the norm form for L in the basis $na, n\frac{b+\sqrt{\Delta}}{2}$. This basis is positively ordered since $a > 0$. By dividing this norm form for L by $n^2 a$ we get a form with integer coefficients and discriminant Δ, namely the form Q. If we change the basis $na, n\frac{b+\sqrt{\Delta}}{2}$ for L to some other positively ordered basis α, β it is still true that the form $\frac{1}{n^2 a}N(\alpha x + \beta y)$ has integer coefficients and discriminant Δ since this just changes Q to a properly equivalent form.

Note that the scaling factor n^2a is the norm $N(L)$ of the ideal $L = nL(a, \frac{b+\sqrt{\Delta}}{2})$. Thus we have shown:

Proposition 8.19. *For an ideal L in R_Δ with positively ordered basis α, β the form $\frac{1}{N(L)}N(\alpha x + \beta y)$ has integer coefficients and discriminant Δ.* □

For an ideal L with positively ordered basis α, β the form $\frac{1}{N(L)}N(\alpha x + \beta y)$ will be denoted by Q_L, although a more precise notation might include α and β since the form depends on the choice of basis.

Different ideals L in R_Δ can give properly equivalent forms Q_L. Obviously a rescaling nL of L gives the same form $Q_{nL} = Q_L$. More generally, suppose we multiply all elements of an ideal $L = L(\alpha, \beta)$ by a fixed nonzero element γ of R_Δ to get a new ideal $\gamma L = L(\gamma\alpha, \gamma\beta)$. Taking norms, we have $N(\gamma\alpha x + \gamma\beta y) = N(\gamma)N(\alpha x + \beta y)$, so if $N(\gamma) > 0$ the new form $N(\gamma\alpha x + \gamma\beta y)$ is just a rescaling of $N(\alpha x + \beta y)$, with rescaling factor $N(\gamma)$. Thus after rescaling to get discriminant Δ we have $Q_{\gamma L} = Q_L$ when $N(\gamma) > 0$. Specifically, if we use the formula $N(\gamma L) = |N(\gamma)|N(L)$ then when $N(\gamma) > 0$ we have:

$$\frac{N(\gamma\alpha x + \gamma\beta y)}{N(\gamma L)} = \frac{N(\gamma)N(\alpha x + \beta y)}{N(\gamma)N(L)} = \frac{N(\alpha x + \beta y)}{N(L)}$$

As a technical point, we should check that $\gamma\alpha, \gamma\beta$ is positively ordered if α, β is positively oriented. When $\Delta < 0$ this is automatic since multiplication by γ just rotates and rescales the plane. When $\Delta > 0$ we can argue as follows. As we saw in the proof of Proposition 8.14, multiplication in $\mathbb{Q}(\sqrt{\Delta})$ by a fixed element $\gamma = p + q\sqrt{\Delta}$ is a linear transformation with matrix $\begin{pmatrix} p & q\Delta \\ q & p \end{pmatrix}$. This has determinant $p^2 - \Delta q^2 = N(p + q\sqrt{\Delta})$, so if $N(\gamma) > 0$ the matrix corresponding to the basis $\gamma\alpha, \gamma\beta$ has positive determinant exactly when the matrix corresponding to α, β has positive determinant.

When $\Delta < 0$ we always have $N(\gamma) > 0$, but when $\Delta > 0$ it is possible to have $N(\gamma) < 0$. In this case the form $N(\gamma\alpha x + \gamma\beta y)$ is the negative of a rescaling of $N(\alpha x + \beta y)$ and the basis $\gamma\alpha, \gamma\beta$ is oppositely ordered from α, β, so $Q_{\gamma L}$ is the negative of the mirror image form of Q_L.

Since the forms Q_L and $Q_{\gamma L}$ are properly equivalent when $N(\gamma) > 0$, we would like to regard the ideals L and γL as being equivalent. Any reasonable notion of equivalence should have the property that two things equivalent to the same thing are equivalent to each other, but this does not seem to hold for the notion of equivalence that we just considered since if two ideals L and L' are equivalent to the same ideal $\gamma L = \gamma' L'$ for some γ and γ' in R_Δ, then it does not follow that $L' = \delta L$ or $L = \delta L'$ for some δ in R_Δ since the quotients γ/γ' and γ'/γ might not lie in R_Δ.

To avoid this difficulty we define two ideals L and L' in R_Δ to be ***equivalent***, written $L \sim L'$, if $\gamma L = \gamma' L'$ for some nonzero elements γ, γ' in R_Δ. If in addition $N(\gamma) > 0$ and $N(\gamma') > 0$ then we say L and L' are ***strictly equivalent*** and write $L \approx L'$. In particular we have $L \sim \gamma L$ for each nonzero γ in R_Δ since if we let $L' = \gamma L$

and $\gamma' = 1$ then the equation $\gamma L = \gamma' L'$ becomes just $\gamma L = L'$. Similarly, $L \approx \gamma L$ for every γ with $N(\gamma) > 0$.

Conversely, a general equivalence $L \sim L'$ can be realized as a pair of equivalences of the special type originally considered, namely $L \sim \gamma L = \gamma' L' \sim L'$ and likewise for strict equivalences. Thus we have not really changed the underlying idea by defining the two kinds of equivalence $\sim$ and $\approx$ as we did. What we have gained is the property that two things equivalent to the same thing are equivalent to each other, which can be expressed as the assertion that if $L \sim L'$ and $L' \sim L''$ then $L \sim L''$. This holds since if $\gamma L = \gamma' L'$ and $\delta L' = \delta' L''$ then $\delta\gamma L = \delta\gamma' L' = \delta'\gamma' L''$ so $L \sim L''$. This reasoning also works with $\approx$ in place of $\sim$ by adding the condition that all of $\gamma, \gamma', \delta, \delta'$ have positive norm, hence all their products have positive norm as well.

For negative discriminants there is no difference between equivalence and strict equivalence of ideals since norms of nonzero elements of R_Δ are always positive, but for positive discriminants there can be a difference. This happens for example when $\Delta = 12$. Here the two forms $x^2 - 3y^2$ and $3x^2 - y^2$ correspond to the ideals $(1, \sqrt{3}) = (1)$ and $(3, \sqrt{3}) = (\sqrt{3})$ in $R_\Delta = \mathbb{Z}[\sqrt{3}]$. These two ideals are equivalent since $(\sqrt{3}) = \gamma(1)$ for $\gamma = \sqrt{3}$. However, $N(\sqrt{3}) = -3$ so this does not show the ideals are strictly equivalent. In fact they are not strictly equivalent since if they were, then the forms $x^2 - 3y^2$ and $3x^2 - y^2$ would be properly equivalent, but this is not the case as one can see from their topographs or from the fact that the character χ_3 takes the value $+1$ on the first form and -1 on the second form.

This example can be contrasted with the case $\Delta = 8$ with $R_\Delta = \mathbb{Z}[\sqrt{2}]$. Here the two forms $x^2 - 2y^2$ and $2x^2 - y^2$ correspond to the ideals $(1, \sqrt{2}) = (1)$ and $(2, \sqrt{2}) = (\sqrt{2})$. Again the two ideals are equivalent since $(\sqrt{2}) = \gamma(1)$ for $\gamma = \sqrt{2}$, with $N(\sqrt{2}) = -2$. There is a unit $\varepsilon = 1 + \sqrt{2}$ of norm -1 so we have $(\sqrt{2}) = (\varepsilon\sqrt{2}) = (2 + \sqrt{2}) = \gamma(1)$ for $\gamma = 2 + \sqrt{2}$ with $N(2 + \sqrt{2}) > 0$ and hence the ideals (1) and $(\sqrt{2})$ are strictly equivalent. In fact the forms $x^2 - 2y^2$ and $2x^2 - y^2$ are properly equivalent as one can see from their topographs.

In the previous example with $\Delta = 12$ there is no unit of norm -1 since -1 is represented by the form $3x^2 - y^2$ but not by the norm form $x^2 - 3y^2$. As we will now see, the distinction between equivalence and strict equivalence of ideals is entirely accounted for by the existence or nonexistence of units of norm -1.

Proposition 8.20. *For positive discriminants Δ the relations of equivalence and strict equivalence of ideals in R_Δ are the same if and only if there is a unit in R_Δ of norm -1.*

Note that it suffices to consider only the fundamental unit since if this has norm $+1$ then all units have norm $+1$.

Proof: Suppose there is a unit ε in R_Δ with $N(\varepsilon) = -1$ and suppose two ideals L and M are equivalent via an equality $\alpha L = \beta M$. We have $\alpha L = \varepsilon\alpha L$ so we can arrange that

$N(\alpha) > 0$ by replacing α with $\varepsilon\alpha$ if necessary. In the same way we can arrange that $N(\beta) > 0$. Thus L and M are strictly equivalent.

For the converse, suppose equivalence is the same as strict equivalence. Since we assume $\Delta > 0$, there exist elements α in R_Δ with $N(\alpha) < 0$. The ideals R_Δ and αR_Δ are equivalent so by hypothesis they are strictly equivalent. This means $\beta R_\Delta = \gamma\alpha R_\Delta$ for some elements β and γ in R_Δ of positive norm. Since β is in $\beta R_\Delta = \gamma\alpha R_\Delta$ we have $\beta = \gamma\alpha\delta$ for some δ in R_Δ. Also $\gamma\alpha$ is in $\gamma\alpha R_\Delta = \beta R_\Delta$ so $\gamma\alpha = \beta\varepsilon$ for some ε in R_Δ. Thus $\beta = \gamma\alpha\delta = \beta\varepsilon\delta$ and hence $1 = \varepsilon\delta$ since $\beta \neq 0$. Thus δ and ε are units. The equation $\gamma\alpha = \beta\varepsilon$ implies that $N(\varepsilon) < 0$ since $N(\gamma) > 0$, $N(\alpha) < 0$, and $N(\beta) > 0$. Since ε is a unit, its norm is then -1. □

Now we come to the main result in this section:

Theorem 8.21. *There is a one-to-one correspondence between the set of strict equivalence classes of ideals in R_Δ and the set of proper equivalence classes of quadratic forms of discriminant Δ. Under this correspondence an ideal L with a positively ordered basis α, β corresponds to the form $Q_L(x,y) = \frac{1}{N(L)}N(\alpha x+\beta y)$, and a form $Q(x,y) = ax^2 + bxy + cy^2$ with $a > 0$ corresponds to the ideal $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$. (When $\Delta < 0$ we are considering only forms with positive values, as usual.)*

For example, when all forms of discriminant Δ are equivalent and hence properly equivalent, the theorem says that all ideals are strictly equivalent. When $\Delta < 0$ this is saying that all ideals have the same shape, or equivalently that all ideals are principal ideals. The negative discriminants for which this happens are $-3, -4, -7, -8, -11$, $-19, -43, -67$, and -163. For the first five of these we already saw that all ideals have the same shape using a geometric argument, but that argument does not apply in the last four cases.

The condition $a > 0$ in the theorem plays a role only when $\Delta > 0$, but its role is sometimes important. For example, the principal form $x^2 + bxy + cy^2$ corresponds to the ideal $L(1, \frac{b+\sqrt{\Delta}}{2})$ which equals R_Δ since it contains 1, but without the condition $a > 0$ the negative of the principal form would correspond to $L(-1, \frac{-b+\sqrt{\Delta}}{2})$ which also equals R_Δ since it contains -1. However, for some values of Δ such as $\Delta = 12$ the principal form is not equivalent to its negative.

Proof: Let Φ be the function from the set of strict equivalence classes of ideals to the set of proper equivalence classes of forms induced by sending an ideal L with a positively ordered basis α, β to the form $Q(x,y) = N(\alpha x + \beta y)/N(L)$. The function Φ is well defined since we have seen that changing one positively ordered basis for L to another changes the associated form to a properly equivalent form, and replacing L with basis α, β by γL with basis $\gamma\alpha, \gamma\beta$ leaves the form unchanged when $N(\gamma) > 0$.

To see that Φ is onto, note first that in each proper equivalence class of forms there are forms $Q(x,y) = ax^2 + bxy + cy^2$ with $a > 0$ since the topograph of an elliptic or hyperbolic form always contains some positive numbers, so we can choose

Q so that $Q(1,0) > 0$. Then $Q = Q_L$ for the ideal $L = L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ since $Q_L = N(ax + \frac{b+\sqrt{\Delta}}{2}y)/N(L) = ax^2 + bxy + cy^2$, using the fact that $N(L) = a$.

To show that Φ is one-to-one, suppose we have two ideals L and L' with positively oriented bases α, β and α', β' such that the associated forms Q_L and $Q_{L'}$ with respect to these bases are properly equivalent. We can assume the basis α, β is chosen so that $Q_L(1,0) > 0$. Since Q_L and $Q_{L'}$ are properly equivalent we can then choose α', β' so that we have actual equality $Q_L(x,y) = Q_{L'}(x,y)$ for all x and y. We have $N(\alpha) = Q_L(1,0)\cdot N(L) > 0$ and $N(\alpha') = Q_{L'}(1,0)\cdot N(L') > 0$ since $Q_L(1,0) = Q_{L'}(1,0) > 0$.

The forms $N(\alpha x + \beta y)$ and $N(\alpha' x + \beta' y)$ are rescalings of each other since they rescale to the same form $Q_L(x,y) = Q_{L'}(x,y)$. Let $\gamma = \beta/\alpha$ and $\gamma' = \beta'/\alpha'$, elements of $\mathbb{Q}(\sqrt{\Delta})$. We have $N(\alpha x + \beta y) = N(\alpha)N(x + \gamma y)$ and $N(\alpha' x + \beta' y) = N(\alpha')N(x+\gamma' y)$ so the two forms $N(x+\gamma y) = N(\alpha x+\beta y)/N(\alpha)$ and $N(x+\gamma' y) = N(\alpha' x + \beta' y)/N(\alpha')$ are also rescalings of each other. Note that these two forms have rational coefficients, not necessarily integers. Since the forms $N(x + \gamma y)$ and $N(x + \gamma' y)$ are rescalings of each other and take the same value at $(x,y) = (1,0)$, namely $N(1) = 1$, they must actually be equal.

Next we show that in fact $\gamma = \gamma'$. Let $\gamma = r + s\sqrt{\Delta}$ and $\gamma' = r' + s'\sqrt{\Delta}$ with r, s, r', s' in $\mathbb{Q}$. We have $N(x + \gamma y) = N(x + \gamma' y)$ for all integers x and y so in particular $N(\gamma) = N(\gamma')$ which means $r^2 - s^2\Delta = r'^2 - s'^2\Delta$. Also $N(1+\gamma) = N(1+\gamma')$ so the difference $N(1 + \gamma) - N(\gamma) = ((r+1)^2 - s^2\Delta) - (r^2 - s^2\Delta) = 2r + 1$ equals the difference $N(1 + \gamma') - N(\gamma') = 2r' + 1$ and hence $r = r'$. From the earlier equation $r^2 - s^2\Delta = r'^2 - s'^2\Delta$ we then get $s = \pm s'$. The bases $1, \gamma$ and $1, \gamma'$ are positively ordered since this was true for α, β and α', β' and multiplication by α and α' preserves orientation of the plane since $N(\alpha) > 0$ and $N(\alpha') > 0$. Since both $1, \gamma$ and $1, \gamma'$ are positively ordered we must have $s > 0$ and $s' > 0$ so $s = s'$. Thus $\gamma = \gamma'$ as claimed.

The lattice $L(1,\gamma)$ may not lie in R_Δ since γ is only an element of $\mathbb{Q}(\sqrt{\Delta})$, but we can rescale $L(1,\gamma)$ to a lattice $nL(1,\gamma) = L(n, n\gamma)$ in R_Δ by multiplying by a positive integer n such that $n\gamma$ is in R_Δ. Using the symbol $\approx$ to denote strict equivalence of ideals, we then have:

$$L = L(\alpha,\beta) \approx nL(\alpha,\beta) = L(n\alpha, n\beta) = L(n\alpha, n\alpha\gamma) = \alpha L(n, n\gamma) \approx L(n, n\gamma)$$

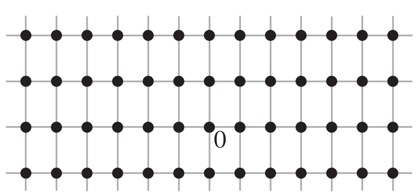
Similarly, $L' \approx L(n', n'\gamma')$ for some positive integer n', but we can choose $n' = n$ since $\gamma = \gamma'$. Thus both L and L' are strictly equivalent to $L(n, n\gamma)$ so they are strictly equivalent to each other. This finishes the proof that Φ is one-to-one. □

To illustrate the theorem consider the case $\Delta = 60$ where there are four proper equivalence classes of forms, given by $x^2 - 15y^2$, $15x^2 - y^2$, $3x^2 - 5y^2$, and $5x^2 - 3y^2$. The corresponding ideals in $R_\Delta = \mathbb{Z}[\sqrt{15}]$ are $(1, \sqrt{15}) = (1)$, $(15, \sqrt{15}) = (\sqrt{15})$, $(3, \sqrt{15})$, and $(5, \sqrt{15})$. According to the theorem no two of these ideals are strictly equivalent, although the first two are equivalent since $(\sqrt{15}) = \sqrt{15}(1)$ and the second two are equivalent since $\sqrt{15}(3,\sqrt{15}) = 3(5,\sqrt{15})$. This corresponds to the fact that the two forms in each pair are negative mirror images of each other, although all four forms have mirror symmetry so taking mirror images makes no difference.

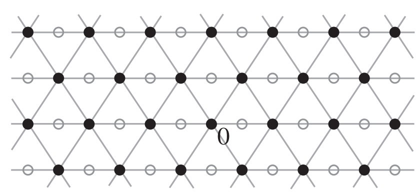
For another example take $\Delta = 136$ with class number 4 realized by the forms x^2-34y^2, $34x^2-y^2$, and $3x^2\pm 2xy-11y^2$ as we saw in an example in Section 7.4 that displayed an interesting combination of symmetry and skew symmetry properties. In $R_\Delta = \mathbb{Z}[\sqrt{34}]$ the four forms correspond to the ideals $(1,\sqrt{34}) = (1)$, $(34,\sqrt{34}) = (\sqrt{34})$, and $(3, 1\pm\sqrt{34})$. The first two ideals are obviously equivalent. For the second two, if we multiply $(3, 1+\sqrt{34})$ by some γ with $N(\gamma) < 0$, for example $\gamma = \sqrt{34}$, we get an ideal corresponding to the negative mirror image of the form $3x^2+2xy-11y^2$. The topograph of this form has rotational skew symmetries but no mirror symmetries, so its negative mirror image is $3x^2-2xy-11y^2$. Thus $\sqrt{34}(3,1+\sqrt{34})$ must be strictly equivalent to $(3,1-\sqrt{34})$, so $(3,1+\sqrt{34})$ and $(3,1-\sqrt{34})$ are equivalent but not strictly equivalent. This is true also for the other two ideals (1) and $(\sqrt{34})$ but for a different reason since the forms x^2-34y^2 and $34x^2-y^2$ have mirror symmetry but no skew symmetries rather than vice versa.

The correspondence between forms and ideals includes nonprimitive forms as well as primitive forms, but the ideals corresponding to primitive and nonprimitive forms behave somewhat differently. Let us illustrate this by the example of discriminant $\Delta = -12$ where there are two equivalence classes of forms, given by the primitive form x^2+3y^2 and the nonprimitive form $2x^2+2xy+2y^2$.

$x^2+3y^2 \longleftrightarrow L(1,\sqrt{-3})$



$2x^2+2xy+2y^2 \longleftrightarrow L(2,1+\sqrt{-3})$



The ideal for $2x^2+2xy+2y^2$ is a lattice of equilateral triangles, and this lattice has the special property that it is taken to itself not just by multiplication by elements of $R_\Delta = \mathbb{Z}[\sqrt{-3}]$ but also by the 60 degree rotation given by multiplication by the element $\omega = (1+\sqrt{-3})/2$ in the larger ring $\mathbb{Z}[\omega]$ which is R_Δ for $\Delta = -3$. Hence the lattice $L(2,1+\sqrt{-3})$ is taken to itself by all elements of $\mathbb{Z}[\omega]$ and so this lattice is an ideal in $\mathbb{Z}[\omega]$, not just in the original ring $\mathbb{Z}[\sqrt{-3}]$.

More generally, suppose we start with a form $Q = ax^2+bxy+cy^2$ of discriminant Δ and then consider the nonprimitive form $kQ = kax^2+kbxy+kcy^2$ of discriminant $k^2\Delta$ for some integer $k > 1$. The associated ideal L_{kQ} is then $L(ka, \frac{kb+k\sqrt{\Delta}}{2}) = kL(a,\frac{b+\sqrt{\Delta}}{2}) = kL_Q$. This is an ideal not just in $R_{k^2\Delta}$ but also in the larger ring R_Δ since it is k times an ideal in R_Δ, namely k times L_Q.

Let us say that an element α in $\mathbb{Q}(\sqrt{\Delta})$ ***stabilizes*** an ideal L in R_Δ if αL is contained in L, and let us call the set of all such elements α the ***stabilizer*** of L. The stabilizer of L contains R_Δ and is a ring itself since if two elements α and β in $\mathbb{Q}(\sqrt{\Delta})$ stabilize L then so do $\alpha \pm \beta$ and $\alpha\beta$. If the stabilizer of L is exactly R_Δ then we will say that L is ***stable.***

For example, principal ideals (γ) are stable since if $\alpha(\gamma)$ is contained in (γ) then in particular $\alpha\gamma$ is in (γ) and so we have $\alpha\gamma = \beta\gamma$ for some β in R_Δ. Canceling γ, we then have $\alpha = \beta$ so α is an element of R_Δ.

Proposition 8.22. *A form Q of discriminant Δ is primitive if and only if the corresponding ideal L_Q in R_Δ is stable.*

Proof: We observed above that a nonprimitive form Q of discriminant Δ gives an ideal L_Q with stabilizer larger than R_Δ. For the converse we wish to show that if $Q = ax^2 + bxy + cy^2$ is a primitive form of discriminant Δ then L_Q is not an ideal in any larger ring than R_Δ in $\mathbb{Q}(\sqrt{\Delta})$. Let us write L_Q as $L(a, \tau)$ for $\tau = \frac{b+\sqrt{\Delta}}{2}$. Note that $R_\Delta = \mathbb{Z}[\tau]$ since b has the same parity as Δ. Also $\mathbb{Q}(\sqrt{\Delta}) = \mathbb{Q}(\tau)$.

Suppose we have an element $\alpha = r + s\tau$ in $\mathbb{Q}(\tau)$ such that $\alpha L(a, \tau)$ is contained in $L(a, \tau)$. Here r and s are rational numbers. Our goal is to show that Q being primitive forces r and s to be integers. This will say that α is in $\mathbb{Z}[\tau] = R_\Delta$, and hence that R_Δ is the stabilizer of $L(a, \tau)$.

Since $\alpha L(a, \tau)$ is contained in $L(a, \tau)$, both αa and $\alpha\tau$ are in $L(a, \tau)$. We have $\alpha a = ra + sa\tau$, and for this to be in $L(a, \tau)$, which consists of the linear combinations $xa + y\tau$ with x and y integers, means that r is an integer and sa is an integer. It remains to show that $\alpha\tau$ being in $L(a, \tau)$ implies that s is an integer.

To do this we first compute $\alpha\tau$ using the fact that τ is a root of the equation $x^2 - bx + ac = 0$ so $\tau^2 = b\tau - ac$. Then we have:

$$\alpha\tau = r\tau + s\tau^2 = r\tau + s(b\tau - ac) = -sac + (r + sb)\tau$$

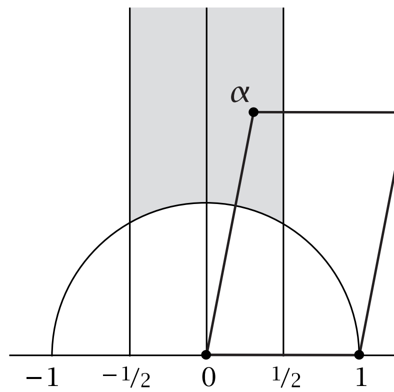
For this to be in $L(a, \tau)$ means that sc and $r + sb$ are integers. We already know that r is an integer, so $r + sb$ being an integer is equivalent to sb being an integer. Thus we know that all three of sa, sb, and sc are integers. Let us write s as a fraction $\frac{m}{n}$ in lowest terms. Then $sa = \frac{m}{n}a$ is an integer, so n must divide a. Similarly sb and sc being integers implies that n divides b and c. But 1 is the only common divisor of a, b and c since the form $ax^2 + bxy + cy^2$ is primitive, so $n = 1$. Thus s is an integer and we are done. □

A Digression on Shapes of Lattices

Let us go into a little more detail about the shapes of lattices in the plane. This will not be used in the rest of the chapter, although it does provide some enlightening context. Lattice shapes are mostly of interest for negative discriminants, but for the following discussion we will consider all possible lattices in the plane, without regard to whether they lie in some ring R_Δ or not.

Recall that we say two lattices have the same shape if one can be transformed into the other by rotation and rescaling of the plane. With this definition of shape one can ask whether it is possible to characterize exactly all the different shapes of lattices. We will give such a characterization and then see how this relates to forms of negative discriminant.

First let us get a global picture of all the possible shapes of lattices in the plane. Given a lattice L, choose a point in L that is closest to the origin, other than the origin itself. We can rotate L about the origin until this point lies on the positive x-axis, and then we can rescale L until this point is at distance 1 from the origin, so it is the point $(1,0)$, or in other words the complex number 1. Now choose a point α in L closest to the origin among all points of L above the x-axis. Thus α lies on or outside the unit circle $x^2+y^2=1$. Also, α must lie in the vertical strip consisting of points $x+yi$ with $-1/2\le x\le 1/2$, otherwise there would be another point of L inside this strip that had the same y-coordinate as α and was closer to the origin than α. This is because all points of L lie in horizontal rows of points of distance 1 apart. The lattice $L(1,\alpha)$ is contained in L and in fact must equal L by the way that we chose α. (There are no other points of L above the x-axis and inside the circle $x^2+y^2=r^2$ passing through α.)



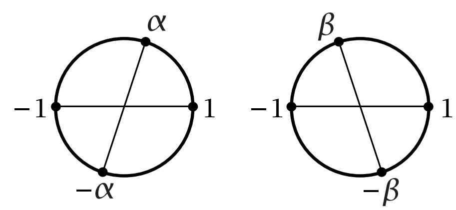
Let R be the region of the plane consisting of the points α as above, that is, all $\alpha=x+yi$ with $x^2+y^2\ge 1$, $-1/2\le x\le 1/2$, and $y>0$.

Proposition 8.23. *The lattices $L(1,\alpha)$ with α in R realize all lattice shapes, and of these lattices the only ones having the same shape are the pairs $L(1,1/2+yi)$ and $L(1,-1/2+yi)$ and the pairs $L(1,x+yi)$ and $L(1,-x+yi)$ with $x^2+y^2=1$.*

Note that these pairs all lie on the boundary of R, either on the vertical edges or on the circular arc forming the lower edge of R. The two points of each pair are mirror reflections of each other across the y-axis.

Proof: We have already seen that all lattices have the shape of a lattice $L(1,\alpha)$ for some α in R, and it remains to see when two of these lattices $L(1,\alpha)$ have the same shape. A more basic question is when two of the lattices $L(1,\alpha)$ and $L(1,\beta)$ with α and β in R are the same lattice. If this happens, the y-coordinates of α and β must be the same since this is the coordinate of points in the first row of the lattice above the x-axis. The x-coordinates of α and β must then differ by an integer if $L(1,\alpha)=L(1,\beta)$, so if α and β are both in R the only possibility is that α and β are points $1/2+yi$ and $-1/2+yi$ on the two vertical edges of R.

For $L(1,\alpha)$ and $L(1,\beta)$ to have the same shape means that there is a rotation and rescaling taking one to the other. However, there can be no rescaling since the smallest distance from nonzero points in these two lattices to the origin is 1 in both cases. To see what sorts of rotations are possible, consider the subsets C_α of $L(1,\alpha)$ and C_β of $L(1,\beta)$ consisting of the lattice points at distance 1 from the origin. If there is a rotation taking $L(1,\alpha)$ to $L(1,\beta)$ then this rotation carries C_α onto C_β. In particular, C_α and C_β must have the same number of points. The points 1 and -1 always belong to C_α and C_β. If these are the only points in C_α and C_β then the only rotations taking C_α to C_β are rotations by 0 and 180 degrees, but these do not affect the lattices so we must have $L(1,\alpha) = L(1,\beta)$ in this case. If C_α and C_β have more than two points then C_α will include $\pm\alpha$ and C_β will include $\pm\beta$. If $C_\alpha = \{\pm 1, \pm\alpha\}$ and $C_\beta = \{\pm 1, \pm\beta\}$ then the only way for C_α to be a rotation of C_β is for the two arcs in the upper half of the unit circle joining α to 1 and to -1 to have the same lengths as the two arcs from β to 1 and -1, after possibly interchanging the two arcs for α or β as in the figure. This implies that β is equal to either α or the reflection of α across the y-axis. Thus $L(1,\alpha)$ and $L(1,\beta)$ are $L(1, x+yi)$ and $L(1, -x+yi)$ for some x and y with $x^2+y^2 = 1$. The remaining possibility is that C_α and C_β contain more that four points, but this only happens when they are the vertices of regular hexagons inscribed in the unit circle since the points of C_α must be of distance at least 1 apart, and likewise for C_β. In this hexagonal case we have $L(1,\alpha) = L(1,\beta)$, finishing the proof. □



Let us see now how the lattices $L_Q = L(a, \frac{b+\sqrt{\Delta}}{2})$ associated to elliptic forms $Q = ax^2 + bxy + cy^2$ fit into this picture. Here a and c are positive since we only consider positive elliptic forms. For the two basis elements of L_Q we have $N(a) = a^2$ and $N(\frac{b+\sqrt{\Delta}}{2}) = \frac{b+\sqrt{\Delta}}{2} \cdot \frac{b-\sqrt{\Delta}}{2} = \frac{b^2-\Delta}{4} = ac$. If we assume that Q is reduced, so $0 \le b \le a \le c$, then $N(a) \le N(\frac{b+\sqrt{\Delta}}{2})$. Also the x-coordinate of $\frac{b+\sqrt{\Delta}}{2}$, which is $b/2$, is at most $a/2$. From these facts we can deduce that a is the closest point in L_Q to the origin. Then when we rescale L_Q by shrinking by a factor of a we get the lattice $L(1,\alpha)$ with $\alpha = \frac{b+\sqrt{\Delta}}{2a}$, with α lying in the right half of the region R since $N(\alpha) \le 1$ and $0 \le \frac{b}{2a} \le \frac{1}{2}$. Conversely, if $\frac{b+\sqrt{\Delta}}{2a}$ is in the right half of R then we have $0 \le b \le a \le c$. Thus Q is reduced exactly when the rescaled L_Q is $L(1,\alpha)$ with α in the right half of R.

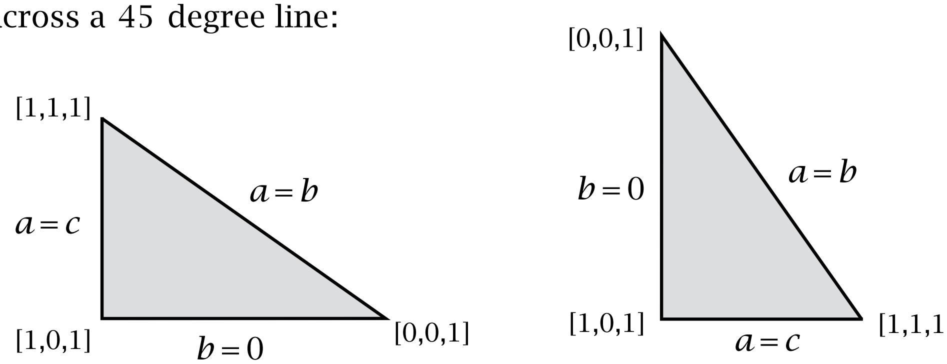
If we replace Q by nQ then L_Q is replaced by $L(na, \frac{nb+\sqrt{n^2\Delta}}{2}) = nL_Q$ so this is just a rescaling of L_Q with the same shape and hence corresponding to the same point α in R. Apart from rescaling Q in this way, different reduced forms give different points α in R since the x-coordinate $b/2a$ of α determines the ratio b/a and the norm of α gives the ratio c/a.

Any point α in the right half of R with rational x-coordinate and rational norm arises in this way from a reduced elliptic form Q. For example for an x-coordinate of $1/3$ and a norm of $5/4$ we have $b/2a = 1/3$ and $c/a = 5/4$. Rewriting these two fractions with a common denominator, we get $4/12$ and $15/12$. Then after writing $4/12$ as $8/24$ we can choose $a = 12$, $b = 8$, and $c = 15$, producing the form $12x^2 + 8xy + 15y^2$.

Points in the left half of the region R are realized by replacing b by $-b$, so the form $ax^2 + bxy + cy$ is replaced by its mirror image form $ax^2 - bxy + cy^2$ which is equivalent but not properly equivalent unless $ax^2 + bxy + cy^2$ has mirror symmetry. The reduced forms with mirror symmetric topographs are those where one of the inequalities $0 \le b \le a \le c$ becomes an equality. When $b = 0$ we have the forms $ax^2 + cy^2$ corresponding to the lattices $L(1, \frac{\sqrt{\Delta}}{2a})$ along the y-axis in R. These are the rectangular lattices, with mirror symmetry across the y-axis. When $b = a$ we have the forms $ax^2 + axy + cy^2$ whose associated lattices $L(1, \frac{a+\sqrt{\Delta}}{2a}) = L(1, \frac{1}{2} + \frac{\sqrt{\Delta}}{2a})$ lie along the right-hand edge of R. These lattices also have mirror symmetry across the y-axis since they equal their mirror image lattices $L(1, -\frac{1}{2} + \frac{\sqrt{\Delta}}{2a})$. Finally, if $a = c$ we have forms $ax^2 + bxy + ay^2$ corresponding to lattices $L(1, \frac{b+\sqrt{\Delta}}{2a})$ with $\frac{b+\sqrt{\Delta}}{2a}$ having norm $c/a = 1$ and hence lying on the arc of the unit circle forming the bottom border of R. These lattices also have mirror symmetry since they form grids of rhombuses, the distances from both basis elements 1 and $\frac{b+\sqrt{\Delta}}{2a}$ to the origin being equal.

Thus forms with mirror symmetric topographs give rise to mirror symmetric lattices. The converse is also true since none of the lattices $L(1, \alpha)$ with α in the interior of R but not on the y-axis have mirror symmetry. One can see this by noting that for points α in the interior of R the only points in lattices $L(1, \alpha)$ of unit distance apart lie on horizontal lines, so mirror symmetries of these lattices must take horizontal lines to horizontal lines, which forces these symmetries to be reflections across either horizontal or vertical lines. The only time such a reflection takes a lattice $L(1, \alpha)$ to itself for some α in the interior of R is when α is on the y-axis, so the lattice is rectangular.

It is interesting to compare the picture of the region R shown earlier with the figure in Section 5.5 showing the location of reduced elliptic forms in a triangle inside the Farey diagram. Here is this triangle, first as it appeared in Section 5.5 and then reflected across a 45 degree line:



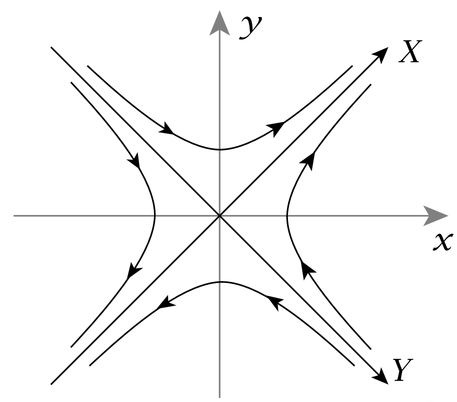
The three sides of the triangle are specified by the equations $a = c$, $a = b$, and $b = 0$, so we see that the triangle corresponds exactly to the right half of the region R, with the edge $a = b$ corresponding to the right edge of R, the edge $a = c$ to an arc of the unit circle, and the edge $b = 0$ to the central vertical axis of R.

A Digression on Hyperbolic Motions

For negative discriminants the relation of strict equivalence of ideals corresponds geometrically to rotation and rescaling of lattices. There is an analogous interpretation for positive discriminants but it involves replacing rotations by somewhat more complicated motions of the plane involving hyperbolas, as we shall now see.

What we want is a geometric description of the transformation T_γ of $\mathbb{Q}(\sqrt{\Delta})$ defined by multiplying by a fixed nonzero element γ, so $T_\gamma(\alpha) = \gamma\alpha$. For a positive discriminant Δ we are regarding $\mathbb{Q}(\sqrt{\Delta})$ as a subset of the plane by giving an element $\alpha = a + b\sqrt{\Delta}$ the coordinates $(x, y) = (a, b\sqrt{\Delta})$. The norm $N(\alpha) = a^2 - \Delta b^2$ is then equal to $x^2 - y^2$ and T_γ takes each hyperbola $x^2 - y^2 = k$ to a hyperbola $x^2 - y^2 = N(\gamma)k$ since $N(\gamma\alpha) = N(\gamma)N(\alpha)$.

To picture linear transformations of the plane that take hyperbolas $x^2 - y^2 = k$ to hyperbolas $x^2 - y^2 = k'$ it will be convenient to change the coordinates x and y to $X = x + y$ and $Y = x - y$. This changes the hyperbolas $x^2 - y^2 = k$ to the hyperbolas $XY = k$ whose asymptotes are the X-axis and the Y-axis, at a 45 degree angle from the x-axis and the y-axis. Notice that since $(x, y) = (a, b\sqrt{\Delta})$, the coordinate $X = x + y$ is just $a + b\sqrt{\Delta}$, the real number α we started with, while $Y = x - y$ is $a - b\sqrt{\Delta}$, its conjugate $\overline{\alpha}$.



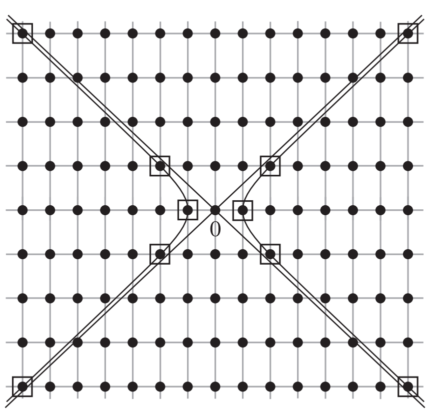
The transformation T_γ sends α to $\gamma\alpha$ so T_γ multiplies the X-coordinate α by γ. To see how T_γ acts on the Y-coordinate, observe that since the Y-coordinate of α is $\overline{\alpha}$, the Y-coordinate of $T_\gamma(\alpha)$ is $\overline{T_\gamma(\alpha)} = \overline{\gamma\alpha} = \overline{\gamma}\,\overline{\alpha}$, so the Y-coordinate of $T_\gamma(\alpha)$ is $\overline{\gamma}$ times the Y-coordinate of α. Thus T_γ multiplies the Y-coordinate by $\overline{\gamma}$, so we have the simple formula $T_\gamma(X, Y) = (\gamma X, \overline{\gamma}Y)$.

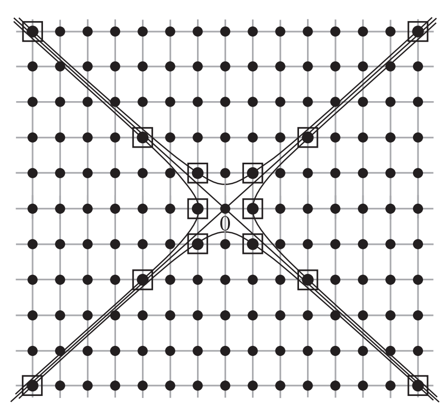
A consequence of the formula $T_\gamma(X, Y) = (\gamma X, \overline{\gamma}Y)$ is that T_γ takes the X-axis to itself since the X-axis is the points (X, Y) with $Y = 0$. Similarly, T_γ takes the Y-axis to itself, the points where $X = 0$. In general, a linear transformation that takes both the X-axis and the Y-axis to themselves has the form $T(X, Y) = (\lambda X, \mu Y)$ for real constants λ and μ. In particular when $\mu = \lambda^{-1}$ we have the transformation $T(X, Y) = (\lambda X, \lambda^{-1}Y)$ taking each hyperbola $XY = k$ to itself. When $\lambda > 1$ this transformation stretches the X-coordinate by a factor of λ and shrinks the Y-coordinate by the same factor. Thus each hyperbola $XY = k$ slides along itself in the direction indicated by the arrows in the figure above. When λ is between 0 and 1 the situation is reversed and the Y-coordinate is stretched while the X-coordinate is shrunk.

When $\lambda > 0$ and $\mu > 0$ we can rescale the transformation $T(X, Y) = (\lambda X, \mu Y)$ to $(1/\sqrt{\lambda\mu})T(X, Y) = (\sqrt{\lambda/\mu}\,X, \sqrt{\mu/\lambda}\,Y)$ which is a transformation of the type considered in the preceding paragraph, sliding each hyperbola along itself. Thus a transformation $T(X,Y) = (\lambda X, \mu Y)$ with λ and μ positive is a composition of a "hyperbola slide" and a rescaling. This is analogous to compositions of rotations and rescalings in the situation of negative discriminants. Allowing λ or μ to be negative then allows reflections across the X-axis or the Y-axis as well. If both λ and μ are negative the composition of these two reflections is a 180 degree rotation of the plane.

Now we specialize to the situation of a transformation T_γ of R_Δ given by multiplication by an element γ in R_Δ with $N(\gamma) > 0$. The condition $N(\gamma) > 0$ implies that T_γ preserves the orientation of the plane and also the sign of the norm, so it takes each quadrant of the XY-plane (north, south, east, or west) either to itself or to the opposite quadrant. In the former case T_γ is a composition of a hyperbola slide and a rescaling, while in the latter case there is also a composition with a 180 degree rotation of the plane, which is just T_γ for $\gamma = -1$. The sign of γ distinguishes these two cases since if $\gamma > 0$ the transformation T_γ takes positive numbers to positive numbers so the positive X-axis goes to itself, while if $\gamma < 0$ the positive X-axis goes to the negative X-axis.

If γ is a unit with $N(\gamma) = +1$ then each hyperbola $x^2 - y^2 = k$ is taken to itself by T_γ. The two branches of the hyperbola are distinguished by the sign of X, so if γ is positive then T_γ slides each branch along itself while if γ is negative this slide is combined with a 180 degree rotation of the plane. If we choose γ to be the smallest unit greater than 1 with $N(\gamma) = +1$ then the powers γ^n for integers n lie along the right-hand branch of the hyperbola $x^2 - y^2 = 1$, becoming farther and farther apart as one moves away from the origin, and T_γ slides each one of these points along the hyperbola to the next one, increasing the X-coordinate. The case $\Delta = 12$ is shown in the first figure below, with $R_\Delta = \mathbb{Z}[\sqrt{3}]$. The unit γ is $2 + \sqrt{3}$, and the figure shows the units $\pm\gamma^n$ for $|n| \le 2$ positioned along the two branches of the hyperbola $x^2 - y^2 = 1$, with $\gamma^2 = 7 + 4\sqrt{3}$ in the upper right corner of the figure.





For some discriminants there are units γ with $N(\gamma) = -1$ in addition to those with $N(\gamma) = +1$. The transformation T_γ for the smallest $\gamma > 1$ of norm -1 is a composition of a hyperbola slide and reflection across the X-axis. The powers γ^n then lie alternately on $x^2 - y^2 = +1$ and $x^2 - y^2 = -1$. This happens for example in

$\mathbb{Z}[\sqrt{2}]$ with $\gamma = 1 + \sqrt{2}$ as shown in the second figure above, where $\gamma^2 = 3 + 2\sqrt{2}$ and $\gamma^3 = 7 + 5\sqrt{2}$.

Each ideal in R_Δ is taken into itself by the transformations T_γ for γ in R_Δ, but when γ is a unit each ideal is taken *onto* itself since the inverse transformation $(T_\gamma)^{-1}$ is just $T_{\gamma^{-1}}$ which also takes the ideal to itself. Thus all ideals in R_Δ have "hyperbolic symmetries", the hyperbola-preserving transformations T_γ for units γ.

Although we can describe how the ideals corresponding to properly equivalent quadratic forms of positive discriminant are related in geometric terms via hyperbola slides and rescaling, the result is somehow less satisfying than in the negative discriminant case. Hyperbola slides are not nearly as simple visually as rotations, making it harder to see at a glance whether two lattices are related by hyperbola slides and rescaling or not. This may be a reflection of the fact that hyperbolic forms do not have a canonical reduced form as elliptic forms do, making it a little more difficult to determine whether two hyperbolic forms are equivalent.

Exercises

1. For discriminant $\Delta = -23$ draw the lattice L_Q for one form in each proper equivalence class of forms. Prove that no two of these lattices have the same shape by computing ratios of distances from the origin to nearby points in the lattice, with an extra argument to deal with mirror image lattices that do not have the same shape.

2. Do the same things for $\Delta = -39$.

3. (a) Given a lattice L in R_Δ and a nonzero element α in R_Δ, show that there is a positive integer multiple $n\alpha$ that is in L.

(b) Show that the intersection of two lattices in R_Δ is a lattice.

4. (a) For a form $ax^2 + bxy + cy^2$ of discriminant Δ we have the associated ideal $L(a, \frac{b+\sqrt{\Delta}}{2})$ whose basis $a, \frac{b+\sqrt{\Delta}}{2}$ determines a parallelogram P. When $\Delta < 0$ show that P is a rhombus if and only if $a = c$.

(b) Give an example of a form $ax^2 + bxy + ay^2$ with $\Delta > 0$ for which P is not a rhombus.

5. Show that the norm $N(L)$ of a lattice L in R_Δ can be computed in the following way. Choose a basis α, β for L and let $P_{\alpha,\beta}$ be the parallelogram with vertices 0, α, β, and $\alpha + \beta$. Then $N(L)$ is the total number of points of R_Δ in the interior of $P_{\alpha,\beta}$ plus the number of points of R_Δ in the interiors of two adjacent edges of $P_{\alpha,\beta}$, plus an additional 1 for the vertex of $P_{\alpha,\beta}$ between these two edges.

6. Show that if L and L' are lattices in R_Δ with L' a subset of L then $N(L)$ divides $N(L')$.

7. Show that the number of lattices in R_Δ of norm n is equal to the divisor sum $\sigma(n)$, the sum of all the divisors of n including 1 and n itself.

8. Show that $L(a,\sqrt{n})$ is an ideal in $\mathbb{Z}[\sqrt{n}]$ if and only if a divides n.

9. (a) We know that if L is an ideal in R_Δ then so is γL for each nonzero γ in R_Δ. Show the converse, that L is an ideal if γL is an ideal.
(b) Show that if γL is a principal ideal then so is L.

10. Find the four ideals in $\mathbb{Z}[\sqrt{-14}]$ of norm 15 and show that only two are principal ideals, giving explicit generators for these two. (The relevant topographs are shown in Section 6.1.)

11. (a) For $\Delta = 105$ determine all the equivalence and strict equivalence classes of ideals in R_Δ.
(b) Do the same for $\Delta = 145$.

12. For discriminant $\Delta = -64$ determine the stabilizers for all the ideals L_Q associated to reduced forms Q, whether primitive or not.

13. Show that for each ideal L in R_Δ the stabilizer of L is the same as the stabilizer of αL for each nonzero α in R_Δ.

14. Show that all ideals in R_Δ are stable if and only if Δ is a fundamental discriminant.

8.4 The Ideal Class Group

An important feature of ideals is that there is a natural way to define a multiplicative structure in the set of all ideals in R_Δ. Thus every pair of ideals L and M in R_Δ has a product LM which is again an ideal in R_Δ. We will see that this leads to a group structure on the set of strict equivalence classes of stable ideals, which, under the correspondence between ideals and forms, turns out to be the same as the group structure on the class group of forms studied in the previous chapter. If the procedure for defining the product of forms seemed perhaps a little complicated, the viewpoint of ideals provides an alternative that may seem more obvious and direct.

In order to form the product LM of two ideals L and M in R_Δ one's first guess might be to let LM consist of all products $\alpha\beta$ of elements α in L and β in M. This does not always work, however, as we will see in an example later in this section. The difficulty is that for two products $\alpha_1\beta_1$ and $\alpha_2\beta_2$ the sum $\alpha_1\beta_1 + \alpha_2\beta_2$ might not be equal to a product $\alpha\beta$ of an element of L with an element of M, as it would have to be if the set of all products $\alpha\beta$ was an ideal. This difficulty can be avoided by defining LM to be the set of all sums $\alpha_1\beta_1 + \cdots + \alpha_n\beta_n$ with each α_i in L and each β_i in M. With this definition LM is obviously closed under addition as well as subtraction. Also, multiplying such a sum $\sum_i \alpha_i\beta_i$ by an element γ in R_Δ gives an element of LM since $\gamma\sum_i \alpha_i\beta_i = \sum_i \gamma\alpha_)\beta_i$ and the latter sum is in LM since each product $\gamma\alpha_i$ is in L because L is an ideal. To finish the verification that LM is an ideal we need to check that it is a lattice since we defined ideals in R_Δ to be lattices that are taken to themselves by multiplication by arbitrary elements of R_Δ. To check that LM is a lattice we need to explain a few more things about lattices.

We defined a lattice in R_Δ to be a set $L(\alpha,\beta)$ of elements $x\alpha+y\beta$ as x and y range over all integers, where α and β are two elements of R_Δ that do not lie on the same line through the origin. More generally we could define $L(\alpha_1,\cdots,\alpha_n)$ to be the set of all linear combinations $x_1\alpha_1+\cdots+x_n\alpha_n$ with coefficients x_i in $\mathbb{Z}$, where not all the α_i's lie on the same line through the origin (so in particular at least two α_i's are nonzero). It is not immediately obvious that $L(\alpha_1,\cdots,\alpha_n)$ is a lattice, but this is true and can be proved by a generalization of the procedure that converts an arbitrary basis for a lattice into a reduced basis, as we will now describe.

There are three ways in which the set of generators α_i for $L(\alpha_1,\cdots,\alpha_n)$ can be modified without changing the set $L(\alpha_1,\cdots,\alpha_n)$:

(1) Replace one generator α_i with $\alpha_i+k\alpha_j$, adding an integer k times some other generator α_j to α_i.
(2) Replace some α_i by $-\alpha_i$.
(3) Interchange two generators α_i and α_j, or more generally permute the α_i's in any way.

After a modification of type (1) each integer linear combination of the new generators is also a linear combination of the old generators so the new $L(\alpha_1,\cdots,\alpha_n)$ is a subset of the old one, but the process can be reversed by another type (1) operation subtracting $k\alpha_j$ from the new α_i so the new $L(\alpha_1,\cdots,\alpha_n)$ also contains the old one hence must equal it. For the operations (2) and (3) this is also true, more obviously.

Lemma 8.24. *By applying a suitably chosen sequence of operations* (1)–(3) *to a set of generators α_i for $L(\alpha_1,\cdots,\alpha_n)$ it is always possible to produce a new set of generators $\beta_1,\cdots,\beta_n$ which are all zero except for β_1 and β_2. In particular $L(\alpha_1,\cdots,\alpha_n)$ is a lattice.*

Proof: Let us write R_Δ as $\mathbb{Z}[\tau]$ in the usual way. Each α_i can be written as $a_i+b_i\tau$ for integers a_i and b_i. We then form a $2\times n$ matrix $\begin{pmatrix} a_1 & \cdots & a_n \\ b_1 & \cdots & b_n \end{pmatrix}$ whose columns $\begin{pmatrix} a_i \\ b_i \end{pmatrix}$ correspond to the α_i's. The operations (1)–(3) correspond to adding an integer times one column to another column, changing the sign of a column, and permuting columns.

These three column operations can be used to simplify the matrix until only the first two columns are nonzero. To do this we first focus on the second row. This must have a nonzero entry since the α_i's are not all contained in the x-axis. The nonzero entries in the second row can be made all positive by changing the sign of some columns. Choose a column with smallest positive entry b_i. By subtracting suitable multiples of this column from the other columns with positive b_j's we can make all other b_j's either zero or positive integers less than b_i. This process can be repeated using columns with successively smaller second entries until only one nonzero b_i remains. Switching this column with the first column, we can then assume that $b_i = 0$ for all $i > 1$.

Now we do the same procedure for columns 2 through n using the entries a_i rather than b_i. Since these columns have $b_i = 0$, nothing changes in the second row. After this step is finished, only the first two columns will be nonzero. Note that neither of these columns can have both entries zero, otherwise $L(\alpha_1, \cdots, \alpha_n)$ would be entirely contained in a line through the origin. □

Let us restrict attention now to lattices that are ideals. One way to generate such a lattice is to start with elements $\alpha_1, \cdots, \alpha_n$ in R_Δ which we can assume are nonzero and then consider the set of all elements $\sum_i \gamma_i \alpha_i$ for arbitrary coefficients γ_i in R_Δ rather than just taking integer coefficients as we would be doing for the lattice $L(\alpha_1, \cdots, \alpha_n)$. The usual notation for this set of all sums $\sum_i \gamma_i \alpha_i$ is $(\alpha_1, \cdots, \alpha_n)$, generalizing the earlier notation (α) for a principal ideal. The ideal $(\alpha_1, \cdots, \alpha_n)$ is equal to the lattice $L(\alpha_1, \alpha_1\tau, \alpha_2, \alpha_2\tau, \cdots, \alpha_n, \alpha_n\tau)$ where $R_\Delta = \mathbb{Z}[\tau]$ since each coefficient γ_i in a sum $\sum_i \gamma_i \alpha_i$ can be written as $x_i + y_i\tau$ for integers x_i and y_i. To be sure that $(\alpha_1, \cdots, \alpha_n)$ really is a lattice, we should check that $\alpha_1, \alpha_1\tau, \cdots, \alpha_n, \alpha_n\tau$ do not all lie on the same line through the origin. But this is true already for α_1 and $\alpha_1\tau$ since (α_1) is an ideal as we saw in the previous section.

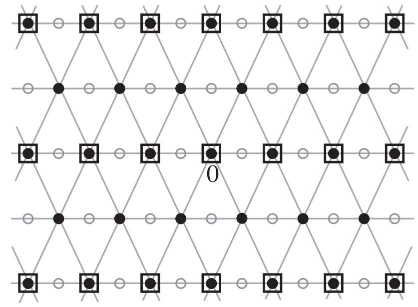
Observe that if a lattice $L(\alpha_1, \cdots, \alpha_n)$ is an ideal, then $L(\alpha_1, \cdots, \alpha_n)$ is equal to $(\alpha_1, \cdots, \alpha_n)$ since every product $\gamma\alpha_i$ with γ in R_Δ can be rewritten as an integer linear combination of $\alpha_1, \cdots, \alpha_n$ if $L(\alpha_1, \cdots, \alpha_n)$ is an ideal. A consequence of this, using Lemma 8.24, is that every ideal $(\alpha_1, \cdots, \alpha_n)$ with $n > 2$ can be rewritten as an ideal (β_1, β_2).

Now we return to products of ideals. For ideals $L = (\alpha_1, \alpha_2)$ and $M = (\beta_1, \beta_2)$ the product LM is the ideal $(\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \alpha_2\beta_2)$ since each of the four products $\alpha_i\beta_j$ is in LM and every element of LM is a sum of terms $\alpha\beta$ for $\alpha = \gamma_1\alpha_1 + \gamma_2\alpha_2$ and $\beta = \delta_1\beta_1 + \delta_2\beta_2$, so $\alpha\beta$ is a linear combination of the products $\alpha_i\beta_j$ with coefficients in R_Δ. Similarly, the product of ideals $(\alpha_1, \cdots, \alpha_n)$ and $(\beta_1, \cdots, \beta_k)$ is the ideal generated by all the products $\alpha_i\beta_j$.

As examples let us compute some products of ideals in $\mathbb{Z}[\sqrt{-5}]$, which is R_Δ for $\Delta = -20$. Consider first the ideal corresponding to the form $2x^2 + 2xy + 3y^2$, the lattice $L(2, 1 + \sqrt{-5})$. Since this is an ideal it is the same as the ideal $(2, 1 + \sqrt{-5})$. Denoting this ideal as P, let us compute its square $P^2 = PP$. We have:

$$P^2 = (2, 1 + \sqrt{-5})(2, 1 + \sqrt{-5}) = (4, 2 + 2\sqrt{-5}, 6)$$

In this ideal each generator is a multiple of 2 so we can pull out a factor of 2 to get $P^2 = 2(2, 1+\sqrt{-5}, 3)$. The ideal $(2, 1+\sqrt{-5}, 3)$ contains 3 and 2 so it contains their difference 1. Once an ideal contains 1 it must be the whole ring, so $(2, 1+\sqrt{-5}, 3) = (1) = \mathbb{Z}[\sqrt{-5}]$ hence $P^2 = 2(1) = (2)$. The figure at the right shows these ideals as lattices, with $(2, 1+\sqrt{-5})$ indicated by the heavy dots and its square (2) by the dots in squares. Notice that P^2 is a sublattice of P. In fact it is always true that a product LM of two ideals L and M is a sublattice of both L and M since each term of a typical element $\sum_i \alpha_i\beta_i$ of LM lies in both L and M by the defining property of ideals.



This example also illustrates the fact that a product LM of two ideals need not consist just of all products $\alpha\beta$ of an element of L with an element of M since the number 2 belongs to P^2 but if we had $2 = \alpha\beta$ with α and β in P then, computing norms, we would have $4 = N(\alpha)N(\beta)$. There are no elements of $\mathbb{Z}[\sqrt{-5}]$ of norm ± 2 since $N(x + y\sqrt{-5}) = x^2 + 5y^2 = \pm 2$ has no integer solutions. Thus either α or β would have norm ± 1 and hence be one of the two units ± 1 in $\mathbb{Z}[\sqrt{-5}]$. However, neither 1 nor -1 is in P, otherwise we would have $P = \mathbb{Z}[\sqrt{-5}]$.

Continuing with the ring $\mathbb{Z}[\sqrt{-5}]$, we consider next the ideal $Q = (3, 1+\sqrt{-5})$ corresponding to the form $3x^2 + 2xy + 2y^2$. For the product PQ we have:

$$PQ = (2, 1+\sqrt{-5})(3, 1+\sqrt{-5}) = (6, 2+2\sqrt{-5}, 3+3\sqrt{-5}, -4+2\sqrt{-5})$$

The last generator $-4+2\sqrt{-5}$ can be discarded since it is the second generator minus the first generator. The difference between the second and third generators is $1+\sqrt{-5}$ so this is in PQ, and these two generators are multiples of $1+\sqrt{-5}$ so we now have $PQ = (6, 1+\sqrt{-5})$. But 6 is in the ideal $(1+\sqrt{-5})$ since it is $1-\sqrt{-5}$ times $1+\sqrt{-5}$, the norm of $1+\sqrt{-5}$, so we have finally $PQ = (1+\sqrt{-5})$.

Next we calculate $Q\overline{Q}$ where the conjugate $\overline{L}$ of an ideal $L = (\alpha, \beta)$ is the ideal consisting of all the conjugates of elements of L, so $\overline{L} = (\overline{\alpha}, \overline{\beta})$. We have:

$$\begin{aligned} Q\overline{Q} = (3, 1+\sqrt{-5})(3, 1-\sqrt{-5}) &= (9, 3+3\sqrt{-5}, 3-3\sqrt{-5}, 6) \\ &= 3(3, 1+\sqrt{-5}, 1-\sqrt{-5}, 2) = (3) \end{aligned}$$

For the product $P\overline{P}$ there is no need to do a separate calculation since $P = \overline{P}$ as one can see in the previous figure, so $P\overline{P} = P^2 = (2)$.

Using these calculations we can see how the two different factorizations of (6) in $\mathbb{Z}[\sqrt{-5}]$ as $(2)(3)$ and as $(1+\sqrt{-5})(1-\sqrt{-5})$ arise:

$$(6) = (2)(3) = P\overline{P} \cdot Q\overline{Q} = P\overline{P}Q\overline{Q}$$

$$(6) = (1+\sqrt{-5})(1-\sqrt{-5}) = PQ \cdot \overline{PQ} = PQ\overline{P}\,\overline{Q}$$

For the last equality we are using the general identity $\overline{LM} = \overline{L}\,\overline{M}$ which follows easily from the definitions.

We defined the norm of an ideal L in R_Δ geometrically as the number of parallel translates of L, including L itself, that are needed to fill up all of R_Δ, and we found other ways to view these norms in terms of areas and determinants. For the ideals we will be most interested in, the stable ideals in Proposition 8.22, there is yet another interpretation of the norm $N(L)$ that is more like the definition of the norm of an element α as $N(\alpha) = \alpha\overline{\alpha}$.

Proposition 8.25. *If the ideal L in R_Δ is stable then $L\overline{L} = (N(L))$, the principal ideal generated by the norm $N(L)$.*

In the preceding example the calculations of $P\overline{P}$ and $Q\overline{Q}$ are consequences of this general result since the norm of an ideal $(a, 1+\sqrt{-5})$ is a.

Proof: By Proposition 8.16 the ideal L is equal to $nL(a, \frac{b+\sqrt{\Delta}}{2})$ for some integer $n \geq 1$ and some form $ax^2+bxy+cy^2$ of discriminant Δ with $a > 0$. It will suffice to prove the proposition in the case $n = 1$ since replacing an ideal L by nL does not affect the stabilizer and it multiplies $N(L)$ by n^2, so both sides of the equation $L\overline{L} = (N(L))$ are multiplied by n^2. Thus we may take $L = L(a, \frac{b+\sqrt{\Delta}}{2})$ for the rest of the proof. Since we assume L is stable, the form $ax^2 + bxy + cy^2$ is primitive by Proposition 8.22.

Let $\tau = \frac{b+\sqrt{\Delta}}{2}$ so τ is a root of the equation $x^2 - bx + ac = 0$ and $\tau\overline{\tau} = ac$. We have $L = (a, \tau)$ and $\overline{L} = (a, \overline{\tau})$. The product $L\overline{L}$ is then:

$$L\overline{L} = (a^2, a\tau, a\overline{\tau}, \tau\overline{\tau}) = (a^2, a\tau, a\overline{\tau}, ac) = a(a, \tau, \overline{\tau}, c)$$

The ideal $(a, \tau, \overline{\tau}, c)$ contains the ideal $(a, \tau+\overline{\tau}, c) = (a, b, c)$. The latter ideal is all of R_Δ since it contains all integral linear combinations $ma + nb + qc$ and there is one such combination that equals 1 since the greatest common divisor of a, b, and c is 1 because the form $ax^2 + bxy + cy^2$ is primitive. (We know from Chapter 2 that the greatest common divisor d of a and b can be written as $d = ma + nb$, and then the greatest common divisor of d and c, which is the greatest common divisor of a, b, and c, can be written as an integral linear combination of d and c and hence also of a, b, and c.)

Thus the ideal $(a, \tau, \overline{\tau}, c)$ contains R_Δ and so must equal it. Hence we have $L\overline{L} = aR_\Delta = (a)$ and this equals $(N(L))$ since $N(L) = a$ for $L = L(a, \frac{b+\sqrt{\Delta}}{2})$. ◻

Proposition 8.26. *An ideal L in R_Δ is stable if and only if there exists an ideal M in R_Δ such that LM is a principal ideal.*

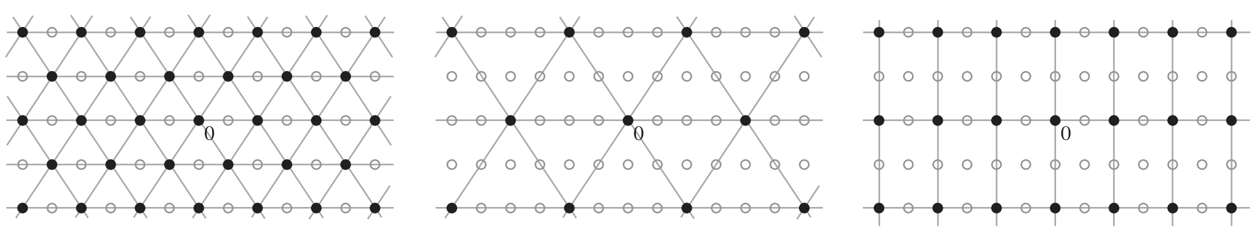
Proof: The forward implication follows from Proposition 8.25 by choosing $M = \overline{L}$. For the opposite implication, suppose that $LM = (\alpha)$, and let β be an element of $\mathbb{Q}(\sqrt{\Delta})$ such that βL is contained in L. Then $\beta(\alpha) = \beta LM$ is contained in $LM = (\alpha)$. In particular this says that $\beta\alpha$ is in (α) so $\beta\alpha = \gamma\alpha$ for some element γ of R_Δ. Since α is nonzero this implies $\beta = \gamma$ and so β is an element of R_Δ. This shows that the stabilizer of L is R_Δ, so L is stable. ◻

Proposition 8.27. *If L and M are stable ideals in R_Δ then $N(LM) = N(L)N(M)$.*

Proof: If L and M are stable then so is LM by Proposition 8.26 since the product of two principal ideals is principal. Since $\overline{LM} = \overline{L}\,\overline{M}$ we have $LM\overline{LM} = L\overline{L}M\overline{M}$ which means $(N(LM)) = (N(L))(N(M))$. We also have $(N(L))(N(M)) = (N(L)N(M))$ since for principal ideals we always have $(\alpha)(\beta) = (\alpha\beta)$. Thus $(N(LM)) = (N(L)N(M))$, and this implies $N(LM) = N(L)N(M)$ since if $(a) = (b)$ for positive integers a and b then $a = b$, as is evident from the lattices $(a) = L(a, a\tau)$ and $(b) = L(b, b\tau)$ for $R_\Delta = \mathbb{Z}[\tau]$. ☐

The formula $L\overline{L} = (N(L))$ and the multiplicative property $N(LM) = N(L)N(M)$ can fail to hold for ideals with stabilizer larger than R_Δ. A simple example is provided by taking L to be the ideal $(2, 1+\sqrt{-3})$ in $\mathbb{Z}[\sqrt{-3}]$ which we considered in the previous section, before Proposition 8.22, as an example of an ideal corresponding to the nonprimitive form $2x^2 + 2xy + 2y^2$ of discriminant -12. Here $L = \overline{L}$ and the ideal $L^2 = L\overline{L}$ is $(2, 1+\sqrt{-3})(2, 1-\sqrt{-3}) = (4, 2+2\sqrt{-3}, 2-2\sqrt{-3}, 4)$. Of these four generators we can obviously drop the repeated 4, and we can also omit the third generator which is expressible as the first generator minus the second. We are left with the ideal $(4, 2+2\sqrt{-3}) = 2(2, 1+\sqrt{-3})$. Thus we have $L^2 = L\overline{L} = 2L$.

$L = L(2, 1+\sqrt{-3})$ $\qquad$ $L^2 = 2L = L(4, 2+2\sqrt{-3})$ $\qquad$ $(2) = L(2, 2\sqrt{-3})$



From the figure we see that $N(L) = 2$ and hence $N(2L) = 2^2N(L) = 8$ so $N(L^2) \neq N(L)^2 = 4$. This shows that $N(LM)$ need not equal $N(L)N(M)$ in general. Also we see from the figure that $L\overline{L} \neq (N(L))$ since $L\overline{L} = L^2 = 2L \neq (2) = (N(L))$. In fact $L\overline{L}$ is not even a principal ideal since $2L$ is a lattice of equilateral triangles while principal ideals have the same shape as the rectangular lattice $\mathbb{Z}[\sqrt{-3}]$.

Now at last we come to the construction of the ideal class group, which we will denote $ICG(\Delta)$ until we show that it coincides with the class group $CG(\Delta)$ defined in Chapter 7 in terms of forms. Let $[L]$ denote the strict equivalence class of a stable ideal L in R_Δ and let $ICG(\Delta)$ be the set of such classes $[L]$. The multiplication operation in $ICG(\Delta)$ is defined by taking products of ideals, so we set $[L][M] = [LM]$, recalling the fact that the product of two stable ideals is stable by Proposition 8.26. To check that this product in $ICG(\Delta)$ is well defined we need to see that choosing different ideals L' and M' in the classes $[L]$ and $[M]$ does not affect $[LM]$. This is true because $[L] = [L']$ means $\alpha L = \alpha' L'$ for some α and α', and $[M] = [M']$ means $\beta M = \beta' M'$ for some β and β', hence $\alpha\beta LM = \alpha'\beta' L'M'$, so $[LM] = [L'M']$.

Here we are dealing with strict equivalence classes of ideals so we are assuming all of $\alpha, \beta, \alpha', \beta'$ have positive norms, hence so do $\alpha\beta$ and $\alpha'\beta'$. (As always this condition is automatic when Δ is negative.)

Proposition 8.28. *$ICG(\Delta)$ is a commutative group with respect to the multiplication $[L][M] = [LM]$.*

Proof: The commutativity property $[L][M] = [M][L]$ is easy since this amounts to saying $[LM] = [ML]$, which holds since multiplication of ideals is commutative, $LM = ML$, because multiplication in R_Δ is commutative.

To have a group there are three things to check. First, the multiplication should be associative, so $([L][M])[N] = [L]([M][N])$. By the definition of the product in $ICG(\Delta)$ this is equivalent to saying $[LM][N] = [L][MN]$ which in turn means $[(LM)N] = [L(MN)]$, so it suffices to check that multiplication of ideals is associative, $(LM)N = L(MN)$. The claim is that each of these two products consists of all the finite sums $\sum_i \alpha_i\beta_i\gamma_i$ with α_i, β_i, and γ_i elements of L, M, and N respectively. Every such sum is in both $(LM)N$ and $L(MN)$ since each term $\alpha_i\beta_i\gamma_i$ is in both of the ideals $(LM)N$ and $L(MN)$. Conversely, each element of $(LM)N$ is a sum of terms $(\sum_j \alpha_j\beta_j)\gamma$ so it can be written as a sum $\sum_i \alpha_i\beta_i\gamma_i$, and similarly each element of $L(MN)$ can be written as a sum $\sum_i \alpha_i\beta_i\gamma_i$. Thus we have $(LM)N = L(MN)$.

Next, a group must have an identity element, and the class $[(1)]$ of the ideal $(1) = R_\Delta$ obviously serves this purpose since $(1)L = L$ for all ideals L, hence $[(1)][L] = [L]$. There is no need to check that $[L][(1)] = [L]$ as one would have to do for a noncommutative group since we have already observed that multiplication in $ICG(\Delta)$ is commutative.

The last thing to check is that each element of $ICG(\Delta)$ has a multiplicative inverse, and this is where we use the condition that we are considering only stable ideals in the definition of $ICG(\Delta)$. As we showed in Proposition 8.25, each stable ideal L satisfies $L\overline{L} = (n)$ where the integer n is the norm of L. Then we have $[L][\overline{L}] = [(n)] = [(1)]$ where this last equality holds since the ideals (n) and (1) are strictly equivalent, the norm of n being n^2, a positive integer. Thus the multiplicative inverse of $[L]$ is $[\overline{L}]$. Again commutativity of the multiplication means that we do not have to check that $[\overline{L}]$ is an inverse for $[L]$ for multiplication both on the left and on the right. □

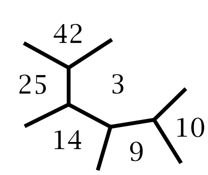
There is a variant of the ideal class group in which the relation of strict equivalence of ideals is modified by deleting the word "strict", so an ideal L is considered equivalent to αL for all nonzero elements α of R_Δ without the condition that $N(\alpha) > 0$. The preceding proof that $ICG(\Delta)$ is a group applies equally well in this setting by just omitting any mention of norms being positive. Sometimes the resulting group is called the class group while $ICG(\Delta)$ is called the strict class group or narrow class group. However, for studying quadratic forms the more appropriate notion is strict equivalence, which is why we are using this for the class group $ICG(\Delta)$.

Next we check that the one-to-one correspondence $\Phi\colon CG(\Delta) \to ICG(\Delta)$ induced by sending a form $Q = ax^2 + bxy + cy^2$ with $a > 0$ to the ideal $L_Q = (a, \frac{b+\sqrt{\Delta}}{2})$ respects the group structures defined on $CG(\Delta)$ and $ICG(\Delta)$. Given two classes $[Q_1]$ and $[Q_2]$ in $CG(\Delta)$, we can realize them by concordant forms $[a_1, b, a_2c]$ and $[a_2, b, a_1c]$ with a_1 and a_2 coprime and positive. The product $[Q_1][Q_2]$ in $CG(\Delta)$ is then the class of $[a_1a_2, b, c]$. The ideals corresponding to these three forms are $L_1 = (a_1, \frac{b+\sqrt{\Delta}}{2})$, $L_2 = (a_2, \frac{b+\sqrt{\Delta}}{2})$, and $L_3 = (a_1a_2, \frac{b+\sqrt{\Delta}}{2})$. To show that multiplication in $CG(\Delta)$ corresponds under Φ to multiplication in $ICG(\Delta)$ it will suffice to show that $L_1L_2 = L_3$. The product L_1L_2 is the ideal $(a_1a_2, a_1\frac{b+\sqrt{\Delta}}{2}, a_2\frac{b+\sqrt{\Delta}}{2}, \frac{b+\sqrt{\Delta}}{2}\cdot\frac{b+\sqrt{\Delta}}{2})$. This is certainly contained in L_3 since the first generator a_1a_2 is in L_3 and the other three generators are multiples of $\frac{b+\sqrt{\Delta}}{2}$ by elements of R_Δ hence are in L_3. On the other hand L_3 is contained in L_1L_2 since a_1a_2 is in L_1L_2 and so is $\frac{b+\sqrt{\Delta}}{2}$ which can be written as a linear combination $ma_1\frac{b+\sqrt{\Delta}}{2} + na_2\frac{b+\sqrt{\Delta}}{2}$ for some integers m and n, using the fact that a_1 and a_2 are coprime so we have $ma_1 + na_2 = 1$ for some integers m and n.

The identity element of $CG(\Delta)$ is the class of the principal form $[1, b, c]$ and this is sent by Φ to the class of the ideal $(1, \frac{b+\sqrt{\Delta}}{2}) = (1)$ which is the identity element of $ICG(\Delta)$. The inverse of an element of $CG(\Delta)$ determined by a form $[a, b, c]$ is the class of the mirror image form $[a, -b, c]$, so under Φ these forms correspond to the ideals $(a, \frac{b+\sqrt{\Delta}}{2})$ and $(a, \frac{-b+\sqrt{\Delta}}{2})$. The latter ideal is the same as $(a, \frac{b-\sqrt{\Delta}}{2})$ which is the conjugate of $(a, \frac{b+\sqrt{\Delta}}{2})$ so it gives the inverse of $(a, \frac{b+\sqrt{\Delta}}{2})$ in $ICG(\Delta)$.

Thus the group structures on $CG(\Delta)$ and $ICG(\Delta)$ are really the same, and we can use the notation $CG(\Delta)$ for both without any conflict.

To illustrate this let us consider $CG(\Delta)$ for $\Delta = -104$, so $R_\Delta = \mathbb{Z}[\sqrt{-26}]$. We looked at this example in Section 7.2 and found that $CG(\Delta)$ is a cyclic group of order 6 generated by the form $Q_4 = [5, 4, 6]$. From the topographs we could see that Q_4^2 was either $Q_3 = [3, 2, 9]$ or $Q_3^{-1} = [3, -2, 9]$, but to determine which, we had to find a pair of concordant forms equivalent to Q_4 and multiply them together. Now we can use ideals to do the same calculation. The ideal corresponding to $Q_4 = [5, 4, 6]$ is $(5, 2 + \sqrt{-26})$ so for Q_4^2 the ideal is $(5, 2 + \sqrt{-26})(5, 2 + \sqrt{-26})$ which equals $(25, 10 + 5\sqrt{-26}, -22 + 4\sqrt{-26})$. The next step is to find a reduced basis for this ideal. As a lattice this ideal is generated by these three elements and their products with $\sqrt{-26}$. Thus we have the matrix $\begin{pmatrix} 25 & 0 & 10 & -130 & -22 & -104 \\ 0 & 25 & 5 & 10 & 4 & -22 \end{pmatrix}$ which reduces to $\begin{pmatrix} 25 & 7 \\ 0 & 1 \end{pmatrix}$ so the ideal is $(25, 7 + \sqrt{-26})$. The corresponding form is $[25, 14, c]$ and we can determine c from the discriminant equation $b^2 - 4ac = -104$ which gives $c = 3$. The form is thus $[25, 14, 3]$. A small portion of the topograph of this form is shown at the right. There is a source vertex surrounded by the three values $3, 9, 10$ in counterclockwise order. The form $[3, -2, 9]$ has exactly this same configuration at its source vertex, so we conclude that $Q_4^2 = Q_3^{-1}$, the same answer we got in Section 7.2.



Exercises

1. Corresponding to a lattice $L(\alpha_1,\cdots,\alpha_n)$ in $\mathbb{Z}[\tau]$ there is a matrix $\left(\begin{smallmatrix} a_1 & \cdots & a_n \\ b_1 & \cdots & b_n \end{smallmatrix}\right)$ with $\alpha_i = a_i + b_i\tau$ as in Lemma 8.24. Show that the three operations of adding a multiple of one column to another, changing the sign of a column, and permuting columns do not change the greatest common divisor of the numbers in each row of the matrix. Deduce from this that if $a, b + c\tau$ is the reduced basis for the lattice then c is the greatest common divisor of the entries in the second row of the matrix.

2. In $\mathbb{Z}[\sqrt{-6}]$ compute the powers of the ideal $(2,\sqrt{-6})$ and determine which powers are principal ideals.

3. In $\mathbb{Z}[\sqrt{-14}]$ do the following:
(a) Compute the square of the ideal $(2,\sqrt{-14})$.
(b) For the ideal $L = (3, 1+\sqrt{-14})$ find a reduced basis for L^2 and use this to draw a picture of the lattice L^2.
(c) Find nonzero elements α and β in $\mathbb{Z}[\sqrt{-14}]$ such that $\alpha L^2 = \beta(2,\sqrt{-14})$.

4. In $\mathbb{Z}[\sqrt{-5}]$ compute Q^2 for $Q = (3, 1+\sqrt{-5})$ as a principal ideal (α) after first determining what $N(\alpha)$ must be.

5. Use the formula $N(LM) = N(L)N(M)$ with $L = (\alpha)$ and $M = (\overline{\alpha})$ to give another proof that $N\big((\alpha)\big) = |N(\alpha)|$.

8.5 Unique Factorization of Ideals

In this section we will be restricting our attention exclusively to discriminants Δ that are fundamental discriminants, so all forms will be primitive and hence all ideals in R_Δ will be stable. This means that we will be able to make free use of the formulas $N(LM) = N(L)N(M)$ and $L\overline{L} = (N(L))$.

Our main goal in this section is to show that all ideals in R_Δ, with the trivial exception of R_Δ itself, have unique factorizations as products of prime ideals, where an ideal P different from R_Δ is called a ***prime ideal*** if whenever it is expressed as a product LM of two ideals in R_Δ, either L or M must equal R_Δ, so the factorization becomes the trivial factorization $P = R_\Delta P$ that every ideal has. Note that R_Δ, considered as an ideal in R_Δ, satisfies this condition but we do not allow R_Δ as a prime ideal, just as the number 1 is not considered a prime number.

For an element α of R_Δ we know that α is prime if its norm $N(\alpha)$ is prime in $\mathbb{Z}$, either positive or negative. The analogue for ideals also holds:

Proposition 8.29. *If the norm $N(P)$ of an ideal P is prime then P is a prime ideal.*

Proof: Suppose $P = LM$. Then $N(P) = N(L)N(M)$. If $N(P)$ is prime then since $N(L)$ and $N(M)$ are positive integers, one of them must be 1. The only ideal of norm 1 is R_Δ so this means L or M must be R_Δ. Thus P is a prime ideal. □

Proposition 8.30. *For each prime p the principal ideal (p) in R_Δ is either a prime ideal or it factors as $(p) = P\overline{P}$ for prime ideals P and $\overline{P}$ of norm p.*

As we will see later in Corollary 8.34, all prime ideals in R_Δ are accounted for by this proposition, so every prime ideal is either a principal ideal (p) with p prime or a factor P or $\overline{P}$ when $(p) = P\overline{P}$.

Proof: If (p) is not a prime ideal then it factors as $(p) = PQ$ for ideals P and Q not equal to R_Δ. Since the norm of (p) is p^2 we must have $N(P) = p$ and $N(Q) = p$. From the general formula $L\overline{L} = \big(N(L)\big)$ we have $P\overline{P} = \big(N(P)\big) = (p)$. Since $N(P) = p$ we must also have $N(\overline{P}) = p$ so P and $\overline{P}$ are both prime ideals. (From the unique prime factorization property of ideals it will follow that $Q = \overline{P}$, but we do not need to know this here.) □

In the case that $(p) = P\overline{P}$ the prime p is said to ***split*** in R_Δ. The primes that split in R_Δ are the primes that are norms of ideals in R_Δ, and as we saw in Section 8.3 these are exactly the primes that are represented by forms of discriminant Δ. For a split prime p we saw in Proposition 8.18 how to find an ideal P of norm p so this now tells us how to factor (p) as $P\overline{P}$.

A further distinction for split primes is whether the two factors of $(p) = P\overline{P}$ are equal or not. If $P = \overline{P}$ then p is said to be ***ramified*** in R_Δ. According to part (c) of Proposition 8.18 the ramified primes are exactly the primes that divide Δ.

Now we turn to proving the unique factorization property for ideals in R_Δ. It will be helpful to have a criterion for when one ideal L in R_Δ divides another ideal M, meaning that $M = LK$ for some ideal K. For individual elements of R_Δ it is easy to tell when one element divides another since α divides β exactly when the quotient β/α lies in R_Δ. For ideals, however, the criterion is rather different:

Proposition 8.31. *An ideal L in R_Δ divides an ideal M if and only if L contains M.*

One can remember this as "to divide is to contain". At first glance the proposition may seem a little puzzling since for ordinary numbers the divisors of a number n, apart from n itself, are smaller than n while for ideals the divisors are larger, where "larger" for sets means that one set contains the other. The puzzle can be resolved by interpreting "m divides n" as "the multiples of m contain the multiples of n".

The proposition gives some insight into the choice of the ideals P and Q in the example preceding Proposition 8.25 where we factored the ideal (6) in $\mathbb{Z}[\sqrt{-5}]$ as $(2)(3) = P\overline{P}\cdot Q\overline{Q}$ and as $(1+\sqrt{-5})(1-\sqrt{-5}) = PQ\cdot\overline{P}\overline{Q}$. Since we want $P\overline{P} = (2)$ and $PQ = (1+\sqrt{-5})$, this means that P should divide both (2) and $(1+\sqrt{-5})$. By the above Proposition 8.31 this is the same as saying that P should contain both (2) and

$(1+\sqrt{-5})$. An obvious ideal with this property is the ideal $(2, 1+\sqrt{-5})$. Similarly one would be led to try $Q = (3, 1+\sqrt{-5})$. Then one could check that these choices for P and Q actually work.

Before proving the proposition let us derive a fact which will be used in the proof, a cancellation property of multiplication of ideals: If $LM_1 = LM_2$ then $M_1 = M_2$. To see this, first multiply the equation $LM_1 = LM_2$ by $\overline{L}$ to get $\overline{L}LM_1 = \overline{L}LM_2$. Since $\overline{L}L = (n)$ for $n = N(L)$, a positive integer, we then have $(n)M_1 = (n)M_2$, which is equivalent to saying $nM_1 = nM_2$. Thus the rescalings nM_1 and nM_2 of M_1 and M_2 are equal, so after rescaling again by the factor $1/n$ we get $M_1 = M_2$.

Now let us prove the proposition.

Proof: Suppose first that L divides M, so $M = LK$ for some ideal K. A typical element of LK is a sum $\sum_i \alpha_i\beta_i$ with $\alpha_i \in L$ and $\beta_i \in K$ for all i. Since L is an ideal, each term $\alpha_i\beta_i$ is then in L and hence so is their sum. This shows that L contains $LK = M$.

For the converse, suppose L contains M. Then $L\overline{L}$ contains $M\overline{L}$. Since $L\overline{L} = (n)$ for $n = N(L)$, this says that (n) contains $M\overline{L}$, so every element of $M\overline{L}$ is a multiple of n by some element of R_Δ. This means that if we write $M\overline{L} = (\alpha, \beta)$ then we can define an ideal K by letting $K = (\alpha/n, \beta/n)$.

Now we have $(n)K = (n)(\alpha/n, \beta/n) = (\alpha, \beta) = M\overline{L}$. Multiplying by L we then have $(n)KL = M\overline{L}L = M(n)$. Canceling the factor (n) gives the equation $KL = M$, which says that L divides M, finishing the proof of the converse. $\square$

When we proved unique prime factorization for $\mathbb{Z}$ and those rings R_Δ that have a Euclidean algorithm, a key step was showing that if a prime p divides a product ab then p must divide either a or b. Now we prove the corresponding fact for ideals:

Lemma 8.32. *If a prime ideal P divides a product LM of two ideals, then P must divide either L or M.*

Proof: We will prove the equivalent statement that if P divides LM but not L, then P divides M. Consider the set $P+L$ of all sums $\alpha+\beta$ of elements $\alpha \in P$ and $\beta \in L$. This set $P+L$ is an ideal since if $P = (\alpha_1, \alpha_2)$ and $L = (\beta_1, \beta_2)$ then $P+L = (\alpha_1, \alpha_2, \beta_1, \beta_2)$. The ideal $P+L$ is strictly larger than P since the assumption that P does not divide L means that P does not contain L, so any element of L not in P is in $P+L$ but not P. Thus $P+L$ contains P, hence divides P, but is not equal to P. Since P is prime we must then have $P+L = R_\Delta$.

In particular $P+L$ contains 1 so we can write $1 = \alpha + \beta$ for some $\alpha \in P$ and $\beta \in L$. For an arbitrary element $\gamma \in M$ we then have $\gamma = \alpha\gamma + \beta\gamma$. The term $\alpha\gamma$ is in P since α is in P and P is an ideal. The term $\beta\gamma$ is in LM since β is in L and γ is in M. We assume P divides LM so P contains LM and it follows that $\beta\gamma$ is in P. Thus both terms on the right side of the equation $\gamma = \alpha\gamma + \beta\gamma$ are in P so γ is in P. Since γ was an arbitrary element of M this shows that M is contained in P, or in other words P divides M, which is what we wanted to prove. $\square$

Now we can prove our main result:

Theorem 8.33. *Every ideal in R_Δ other than R_Δ itself is a product of prime ideals, and this factorization is unique up to the order of the factors.*

Proof: We first show the existence of a prime factorization for each ideal $L \neq R_\Delta$. If L is prime itself there is nothing to prove, so suppose L is not prime, hence there is a factorization $L = KM$ with neither factor equal to R_Δ. Taking norms, we have $N(L) = N(K)N(M)$. Both $N(K)$ and $N(M)$ are greater than 1 since R_Δ is the only ideal of norm 1. Hence $N(K) < N(L)$ and $N(M) < N(L)$. By induction on the norm, both K and M have prime factorizations, hence so does $L = KM$. We can start the induction with the case $N(L) = 2$, a prime, hence L is prime. (The case $N(L) = 1$ does not arise since $L \neq R_\Delta$.)

For the uniqueness, suppose an ideal L has prime factorizations $P_1 \cdots P_k$ and $Q_1 \cdots Q_l$. We can assume $k \leq l$ by a notational change if necessary. The prime ideal P_1 divides the product $Q_1(Q_2 \cdots Q_l)$ so by the preceding lemma it must divide either Q_1 or $Q_2 \cdots Q_l$. In the latter case the same reasoning shows it must divide either Q_2 or $Q_3 \cdots Q_l$. Repeating this argument enough times, we eventually deduce that P_1 must divide some Q_i, and after permuting the factors of $Q_1 \cdots Q_l$ we can assume that P_1 divides Q_1. When one prime ideal divides another prime ideal they must be equal. For if P divides Q then $Q = PM$ for some M, but if Q is prime then either $P = R_\Delta$, which is impossible if P is prime, or $M = R_\Delta$, hence $P = Q$.

Once we have $P_1 = Q_1$ we can cancel this common factor of $P_1 \cdots P_k$ and $Q_1 \cdots Q_l$ to get $P_2 \cdots P_k = Q_2 \cdots Q_l$. Repeating this process often enough, we eventually get, after suitably permuting the Q_i's, that $P_1 = Q_1$, $P_2 = Q_2$, $\cdots$, $P_{k-1} = Q_{k-1}$, and $P_k = Q_k \cdots Q_l$. Since P_k is prime, as are the Q_i's, the equation $P_k = Q_k \cdots Q_l$ can have only one term on the right side, so $k = l$ and $P_k = Q_k$. This finishes the proof of the uniqueness of prime factorizations of ideals. □

From unique factorization we can deduce that there are no other prime ideals beyond those we saw in Proposition 8.30.

Corollary 8.34. *All prime ideals P in R_Δ are factors of ideals (p) for primes p, with either $(p) = P$ or $(p) = P\overline{P}$.*

Proof: Let P be a prime ideal in R_Δ. We have $P\overline{P} = (N(P))$. Writing $N(P)$ as a product $p_1 \cdots p_k$ of primes p_i, we then have $P\overline{P} = (p_1) \cdots (p_k)$. Thus P divides $(p_1) \cdots (p_k)$ so since P is prime it must divide one of the factors. This means there is a prime p such that P divides (p). Proposition 8.30 then finishes the proof. □

Let us consider how one can find the prime factorization of a given ideal. The procedure will be analogous to how we factored Gaussian integers in Section 8.1. We begin with an example in the case $\Delta = -24$ with $R_\Delta = \mathbb{Z}[\sqrt{-6}]$. We looked at this case in Section 8.3 when we considered how to find ideals of a given norm. For the norm

35 we found the two ideals $(35, 8+\sqrt{-6})$ and $(35, 13+\sqrt{-6})$ and their conjugates. The prime factors of these ideals will have norms dividing 35, so either 5 or 7, with one factor of norm 5 and one of norm 7. We found the ideals of norms 5 and 7, which were $(5, 2\pm\sqrt{-6})$ and $(7, 1\pm\sqrt{-6})$, and we need to see now which of these ideals divide $(35, 8+\sqrt{-6})$ and which divide $(35, 13+\sqrt{-6})$, or in other words, which of these ideals contain $(35, 8+\sqrt{-6})$ and which contain $(35, 13+\sqrt{-6})$. This will be easy using the following general fact:

Lemma 8.35. *A lattice $L(a, b+c\tau)$ in $\mathbb{Z}[\tau]$ contains another lattice $L(a', b'+c'\tau)$ if and only if a divides a', c divides c', and $b' \equiv b\,{}^{c'}\!/c \bmod a$.*

Proof: For $L(a, b+c\tau)$ to contain $L(a', b'+c'\tau)$ amounts to asking when a' and $b'+c'\tau$ are in $L(a, b+c\tau)$. For a', the only integers in $L(a, b+c\tau)$ are the multiples of a, so the condition on a' is that it must be a multiple of a. For $b'+c'\tau$ to be in $L(a, b+c\tau)$ means that the equation $b'+c'\tau = ax+(b+c\tau)y$ must have an integer solution. Equating the coefficients of τ gives $c' = cy$ which just says that c' is a multiple of c, with $y = {}^{c'}\!/c$. Then the equation becomes $b' = ax + b\,{}^{c'}\!/c$ which is equivalent to the congruence $b' \equiv b\,{}^{c'}\!/c \bmod a$. □

Applying this lemma to determine which of $(5, 2\pm\sqrt{-6})$ contains $(35, 8+\sqrt{-6})$ we see that the two divisibility conditions are satisfied and the congruence condition is $8 \equiv \pm 2 \bmod 5$ where the sign is the same as in $(5, 2\pm\sqrt{-6})$. The minus sign gives a valid congruence so it is $(5, 2-\sqrt{-6})$ that divides $(35, 8+\sqrt{-6})$. For $(7, 1\pm\sqrt{-6})$ to divide $(35, 8+\sqrt{-6})$ the divisibility conditions are again satisfied and the congruence condition is now $8 \equiv \pm 1 \bmod 7$ so this time the plus sign is correct so $(7, 1+\sqrt{-6})$ divides $(35, 8+\sqrt{-6})$. Thus we obtain the prime factorization of $(35, 8+\sqrt{-6})$ as $(5, 2-\sqrt{-6})(7, 1+\sqrt{-6})$. In similar fashion one finds that $(35, 13+\sqrt{-6})$ factors as $(5, 2-\sqrt{-6})(7, 1-\sqrt{-6})$. Taking the conjugates of these two factorizations gives the factorizations of the other two ideals of norm 35.

The general procedure for finding the prime factorization of an ideal L in R_Δ can be described as follows. As an easy first step one finds the largest positive integer n dividing each generator for L, assuming L is given in terms of generators. This gives a factorization $L = nL' = (n)L'$ with L' a primitive ideal. Factoring (n) into prime ideals is done by first factoring n as a product of primes p_i and then factoring the corresponding principal ideals (p_i) as in Proposition 8.30. This reduces the problem to the case that L is a primitive ideal. To do this case one computes $N(L)$, say by finding a reduced basis for L, then one factors $N(L)$ as $N(L) = p_1^{r_1}\cdots p_k^{r_k}$ for distinct primes p_i. These must be split primes, otherwise L would not be primitive. After factoring each principal ideal (p_i) as $P_i\overline{P}_i$, one can then determine which of P_i or $\overline{P}_i$ divides L by applying the preceding Lemma 8.35. Only one of P_i and $\overline{P}_i$ can divide L since L is primitive, so the prime factorization of L is then obtained from the product $p_1^{r_1}\cdots p_k^{r_k}$ by replacing each p_i by the ideal P_i or $\overline{P}_i$ that divides L.

Unique prime factorization for ideals can be used to determine the number of times each number n appears in a given topograph. Let us illustrate this by returning to the case of discriminant -24 where there are the two forms x^2+6y^2 and $2x^2+3y^2$. As we saw in Section 8.3, the number of appearances of n for both forms together is the same as the number of primitive ideals of norm n. The norms of primitive ideals are the numbers $n = 2^a 3^b p_1^{r_1} \cdots p_k^{r_k}$ with $a \le 1$, $b \le 1$, and the p_i's distinct unramified split primes. The primitive ideals of norm n are then obtained by replacing the factors 2 and 3 in $2^a 3^b p_1^{r_1} \cdots p_k^{r_k}$ by the ideals $(2, \sqrt{-6})$ and $(3, \sqrt{-6})$ and each $p_i^{r_i}$ by either $P_i^{r_i}$ or $\overline{P}_i^{r_i}$ where $(p_i) = P_i \overline{P}_i$. Thus there are exactly 2^k primitive ideals of norm n, so this is the number of times that n appears in at least one of the two topographs. We know from Chapter 6 that no number is represented by both forms, and the form representing n is $x^2 + 6y^2$ or $2x^2 + 3y^2$ according to whether the character values $\chi_3(n)$ and $\chi_8(n)$ are both $+1$ or both -1.

In some cases the unique factorization property for ideals implies unique factorization for elements of R_Δ. The relation between the two situations is obtained by associating to each nonzero element α in R_Δ the principal ideal (α). Multiplication of elements corresponds to multiplication of ideals since $(\alpha\beta) = (\alpha)(\beta)$. A key observation is that $(\alpha) = (\beta)$ if and only if α and β differ only by multiplication by a unit. For if $\beta = \varepsilon\alpha$ for some unit ε then (ε) contains $\varepsilon\varepsilon^{-1} = 1$ so $(\varepsilon) = R_\Delta$ hence $(\beta) = (\varepsilon\alpha) = (\varepsilon)(\alpha) = (\alpha)$. Conversely, if $(\alpha) = (\beta)$ then β is in (α) so $\beta = \varepsilon\alpha$ for some $\varepsilon \in R_\Delta$, and similarly $\alpha = \eta\beta$ for some $\eta \in R_\Delta$. Thus $\alpha = \eta\beta = \eta\varepsilon\alpha$ hence $\eta\varepsilon = 1$ so ε and η are units, showing that α and β differ just by a unit.

Proposition 8.36. *If all ideals in R_Δ are principal ideals then all elements of R_Δ other than units and 0 have unique factorizations as products of prime elements, where the uniqueness is up to order and multiplication by units.*

Proof: This follows immediately from Theorem 8.33 since principal ideals in R_Δ correspond exactly to nonzero elements of R_Δ up to multiplication by units. □

Proposition 8.37. *When $\Delta < 0$ all ideals are principal if and only if all forms are equivalent to the principal form. When $\Delta > 0$ all ideals are principal if and only if all forms are equivalent to either the principal form or its negative.*

Proof: All principal ideals in R_Δ are equivalent since they are equivalent to R_Δ itself. In fact the principal ideals form a complete equivalence class of ideals since any ideal that is equivalent to a principal ideal is also a principal ideal by the following argument. Suppose an ideal L is equivalent to a principal ideal (α), so $\beta L = \gamma(\alpha)$ for nonzero elements β and γ of R_Δ. Then $\gamma\alpha$ is in βL, which means $\gamma\alpha = \beta\delta$ for some δ in L, and hence we have $\beta L = \gamma(\alpha) = (\gamma\alpha) = (\beta\delta) = \beta(\delta)$. Thus $\beta L = \beta(\delta)$, so after multiplying both sides of this equation by β^{-1} in $\mathbb{Q}(\sqrt{\Delta})$ we have $L = (\delta)$, a principal ideal.

To prove the proposition we will use the one-to-one correspondence between proper equivalence classes of forms and strict equivalence classes of ideals. The principal form has mirror symmetry so forms equivalent to this form are properly equivalent to it, and the same holds for the negative of the principal form, which only enters the picture when $\Delta > 0$.

We distinguish three cases:

Case 1: $\Delta < 0$. Here equivalence of ideals is the same as strict equivalence. The principal form has leading coefficient 1 so it corresponds to the principal ideal R_Δ. Thus all forms are equivalent to the principal form exactly when all ideals are equivalent to R_Δ, or in other words, all ideals are principal.

Case 2: $\Delta > 0$ and the principal form is equivalent to its negative. The principal form then represents -1 so equivalence of ideals is again the same as strict equivalence. Thus there is a single equivalence class of forms exactly when there is a single equivalence class of ideals, the principal ideals.

Case 3: $\Delta > 0$ and the principal form is not equivalent to its negative. These forms then give two different equivalence classes of forms, and we will show that they correspond to two different strict equivalence classes of principal ideals (α), those with $N(\alpha) > 0$ and those with $N(\alpha) < 0$.

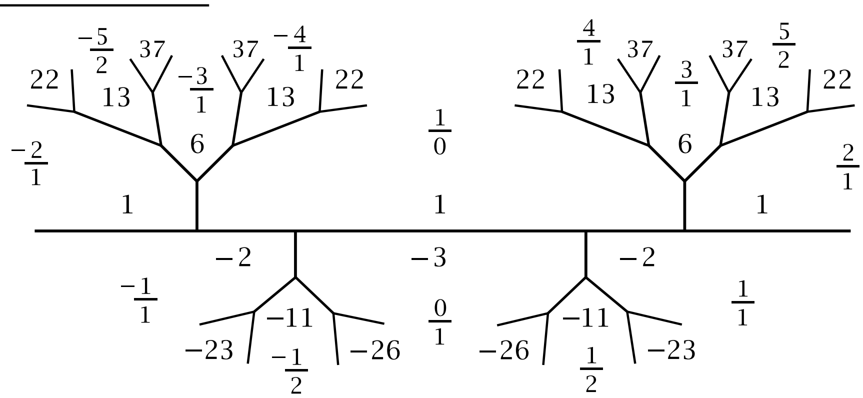
Any two ideals (α) and (β) with $N(\alpha) > 0$ and $N(\beta) > 0$ are strictly equivalent since they are both strictly equivalent to (1). Likewise (α) and (β) are strictly equivalent if $N(\alpha) < 0$ and $N(\beta) < 0$ since if γ is any element with $N(\gamma) < 0$, for example α or β, then (α) and (β) are both strictly equivalent to $(\alpha\beta\gamma)$ since $N(\beta\gamma) > 0$ and $N(\alpha\gamma) > 0$. Now suppose (α) and (β) are strictly equivalent with $N(\alpha)$ and $N(\beta)$ having opposite sign. Then $(\gamma\alpha) = (\delta\beta)$ for some γ and δ of positive norm. This means we have elements $\alpha' = \gamma\alpha$ and $\beta' = \delta\beta$ with $(\alpha') = (\beta')$ and such that the norms of α' and β' have opposite sign. Since $(\alpha') = (\beta')$ we have $\beta' = \varepsilon\alpha'$ for some unit ε. Since $N(\alpha')$ and $N(\beta')$ have opposite sign we must have $N(\varepsilon) < 0$. This means that the principal form represents -1 so its topograph has a skew symmetry, making it equivalent to its negative, contrary to hypothesis. Thus we have shown the the equivalence class of principal ideals (α) splits into two strict equivalence classes according to the sign of $N(\alpha)$.

Now we show that the negative of the principal form corresponds to a principal ideal (α) with $N(\alpha) < 0$. The principal form is $x^2 - dy^2$ if $\Delta = 4d$ and $x^2 + xy - dy^2$ if $\Delta = 4d + 1$. The negative of the principal form has leading coefficient -1 so to find the corresponding ideal as in Theorem 8.21 we first have to choose a properly equivalent form with positive leading coefficient. For this we can choose $dx^2 - y^2$ or $dx^2 + xy - y^2$, obtained from the negative of the principal form by replacing x, y by $-y, x$, rotating the topograph by 180 degrees. For $dx^2 - y^2$ the associated ideal is $L(d, \sqrt{d})$ which is the principal ideal $(\sqrt{d})$ since $d = \sqrt{d} \cdot \sqrt{d}$ so d is an element of $(\sqrt{d})$. For $dx^2 + xy - y^2$ the corresponding ideal is $L(d, \frac{1+\sqrt{\Delta}}{2})$ which is $(\frac{1+\sqrt{\Delta}}{2})$ since

$d = \frac{-1+\sqrt{\Delta}}{2} \cdot \frac{1+\sqrt{\Delta}}{2}$. In both cases the norm of the element $\sqrt{d}$ or $\frac{1+\sqrt{\Delta}}{2}$ generating the ideal is $-d$ so it is negative.

Thus in Case 3 the two strict equivalence classes of principal ideals correspond to the equivalence classes of the principal form and its negative, so these are the only two equivalence classes of forms exactly when all ideals are principal. □

An example for the third case in this proof is $\Delta = 12$ where the class number is 2 corresponding to the principal form $x^2 - 3y^2$ and its negative. The primes represented in discriminant 12 are $2, 3$, and the odd primes p with Legendre symbol $\left(\frac{12}{p}\right) = \left(\frac{3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{3}\right) = +1$ so these are the primes $p \equiv \pm 1 \bmod 12$. The two forms are of different genus, with $x^2 - 3y^2$ representing primes $p \equiv +1 \bmod 12$ and $-x^2 + 3y^2$ representing primes $p \equiv -1 \bmod 12$. By Proposition 8.7 the primes p that factor in $R_\Delta = \mathbb{Z}[\sqrt{3}]$ are the primes represented by either of the two forms, for example $2 = (\sqrt{3}+1)(\sqrt{3}-1)$, $3 = (\sqrt{3})^2$, $11 = (2\sqrt{3}+1)(2\sqrt{3}-1)$, and $13 = (4+\sqrt{3})(4-\sqrt{3})$. Here the factorization of 11 comes from the value -11 in the $\pm 1/2$ regions in the topograph of the principal form while the factorization of 13 comes from the 13 in the $\pm 4/1$ regions.



In this example prime factorizations are unique up to units, but there are infinitely many units for positive discriminants so there can be many factorizations that look rather different but are obtained just by inserting units. For example the topograph also gives $13 = (5+2\sqrt{3})(5-2\sqrt{3})$ from the $\pm 5/2$ regions so $5+2\sqrt{3}$ must be a unit times either $4+\sqrt{3}$ or $4-\sqrt{3}$. One can determine which by computing which of the two quotients $(5+2\sqrt{3})/(4+\sqrt{3})$ and $(5+2\sqrt{3})/(4-\sqrt{3})$ lies in $\mathbb{Z}[\sqrt{3}]$. One finds that the latter quotient is the unit $2+\sqrt{3}$ so $5+2\sqrt{3} = (2+\sqrt{3})(4-\sqrt{3})$. In terms of the topograph, multiplication by the fundamental unit $2+\sqrt{3}$ translates the topograph by one period to the right, while conjugation is reflection across the vertical line through the $1/0$ and $0/1$ regions. So to get from $4/1$ to $5/2$ we first reflect $4/1$ to $-4/1$, then we translate by one period to get $5/2$.

As this example shows, for prime factorizations it makes little difference if the principal form is not equivalent to its negative since changing the sign of an element of R_Δ is just multiplying it by the unit -1. The issue could be avoided entirely by using the version of the ideal class group based on equivalence of ideals rather than strict equivalence.

Let us conclude this section with some comments on what happens when the discriminant Δ is not a fundamental discriminant. One might hope that the unique factorization property for ideals still holds at least for stable ideals, the ideals corresponding to primitive forms. However, this is not the case, and here is an example. Take $\Delta = -12$, so $R_\Delta = \mathbb{Z}[\sqrt{-3}]$. The class number is 1 in this case so all stable ideals are principal (and recall that principal ideals are always stable). Consider the factorizations $(4) = (2)(2) = (1+\sqrt{-3})(1-\sqrt{-3})$. The ideals (2) and $(1\pm\sqrt{-3})$ are prime since their norms are 4 so any nontrivial factorization as $(\alpha)(\beta)$ would have $N(\alpha) = N(\beta) = 2$ but no elements of $\mathbb{Z}[\sqrt{-3}]$ have norm 2 since $x^2 + 3y^2 = 2$ has no integer solutions. The three ideals (2) and $(1\pm\sqrt{-3})$ are distinct since the only units in $\mathbb{Z}[\sqrt{-3}]$ are ± 1. Thus we have two different factorizations of (4) into prime ideals when we restrict attention just to stable ideals. If one drops this restriction then unique prime factorization still fails since for the ideal $L = (2, 1+\sqrt{-3})$ we saw in the discussion following Proposition 8.27 that $L^2 = 2L$, but unique factorization implies the cancellation property so we would then have $L = (2)$, which is false.

One might ask where the proof of unique factorization breaks down for stable ideals in the case of a nonfundamental discriminant. The answer is in the key property in Lemma 8.32 that if a prime ideal P divides a product LM then it must divide one of the factors L or M. In the proof of this we considered the ideal $P+L$, but unfortunately this need not be a stable ideal when P and L are stable. For example, in the preceding paragraph if we take $P = (2)$, $L = (1+\sqrt{-3})$, and $M = (1-\sqrt{-3})$ then $P+L$ is the ideal $(2, 1+\sqrt{-3})$, but this is not stable as we saw after Proposition 8.27. And in fact the ideal (2) does not divide either $(1+\sqrt{-3})$ or $(1-\sqrt{-3})$.

Exercises

1. (a) Find the ideals of norm 39 in $\mathbb{Z}[\sqrt{10}]$ and find the factorizations of these ideals into prime ideals.
(b) Do the same for the ideals in $\mathbb{Z}[\sqrt{10}]$ of norm 10, 15, and 30.

2. Let p_1, p_2, p_3, p_4 be distinct primes represented by the form $2x^2+3y^2$. Show that there is an element of $\mathbb{Z}[\sqrt{-6}]$ of norm $p_1p_2p_3p_4$ having three different factorizations as products of prime elements of $\mathbb{Z}[\sqrt{-6}]$, where factorizations that differ just by units are not regarded as different factorizations.

3. For a fundamental discriminant Δ let us define two ideals L and L' in R_Δ to be scale equivalent if there exist positive integers m and n such that $mL = nL'$. Show that the set of scale equivalence classes of ideals in R_Δ forms a group with respect to the usual multiplication of ideals, and determine the structure of this group.

8.6 Applications to Forms

As we have seen, ideals provide an alternative way of constructing the class group $CG(\Delta)$. One of the main uses of the group structure in $CG(\Delta)$ in Chapter 7 was in Theorem 7.7 which characterized the primitive forms of discriminant Δ representing a given number n in terms of the forms representing the prime factors of n, or prime-power factors in the case of primes dividing the conductor. When Δ is a fundamental discriminant the same characterization can be derived from the unique factorization property of ideals in R_Δ. This viewpoint provides additional insights into the somewhat subtle answer to the representation problem. Here is a restatement of the result we will now prove using ideals:

Theorem 8.38. *Let Δ be a fundamental discriminant and let $n > 1$ be a number represented by at least one form of discriminant Δ. If the prime factorization of n is $n = p_1^{e_1} \cdots p_k^{e_k}$ for distinct primes p_i, with $e_i = 1$ for each p_i dividing Δ and $e_i \geq 1$ otherwise, then the forms of discriminant Δ representing n are exactly the forms $Q_1^{\pm e_1} \cdots Q_k^{\pm e_k}$ where Q_i represents p_i and the product $Q_1^{\pm e_1} \cdots Q_k^{\pm e_k}$ is formed in the class group $CG(\Delta)$.*

There are a few facts that are used in the proof that we will explain in advance to avoid complicating the later arguments. The first is the elementary fact that an element α in R_Δ belongs to an ideal L if and only if the ideal (α) factors as $(\alpha) = LM$ for some ideal M. This is because α is an element of L exactly when the ideal (α) is contained in L, or in other words, when L divides (α), which means $(\alpha) = LM$ for some ideal M.

Next is a reformulation of what it means for a form Q_L to represent a number n. By definition, $Q_L(\alpha) = N(\alpha)/N(L)$ for α in L. Thus if we choose a basis α_1, α_2 for L regarded as a lattice and we let $\alpha = x\alpha_1 + y\alpha_2$ for integers x and y, then $Q_L(x,y) = N(x\alpha_1 + y\alpha_2)/N(L)$. For this to give a representation of n means that x and y are coprime. In terms of α this is saying that α is not a multiple $m\beta$ of any element β of L with $m > 1$. This last condition can be abbreviated to saying just that α is *primitive* in L.

We have also defined what it means for an ideal L to be primitive, namely, L is not a multiple mL' of any other ideal L' with $m > 1$, or equivalently, L is not divisible by any principal ideal (m) with $m > 1$. We could require m to be a prime without affecting the definition since if $L = mL'$ with $m = pq$ for p a prime then L is p times the ideal qL'. By Proposition 8.16 every ideal in R_Δ is equal to nL_Q for some integer $n \geq 1$ and some form Q of discriminant Δ, so the primitive ideals are exactly the ideals L_Q.

An equivalent way of formulating the condition for L to be primitive is to say that the factorization $L = P_1 \cdots P_k$ as a product of prime ideals satisfies the following two conditions:

(1) No P_i is a prime ideal (p) with p a prime integer. Thus each P_i has norm a prime rather than the square of a prime.

(2) There is no pair of factors P_i and P_j with $i \neq j$ such that $P_i = \overline{P}_j$. In particular if $P_i = \overline{P}_i$ then P_i can occur only once in the prime factorization of L.

Proof of Theorem 8.38: Suppose that a number $n > 1$ is represented by a form Q. From the correspondence between proper equivalence classes of forms and strict equivalence classes of ideals we may assume $Q = Q_L$ for some ideal L. Thus we have $n = Q_L(\alpha) = N(\alpha)/N(L)$ for some primitive α in L. Since n and $N(L)$ are positive, so is $N(\alpha)$.

We can reduce to the case that α is a positive integer by the following argument. We have $n = N(\alpha)/N(L) = N(\overline{\alpha}\alpha)/N(\overline{\alpha}L) = Q_{\overline{\alpha}L}(\overline{\alpha}\alpha)$. The element $\overline{\alpha}\alpha$ of $\overline{\alpha}L$ is primitive in $\overline{\alpha}L$ since if $\overline{\alpha}\alpha = q\overline{\alpha}\beta$ for some positive integer q and some β in L, then $\alpha = q\beta$ which forces q to be 1 since α is primitive in L. The integer $m = \overline{\alpha}\alpha$ is $N(\alpha)$ which is positive as noted above. Also, m is in $\overline{\alpha}L$ since α is in L. The ideals L and $\overline{\alpha}L$ are strictly equivalent since $N(\overline{\alpha}) = N(\alpha) > 0$, so the forms $Q_{\overline{\alpha}L}$ and Q_L are properly equivalent. This shows that we may take n to be represented as $n = Q_L(m)$ for some primitive positive integer m in the new L.

Next we reduce to the case that L is a primitive ideal. If L is not primitive we can write it as $L = qL'$ for some integer $q > 1$ with L' primitive. Since m is in $L = qL'$ we have $m = qr$ for some r in L', and in fact r must be an integer since $r = {}^{m}/_{q}$ and the only rational numbers in R_Δ are integers. Since m and q are positive, so is r. Also, r is primitive in L' since m is primitive in L and we are just rescaling m and L by a factor of ${}^{1}/_{q}$ to get r and L'. The equation $n = N(m)/N(L)$ can be written as $n = N(qr)/N(qL') = N(r)/N(L')$ since $qL' = (q)L'$ and $N((q)) = N(q)$. This shows that n is represented as $Q_{L'}(r) = n$. The form $Q_{L'}$ is properly equivalent to Q_L since $L = qL'$ and $N(q) > 0$. The net result of this argument is that we can assume that n is represented as $n = Q_L(m) = N(m)/N(L)$ where L is primitive and m is a positive integer that is a primitive element of L.

Since m is in L we have $(m) = LM$ for some ideal M. This M must also be primitive, otherwise if $M = qM'$ for some ideal M' and some integer $q > 1$, then, arguing as in the preceding paragraph, we would have $m = qr$ for some positive integer r in M' with $(r) = LM'$. This last equality implies that r is in L, so m would not be primitive in L.

Since L and M are both primitive, their factorizations into prime ideals satisfy the earlier conditions (1) and (2). Then since their product is (m) with m an integer, we must have $M = \overline{L}$. Thus $(m) = L\overline{L}$ and so $m = N(L)$. Now we have $n = N(m)/N(L) =$

$m^2/m = m$ so $n = m$ and the representation of n becomes $n = Q_L(n)$ with L primitive and $n = N(L)$.

Let the factorization of L into prime ideals be $L = P_1 \cdots P_k$. Then $N(P_i)$ is a prime p_i and p_i is in P_i since $P_i\overline{P}_i = (p_i)$. Also, p_i is primitive in P_i since p_i is prime so if p_i was not primitive in P_i then P_i would contain 1 which is impossible since $P_i \neq R_\Delta$. If we denote Q_{P_i} by Q_i for simplicity then Q_i represents p_i since $Q_i(p_i) = N(p_i)/N(P_i) = p_i^2/p_i = p_i$.

Since $n = N(L)$ and $L = P_1 \cdots P_k$ we have $n = p_1 \cdots p_k$. The prime factorization $n = p_1 \cdots p_k$ is unique so the prime ideals P_i are uniquely determined by n up to the ambiguity of replacing P_i by $\overline{P}_i$. In $CG(\Delta)$ this amounts to replacing Q_i by Q_i^{-1}. Keeping in mind condition (2), we have now shown that if a form Q represents n then in $CG(\Delta)$ we have $Q = Q_1^{\pm e_1} \cdots Q_k^{\pm e_k}$ where $n = p_1^{e_1} \cdots p_k^{e_k}$ is the factorization of n into powers of distinct primes p_i and the form Q_i represents p_i. Condition (2) implies that $e_i = 1$ for each i with $P_i = \overline{P}_i$, that is, for each p_i that divides the discriminant Δ.

To show the converse, suppose $n = p_1^{e_1} \cdots p_k^{e_k}$ is the factorization of n into powers of distinct primes p_i with $e_i = 1$ when p_i divides Δ, and suppose the form Q_i represents p_i. Our objective is then to show that $Q_1^{\pm e_1} \cdots Q_k^{\pm e_k}$ represents n. By the arguments in the first part of the proof applied to p_i in place of n there is an ideal L_i containing p_i with $N(L_i) = p_i$, so L_i is a prime ideal since its norm p_i is prime. If we set $L = L_1^{e_1} \cdots L_k^{e_k}$ then L is primitive since its factorization into prime ideals satisfies conditions (1) and (2). We have $n \in L$ since each p_i is in L_i. Also we have $N(L) = N(L_1)^{e_1} \cdots N(L_k)^{e_k} = p_1^{e_1} \cdots p_k^{e_k} = n$. Thus $Q_L(n) = N(n)/N(L) = n^2/n = n$ which means that Q_L represents n provided that n is primitive in L. If n is not primitive in L then it factors as $n = qr$ for some integer $q > 1$ and some r in L. By an earlier argument r must be a positive integer. Since r is in L, we have $(r) = LM$ for some ideal M. Then $(n) = (qr) = qLM$. We also have $(n) = L\overline{L}$ since $N(L) = n$. Thus $qLM = L\overline{L}$ so the cancellation property for ideals implies that $\overline{L} = qM$. Taking conjugates, this says $L = q\overline{M}$. This contradicts the fact that L is primitive. Thus we have shown that Q_L represents n.

We have $Q_{L_i}(p_i) = N(p_i)/N(L_i) = p_i^2/p_i = p_i$. Thus both Q_i and Q_{L_i} represent the prime p_i so they must be equivalent, hence in $CG(\Delta)$ we have $Q_{L_i} = Q_i^{\pm 1}$. We can choose the sign of the exponent at will since we are free to replace L_i by $\overline{L}_i$ in the previous arguments. Then $Q_1^{\pm e_1} \cdots Q_k^{\pm e_k} = Q_{L_1}^{e_1} \cdots Q_{L_k}^{e_k} = Q_L$ since $L = L_1^{e_1} \cdots L_k^{e_k}$. Thus $Q_1^{\pm e_1} \cdots Q_k^{\pm e_k}$ represents n since Q_L represents n □

As another application of unique factorization for ideals in the rings R_Δ for fundamental discriminants Δ let us consider again the problem of finding which primitive forms represent powers of primes dividing the conductor in the case of nonfundamental discriminants. The large table in Section 6.2 shows some of the subtleties that can occur for small negative nonfundamental discriminants. Perhaps the most interesting cases are when infinitely many different powers of these primes are represented. The first three cases $\Delta = -28$, -60, and -72 were treated in Sections 6.2, 6.3, and 8.2. Let us consider now the fourth case $\Delta = -92$ where there are some new subtleties.

For $\Delta = -92$ the class number is 3 with the three forms $x^2 + 23y^2$ and $3x^2 \pm 2xy + 8y^2$. The associated fundamental discriminant is $\Delta = -23$ which also has class number 3, corresponding to the forms $x^2 + xy + 6y^2$ and $2x^2 \pm xy + 3y^2$. The conductor is 2 and this is represented in discriminant -23 by $2x^2 \pm xy + 3y^2$, as are all powers of 2 since 2 does not divide -23, so by Proposition 6.13 all powers 2^k for $k \geq 3$ are represented by at least one of the forms $x^2 + 23y^2$ and $3x^2 \pm 2xy + 8y^2$. Our aim is to determine which of these powers are represented by each form.

First consider the form $x^2 + 23y^2$. For elements $x + \sqrt{-23}y$ in $\mathbb{Z}[\sqrt{-23}]$ we have $N(x + \sqrt{-23}y) = x^2 + 23y^2$ so we are looking for coprime integers x and y such that $x + \sqrt{-23}y$ has norm a power of 2. We will use the larger ring $\mathbb{Z}[\omega]$ with $\omega = (1 + \sqrt{-23})/2$ since this has unique factorization of ideals, being the ring R_Δ for the associated fundamental discriminant -23. Using Proposition 8.18 we see that the principal ideal (2) in $\mathbb{Z}[\omega]$ factors as $(2) = P\overline{P}$ for $P = (2, \omega)$, with $P \neq \overline{P}$. Since $N(2) = 4$ we have $N(P) = N(\overline{P}) = 2$, so $N(P^k) = 2^k$. The ideal P is not principal since there is no element of $\mathbb{Z}[\omega]$ of norm 2, for if α in $\mathbb{Z}[\omega]$ had norm 2 then 2α would be an element of $\mathbb{Z}[\sqrt{-23}]$ of norm 8 but the form $x^2 + 23y^2$ does not take on the value 8. Since the class number for discriminant -23 is 3 the class group is cyclic of order 3 and P generates this group. Thus the powers of P that are principal ideals are the powers P^{3n}.

Suppose the element $\alpha = x + \sqrt{-23}y$ of $\mathbb{Z}[\sqrt{-23}]$ has norm 2^k, so $\alpha\overline{\alpha} = 2^k$. Then for ideals we have $(\alpha)(\overline{\alpha}) = P^k\overline{P}^k$ and hence $(\alpha) = P^r\overline{P}^s$ for some r and s with $r + s = k$. We have $x^2 + 23y^2 = 2^k$ so x and y have the same parity. We want them to be coprime so this means they are both odd and hence α is divisible by 2 in $\mathbb{Z}[\omega]$. This is saying that (α) is divisible by both P and $\overline{P}$ since $(2) = P\overline{P}$. Thus $r > 0$ and $s > 0$. On the other hand if $r > 1$ and $s > 1$ this would say that (α) was divisible by (4) and hence α was divisible by 4 in $\mathbb{Z}[\omega]$, so x and y would both be even, a contradiction. Therefore one of r and s must be 1, and so in the class group where $\overline{P}$ is the inverse of P the ideal (α) must be either $2P^{k-2}$ if $s = 1$, or $2\overline{P}^{k-2}$ if $r = 1$. Since (α) is a principal ideal this implies that $k-2$ is a multiple of 3, say $k - 2 = 3m$, or $k = 3m + 2$. Thus the only powers of 2 that could possibly be represented by $x^2 + 23y^2$ are the powers 2^k with $k = 2, 5, 8, \cdots$. Obviously 2^2 is not represented so this leaves $2^5, 2^8, 2^{11}, \cdots$. as the only possibilities.

The other two forms $3x^2 \pm 2xy + 8y^2$ are equivalent, though not properly equivalent, so they represent the same numbers. We will show that they cannot represent any of the powers $2^5, 2^8, 2^{11}, \cdots$. Since each power 2^k with $k \geq 3$ is represented by one of the forms $x^2 + 23y^2$ and $3x^2 \pm 2xy + 8y^2$ we will then know that $x^2 + 23y^2$ represents $2^5, 2^8, 2^{11}, \cdots$ and $3x^2 \pm 2xy + 8y^2$ represents $2^3, 2^4, 2^6, 2^7, 2^9, 2^{10}, \cdots$.

The lattice in $\mathbb{Z}[\sqrt{-23}]$ corresponding to $3x^2+2xy+8y^2$ is $L(3,1+\sqrt{-23})$. This has norm 3 in $\mathbb{Z}[\sqrt{-23}]$ so we have $N(3x+(1+\sqrt{-23})y)/3 = N((3x+y)+\sqrt{-23}y)/3 = (9x^2+6xy+y^2+23y^2)/3 = 3x^2+2xy+8y^2$, the given form.

Suppose that x and y are coprime integers for which $3x^2+2xy+8y^2=2^k$. The element $\alpha = 3x+(1+\sqrt{-23})y = 3x+2\omega y$ in $\mathbb{Z}[\sqrt{-23}]$ then has $N(\alpha)=3\cdot 2^k$. In $\mathbb{Z}[\omega]$ we have $(2)=P\overline{P}$ for $P=(2,\omega)$, and we have $(3)=Q\overline{Q}$ for $Q=(3,\omega)$ from Proposition 8.18. Thus $(\alpha)(\overline{\alpha}) = Q\overline{Q}P^k\overline{P}^k$ and hence (α) is either $QP^r\overline{P}^s$ or $\overline{Q}P^r\overline{P}^s$ for some integers $r\ge 0$ and $s\ge 0$ with $r+s=k$. The equation $3x^2+2xy+8y^2=2^k$ implies that x is even, hence $3x+2\omega y$ is divisible by 2 in $\mathbb{Z}[\omega]$. This implies that $r>0$ and $s>0$. If $r>1$ and $s>1$ then 4 divides $3x+2\omega y$ in $\mathbb{Z}[\omega]$ which implies x and y are even, violating their coprimeness. Thus either $r=1$ or $s=1$, say $s=1$. This means $(\alpha)=2QP^{k-2}$ or $(\alpha)=2\overline{Q}P^{k-2}$. Since (α) is a principal ideal this means that QP^{k-2} or $\overline{Q}P^{k-2}$ is a principal ideal. The product PQ is $(2,\omega)(3,\omega)=(6,2\omega,3\omega,\omega^2)$ with $\omega^2=\omega-6$. It follows that $PQ=(\omega)$ since $\omega = 3\omega-2\omega$ and $6=\omega\overline{\omega}$. Since PQ is a principal ideal, Q is the inverse of P in the class group and $\overline{Q}$ is equivalent to P.

In the case $(\alpha)=2QP^{k-2}$ the ideal (α) is principal and is equivalent to P^{k-3} in the class group so $k-3=3n$ for some integer n, which means $k\equiv 0 \bmod 3$. In the other case $(\alpha)=2\overline{Q}P^{k-2}$ we have (α) equivalent to P^{k-1} in the class group so $k-1=3n$ and $k\equiv 1 \bmod 3$. This finishes the argument that the forms $3x^2\pm 2xy+8y^2$ cannot represent any of the powers $2^5, 2^8, 2^{11}, \cdots$. Hence we know which powers of 2 each form x^2+23y^2 and $3x^2\pm 2xy+8y^2$ represents.

It is easy to be more explicit about representing 2^{3n+2} by x^2+23y^2. As we have seen, this amounts to writing the principal ideal $2P^{3n}$ as $(x+\sqrt{-23}y)$. The ideal P^3 has norm 8 so it must equal (β) for some β in $\mathbb{Z}[\omega]$ of norm 8. From the topograph of the norm form $x^2+xy+6y^2$ in discriminant -23 one can see that $1+\omega$ and $1+\overline{\omega}=2-\omega$ are the only elements of $\mathbb{Z}[\omega]$ of norm 8, up to sign. Thus we obtain solutions of $x^2+23y^2=2^{3n+2}$ by writing $2\cdot(1+\omega)^n$ as $x+\sqrt{-23}y$, and these are the only primitive solutions, up to changing the signs of x and y. We can compute inductively, so if $2\cdot(1+\omega)^n = x+\sqrt{-23}y$ then multiplying this by $1+\omega$ gives the solution for the next value of n. Since $1+\omega=\frac{3+\sqrt{-23}}{2}$ the inductive formula is:

$$(x+\sqrt{-23}y)\left(\frac{3+\sqrt{-23}}{2}\right) = \frac{(3x-23y)+(x+3y)\sqrt{-23}}{2}$$

Here are the first few solutions:

n	1	2	3	4	5
(x,y)	$(3,1)$	$(-7,3)$	$(-45,1)$	$(-79,-21)$	$(123,-71)$

One could also be explicit about solutions of $3x^2+2xy+8y^2=2^k$ but the answers are a little more complicated so we will not do this here.

Tables

1. Forms of Negative Discriminant

This table lists the proper equivalence classes of primitive forms for each negative discriminant down to -120. The first column gives the discriminant (up to sign), with an asterisk when it is not a fundamental discriminant. The second column gives the class number. In most cases in the table the class group is cyclic so the class number determines the class group. The exceptions are indicated by writing the class number as a product corresponding to the factorization of the class group as a product of cyclic groups. Thus $2\cdot 2$ means class number 4 with class group the product of two cyclic groups of order 2. The third column gives the various characters for each discriminant. These correspond to the prime divisors of the discriminant, with a few exceptions for the prime 2 in cases with nonfundamental discriminants. The fourth column gives the reduced form for each equivalence class, with $ax^2+bxy+cy^2$ abbreviated to $[a,b,c]$, followed by signs $+$ and $-$ indicating whether the characters have value $+1$ or -1 on each form. The forms in each genus have the same character values, and these forms are listed consecutively. Forms that lack mirror symmetry have middle coefficients $\pm b$, indicating that the form and its mirror image give distinct elements of the class group.

$\|\Delta\|$	h_Δ	Char.	Forms
3	1	χ_3	$[1,1,1]$ $+$
4	1	χ_4	$[1,0,1]$ $+$
7	1	χ_7	$[1,1,2]$ $+$
8	1	χ_8'	$[1,0,2]$ $+$
11	1	χ_{11}	$[1,1,3]$ $+$
$*$ 12	1	χ_3	$[1,0,3]$ $+$
15	2	χ_3 χ_5	$[1,1,4]$ $++$ $[2,1,2]$ $--$
$*$ 16	1	χ_4	$[1,0,4]$ $+$
19	1	χ_{19}	$[1,1,5]$ $+$
20	2	χ_4 χ_5	$[1,0,5]$ $++$ $[2,2,3]$ $--$

$\|\Delta\|$	h_Δ	Char.	Forms
23	3	χ_{23}	$[1,1,6]$ $+$ $[2,\pm 1,3]$ $+$
24	2	χ_8 χ_3	$[1,0,6]$ $++$ $[2,0,3]$ $--$
$*$ 27	1	χ_3	$[1,1,7]$ $+$
$*$ 28	1	χ_7	$[1,0,7]$ $+$
31	3	χ_{31}	$[1,1,8]$ $+$ $[2,\pm 1,4]$ $+$
$*$ 32	2	χ_4 χ_8	$[1,0,8]$ $++$ $[3,2,3]$ $--$
35	2	χ_5 χ_7	$[1,1,9]$ $++$ $[3,1,3]$ $--$
$*$ 36	2	χ_4 χ_3	$[1,0,9]$ $++$ $[2,2,5]$ $+-$

$\lvert\Delta\rvert$	h_Δ	Char.	Forms
39	4	$\chi_3\ \chi_{13}$	$[1,1,10]$ $++$ $[3,3,4]$ $++$ $[2,\pm1,5]$ $--$
40	2	$\chi_8'\ \chi_5$	$[1,0,10]$ $++$ $[2,0,5]$ $--$
43	1	χ_{43}	$[1,1,11]$ $+$
∗ 44	3	χ_{11}	$[1,0,11]$ $+$ $[3,\pm2,4]$ $+$
47	5	χ_{47}	$[1,1,12]$ $+$ $[2,\pm1,6]$ $+$ $[3,\pm1,4]$ $+$
∗ 48	2	$\chi_4\ \chi_3$	$[1,0,12]$ $++$ $[3,0,4]$ $-+$
51	2	$\chi_3\ \chi_{17}$	$[1,1,13]$ $++$ $[3,3,5]$ $--$
52	2	$\chi_4\ \chi_{13}$	$[1,0,13]$ $++$ $[2,2,7]$ $--$
55	4	$\chi_5\ \chi_{11}$	$[1,1,14]$ $++$ $[4,3,4]$ $++$ $[2,\pm1,7]$ $--$
56	4	$\chi_8\ \chi_7$	$[1,0,14]$ $++$ $[2,0,7]$ $++$ $[3,\pm2,5]$ $--$
59	3	χ_{59}	$[1,1,15]$ $+$ $[3,\pm1,5]$ $+$
∗ 60	2	$\chi_3\ \chi_5$	$[1,0,15]$ $++$ $[3,0,5]$ $--$
∗ 63	4	$\chi_3\ \chi_7$	$[1,1,16]$ $++$ $[4,1,4]$ $++$ $[2,\pm1,8]$ $-+$
∗ 64	2	$\chi_4\ \chi_8$	$[1,0,16]$ $++$ $[4,4,5]$ $+-$
67	1	χ_{67}	$[1,1,17]$ $+$
68	4	$\chi_4\ \chi_{17}$	$[1,0,17]$ $++$ $[2,2,9]$ $++$ $[3,\pm2,6]$ $--$
71	7	χ_{71}	$[1,1,18]$ $+$ $[2,\pm1,9]$ $+$ $[3,\pm1,6]$ $+$ $[4,\pm3,5]$ $+$

$\lvert\Delta\rvert$	h_Δ	Char.	Forms
∗ 72	2	$\chi_8'\ \chi_3$	$[1,0,18]$ $++$ $[2,0,9]$ $+-$
∗ 75	2	$\chi_3\ \chi_5$	$[1,1,19]$ $++$ $[3,3,7]$ $+-$
∗ 76	3	χ_{19}	$[1,0,19]$ $+$ $[4,\pm2,5]$ $+$
79	5	χ_{79}	$[1,1,20]$ $+$ $[2,\pm1,10]$ $+$ $[4,\pm1,5]$ $+$
∗ 80	4	$\chi_4\ \chi_5$	$[1,0,20]$ $++$ $[4,0,5]$ $++$ $[3,\pm2,7]$ $--$
83	3	χ_{83}	$[1,1,21]$ $+$ $[3,\pm1,7]$ $+$
84	2·2	$\chi_4\ \chi_3\ \chi_7$	$[1,0,21]$ $+++$ $[2,2,11]$ $--+$ $[3,0,7]$ $-+-$ $[5,4,5]$ $+--$
87	6	$\chi_3\ \chi_{29}$	$[1,1,22]$ $++$ $[4,\pm3,6]$ $++$ $[2,\pm1,11]$ $--$ $[3,3,8]$ $--$
88	2	$\chi_8\ \chi_{11}$	$[1,0,22]$ $++$ $[2,0,11]$ $--$
91	2	$\chi_7\ \chi_{13}$	$[1,1,23]$ $++$ $[5,3,5]$ $--$
∗ 92	3	χ_{23}	$[1,0,23]$ $+$ $[3,\pm2,8]$ $+$
95	8	$\chi_5\ \chi_{19}$	$[1,1,24]$ $++$ $[4,\pm1,6]$ $++$ $[5,5,6]$ $++$ $[2,\pm1,12]$ $--$ $[3,\pm1,8]$ $--$
∗ 96	2·2	$\chi_4\ \chi_8\ \chi_3$	$[1,0,24]$ $+++$ $[3,0,8]$ $---$ $[4,4,7]$ $-++$ $[5,2,5]$ $+--$
∗ 99	2	$\chi_3\ \chi_{11}$	$[1,1,25]$ $++$ $[5,1,5]$ $-+$

$\lvert\Delta\rvert$	h_Δ	Char.	Forms	
$*$ 100	2	$\chi_4\ \ \chi_5$	$[1,0,25]$	$+\,+$
			$[2,2,13]$	$+\,-$
103	5	χ_{103}	$[1,1,26]$	$+$
			$[2,\pm 1,13]$	$+$
			$[4,\pm 3,7]$	$+$
104	6	$\chi_8\ \ \chi_{13}$	$[1,0,26]$	$+\,+$
			$[3,\pm 2,9]$	$+\,+$
			$[2,0,13]$	$-\,-$
			$[5,\pm 4,6]$	$-\,-$
107	3	χ_{107}	$[1,1,27]$	$+$
			$[3,\pm 1,9]$	$+$
$*$ 108	3	χ_3	$[1,0,27]$	$+$
			$[4,\pm 2,7]$	$+$
111	8	$\chi_3,\ \chi_{37}$	$[1,1,28]$	$+\,+$
			$[4,\pm 1,7]$	$+\,+$
			$[3,3,10]$	$+\,+$
			$[2,\pm 1,14]$	$-\,-$
			$[5,\pm 3,6]$	$-\,-$
$*$ 112	2	$\chi_4,\ \chi_7$	$[1,0,28]$	$+\,+$
			$[4,0,7]$	$-\,+$
115	2	$\chi_5,\ \chi_{23}$	$[1,1,29]$	$+\,+$
			$[5,5,7]$	$-\,-$
116	6	$\chi_4,\ \chi_{29}$	$[1,0,29]$	$+\,+$
			$[5,\pm 2,6]$	$+\,+$
			$[2,2,15]$	$-\,-$
			$[3,\pm 2,10]$	$-\,-$
119	10	$\chi_7,\ \chi_{17}$	$[1,1,30]$	$+\,+$
			$[2,\pm 1,15]$	$+\,+$
			$[4,\pm 3,8]$	$+\,+$
			$[3,\pm 1,10]$	$-\,-$
			$[5,\pm 1,6]$	$-\,-$
			$[6,5,6]$	$-\,-$
120	$2\cdot 2$	$\chi_8,\ \chi_3,\ \chi_5$	$[1,0,30]$	$+\ +\ +$
			$[2,0,15]$	$+\ -\ -$
			$[3,0,10]$	$-\ +\ -$
			$[5,0,6]$	$-\ -\ +$

2. Fully Symmetric Negative Discriminants

Listed below are the 101 known negative discriminants Δ for which every primitive form has a mirror-symmetric topograph. This is equivalent to saying that each genus consists of a single equivalence class of forms, or that the class group is either the trivial group or a product of cyclic groups of order 2. The class number h_Δ is then a power of 2 determined by the number of distinct prime divisors of Δ. Asterisks in the table denote nonfundamental discriminants. Among the 101 discriminants there are 65 fundamental discriminants and, coincidentally, 65 even discriminants.

$\lvert\Delta\rvert$	h_Δ
3	1
$4 = 2^2$	1
7	1
$8 = 2^3$	1
11	1
$* \ 12 = 2^2 \cdot 3$	1
$15 = 3 \cdot 5$	2
$* \ 16 = 2^4$	1
19	1
$20 = 2^2 \cdot 5$	2
$24 = 2^3 \cdot 3$	2
$* \ 27 = 3^3$	1
$* \ 28 = 2^2 \cdot 7$	1
$* \ 32 = 2^5$	2
$35 = 5 \cdot 7$	2
$* \ 36 = 2^2 \cdot 3^2$	2
$40 = 2^3 \cdot 5$	2
43	1
$* \ 48 = 2^4 \cdot 3$	2
$51 = 3 \cdot 17$	2
$52 = 2^2 \cdot 13$	2
$* \ 60 = 2^2 \cdot 3 \cdot 5$	2
$* \ 64 = 2^6$	2
67	1
$* \ 72 = 2^3 \cdot 3^2$	2
$* \ 75 = 3 \cdot 5^2$	2
$84 = 2^2 \cdot 3 \cdot 7$	4
$88 = 2^3 \cdot 11$	2
$91 = 7 \cdot 13$	2
$* \ 96 = 2^5 \cdot 3$	4
$* \ 99 = 3^2 \cdot 11$	2
$* \ 100 = 2^2 \cdot 5^2$	2
$* \ 112 = 2^4 \cdot 7$	2
$115 = 5 \cdot 23$	2

$\lvert\Delta\rvert$	h_Δ
$120 = 2^3 \cdot 3 \cdot 5$	4
$123 = 3 \cdot 41$	2
$132 = 2^2 \cdot 3 \cdot 11$	4
$* \ 147 = 3 \cdot 49$	2
$148 = 4 \cdot 37$	2
$* \ 160 = 2^5 \cdot 5$	4
163	1
$168 = 2^3 \cdot 3 \cdot 7$	4
$* \ 180 = 2^2 \cdot 3^2 \cdot 5$	4
$187 = 11 \cdot 17$	2
$* \ 192 = 2^6 \cdot 3$	4
$195 = 3 \cdot 5 \cdot 13$	4
$228 = 2^2 \cdot 3 \cdot 19$	4
$232 = 2^3 \cdot 29$	2
$235 = 5 \cdot 47$	2
$* \ 240 = 2^4 \cdot 3 \cdot 5$	4
$267 = 3 \cdot 89$	2
$280 = 2^3 \cdot 5 \cdot 7$	4
$* \ 288 = 2^5 \cdot 3^2$	4
$312 = 2^3 \cdot 3 \cdot 13$	4
$* \ 315 = 3^2 \cdot 5 \cdot 7$	4
$340 = 2^2 \cdot 5 \cdot 17$	4
$* \ 352 = 2^5 \cdot 11$	4
$372 = 2^2 \cdot 3 \cdot 31$	4
$403 = 13 \cdot 31$	2
$408 = 2^3 \cdot 3 \cdot 17$	4
$420 = 2^2 \cdot 3 \cdot 5 \cdot 7$	8
$427 = 7 \cdot 61$	2
$435 = 3 \cdot 5 \cdot 29$	4
$* \ 448 = 2^6 \cdot 7$	4
$* \ 480 = 2^5 \cdot 3 \cdot 5$	8
$483 = 3 \cdot 7 \cdot 23$	4
$520 = 2^3 \cdot 5 \cdot 13$	4
$532 = 2^2 \cdot 7 \cdot 19$	4

$\lvert\Delta\rvert$	h_Δ
$555 = 3 \cdot 5 \cdot 37$	4
$595 = 5 \cdot 7 \cdot 17$	4
$627 = 3 \cdot 11 \cdot 19$	4
$660 = 2^2 \cdot 3 \cdot 5 \cdot 11$	8
$* \ 672 = 2^5 \cdot 3 \cdot 7$	8
$708 = 2^2 \cdot 3 \cdot 59$	4
$715 = 5 \cdot 11 \cdot 13$	4
$760 = 2^3 \cdot 5 \cdot 19$	4
$795 = 3 \cdot 5 \cdot 53$	4
$840 = 2^3 \cdot 3 \cdot 5 \cdot 7$	8
$* \ 928 = 2^5 \cdot 29$	4
$* \ 960 = 2^6 \cdot 3 \cdot 5$	8
$1012 = 2^2 \cdot 11 \cdot 23$	4
$1092 = 2^2 \cdot 3 \cdot 7 \cdot 13$	8
$* \ 1120 = 2^5 \cdot 5 \cdot 7$	8
$1155 = 3 \cdot 5 \cdot 7 \cdot 11$	8
$* \ 1248 = 2^5 \cdot 3 \cdot 13$	8
$1320 = 2^3 \cdot 3 \cdot 5 \cdot 11$	8
$1380 = 2^2 \cdot 3 \cdot 5 \cdot 23$	8
$1428 = 2^2 \cdot 3 \cdot 7 \cdot 17$	8
$1435 = 5 \cdot 7 \cdot 41$	4
$1540 = 2^2 \cdot 5 \cdot 7 \cdot 11$	8
$* \ 1632 = 2^5 \cdot 3 \cdot 17$	8
$1848 = 2^3 \cdot 3 \cdot 7 \cdot 11$	8
$1995 = 3 \cdot 5 \cdot 7 \cdot 11$	8
$* \ 2080 = 2^5 \cdot 5 \cdot 13$	8
$3003 = 3 \cdot 7 \cdot 11 \cdot 13$	8
$* \ 3040 = 2^5 \cdot 5 \cdot 19$	8
$3315 = 3 \cdot 5 \cdot 13 \cdot 17$	8
$* \ 3360 = 2^5 \cdot 3 \cdot 5 \cdot 7$	16
$* \ 5280 = 2^5 \cdot 3 \cdot 5 \cdot 11$	16
$5460 = 2^2 \cdot 3 \cdot 5 \cdot 7 \cdot 13$	16
$* \ 7392 = 2^5 \cdot 3 \cdot 7 \cdot 11$	16

3. Forms of Positive Nonsquare Discriminant

This table is similar in layout to Table 1. For positive discriminants there is not a unique reduced form within each equivalence class so we have chosen a form which seemed simplest in some less precise sense.

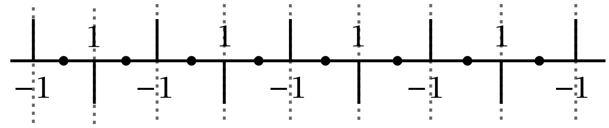
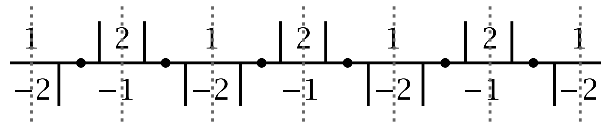
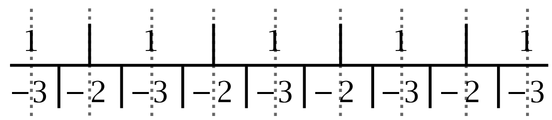
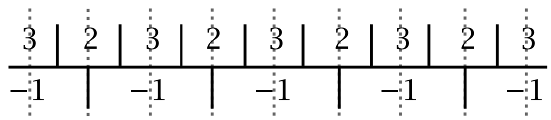
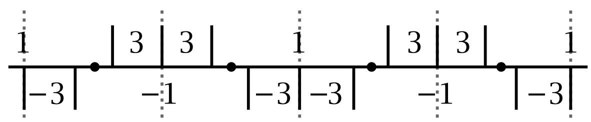
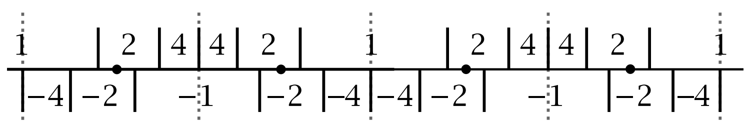
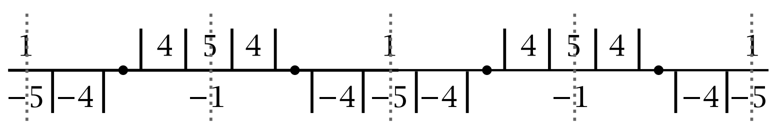
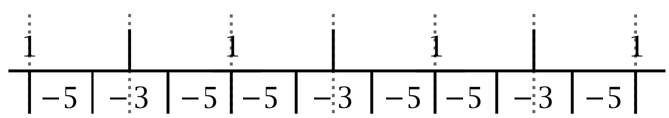
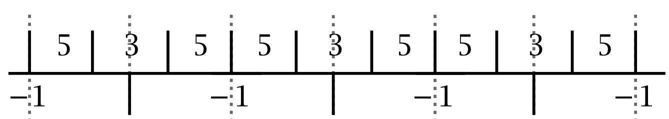
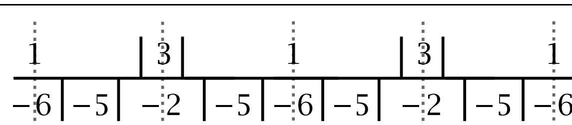
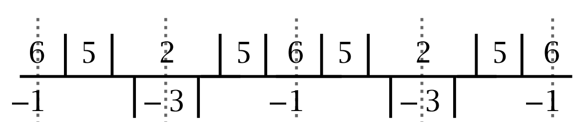
Δ	h_Δ	Char.	Forms
5	1	χ_5	$[1,1,-1]$ $+$
8	1	χ_8	$[1,0,-2]$ $+$
12	2	$\chi_4\ \chi_3$	$[1,0,-3]$ $++$ $[3,0,-1]$ $--$
13	1	χ_{13}	$[1,1,-3]$ $+$
17	1	χ_{17}	$[1,1,-4]$ $+$
$*$ 20	1	χ_5	$[1,0,-5]$ $+$
21	2	$\chi_3\ \chi_7$	$[1,1,-5]$ $++$ $[5,1,-1]$ $--$
24	2	$\chi_8'\ \chi_3$	$[1,0,-6]$ $++$ $[6,0,-1]$ $--$
28	2	$\chi_4\ \chi_7$	$[1,0,-7]$ $++$ $[7,0,-1]$ $--$
29	1	χ_{29}	$[1,1,-7]$ $+$
$*$ 32	2	$\chi_4\ \chi_8$	$[1,0,-8]$ $++$ $[8,0,-1]$ $-+$
33	2	$\chi_3\ \chi_{11}$	$[1,1,-8]$ $++$ $[8,1,-1]$ $--$
37	1	χ_{37}	$[1,1,-9]$ $+$
40	2	$\chi_8\ \chi_5$	$[1,0,-10]$ $++$ $[2,0,-5]$ $--$
41	1	χ_{41}	$[1,1,-10]$ $+$
44	2	$\chi_4\ \chi_{11}$	$[1,0,-11]$ $++$ $[11,0,-1]$ $--$
$*$ 45	2	$\chi_3\ \chi_5$	$[1,1,-11]$ $++$ $[11,1,-1]$ $-+$
$*$ 48	2	$\chi_4\ \chi_3$	$[1,0,-12]$ $++$ $[12,0,-1]$ $--$
$*$ 52	1	χ_{13}	$[1,0,-13]$ $+$
53	1	χ_{53}	$[1,1,-13]$ $+$
56	2	$\chi_8'\ \chi_7$	$[1,0,-14]$ $++$ $[14,0,-1]$ $--$
57	2	$\chi_3\ \chi_{19}$	$[1,1,-14]$ $++$ $[14,1,-1]$ $--$
60	2·2	$\chi_4\ \chi_3\ \chi_5$	$[1,0,-15]$ $+++$ $[15,0,-1]$ $--+$ $[3,0,-5]$ $-+-$ $[5,0,-3]$ $+--$
61	1	χ_{61}	$[1,0,-15]$ $+$
65	2	$\chi_5\ \chi_{13}$	$[1,1,-16]$ $++$ $[2,1,-8]$ $--$
$*$ 68	1	χ_{17}	$[1,0,-17]$ $+$
69	2	$\chi_3\ \chi_{23}$	$[1,1,-17]$ $++$ $[17,1,-1]$ $--$
$*$ 72	2	$\chi_8\ \chi_3$	$[1,0,-18]$ $++$ $[18,0,-1]$ $+-$
73	1	χ_{73}	$[1,1,-18]$ $+$
76	2	$\chi_4\ \chi_{19}$	$[1,0,-19]$ $++$ $[19,0,-1]$ $--$
77	2	$\chi_7\ \chi_{11}$	$[1,1,-19]$ $++$ $[19,1,-1]$ $--$
$*$ 80	2	$\chi_4\ \chi_5$	$[1,0,-20]$ $++$ $[20,0,-1]$ $-+$
$*$ 84	2	$\chi_3\ \chi_7$	$[1,0,-21]$ $++$ $[21,0,-1]$ $--$
85	2	$\chi_5\ \chi_{17}$	$[1,1,-21]$ $++$ $[3,1,-7]$ $--$
88	2	$\chi_8'\ \chi_{11}$	$[1,0,-22]$ $++$ $[22,0,-1]$ $--$
89	1	χ_{89}	$[1,1,-22]$ $+$
92	2	$\chi_4\ \chi_{23}$	$[1,0,-23]$ $++$ $[23,0,-1]$ $--$
93	2	$\chi_3\ \chi_{31}$	$[1,1,-23]$ $++$ $[23,1,-1]$ $--$

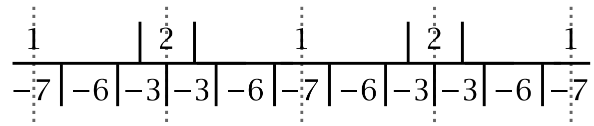
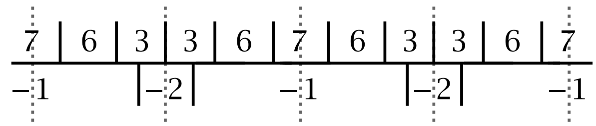
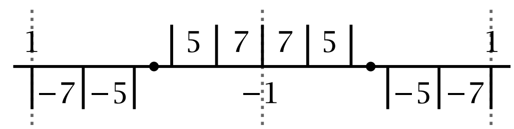
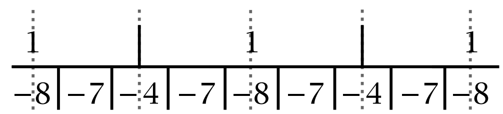
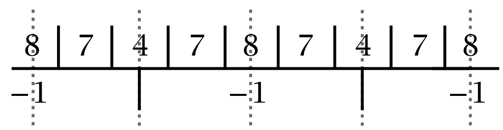
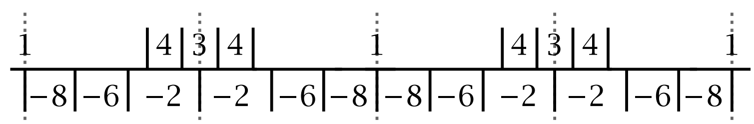
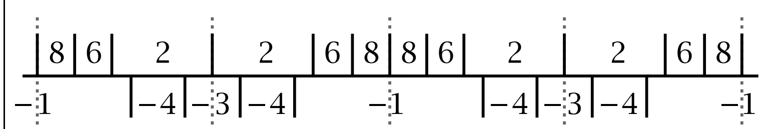
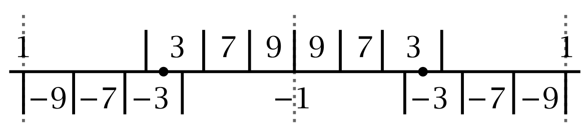
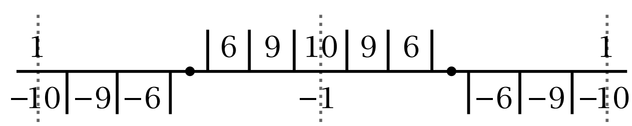
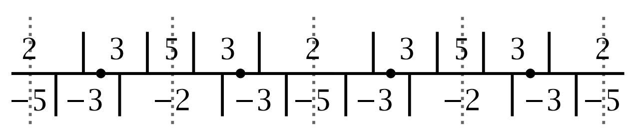
Δ	h_Δ	Char.	Forms	
∗ 96	2·2	$\chi_4\ \chi_8\ \chi_3$	[1, 0, −24]	+ + +
			[24, 0, −1]	− + −
			[3, 0, −8]	− − +
			[8, 0, −3]	+ − −
97	1	χ_{97}	[1, 1, −24]	+
101	1	χ_{101}	[1, 1, −25]	+
104	2	$\chi_8\ \chi_{13}$	[1, 0, −26]	+ +
			[2, 0, −13]	− −
105	2·2	$\chi_3\ \chi_5\ \chi_7$	[1, 1, −26]	+ + +
			[26, 1, −1]	− + −
			[2, 1, −13]	− − +
			[13, 1, −2]	+ − −
∗ 108	2	$\chi_4\ \chi_3$	[1, 0, −27]	+ +
			[27, 0, −1]	− −
109	1	χ_{109}	[1, 1, −27]	+
∗ 112	2	$\chi_4\ \chi_7$	[1, 0, −28]	+ +
			[28, 0, −1]	− −
113	1	χ_{113}	[1, 1, −28]	+
116	1	χ_{29}	[1, 1, −29]	+
∗ 117	2	$\chi_3\ \chi_{13}$	[1, 1, −29]	+ +
			[29, 1, −1]	− +
120	2·2	$\chi_8'\ \chi_3\ \chi_5$	[1, 0, −30]	+ + +
			[30, 0, −1]	− − +
			[2, 0, −15]	+ − −
			[15, 0, −2]	− + −
124	2	$\chi_4\ \chi_{31}$	[1, 0, −31]	+ +
			[31, 0, −1]	− −
∗ 125	1	χ_5	[1, 1, −31]	+
∗ 128	2	$\chi_4\ \chi_8$	[1, 0, −32]	+ +
			[32, 0, −1]	− +
129	2	$\chi_3\ \chi_{43}$	[1, 1, −32]	+ +
			[32, 1, −1]	− −
∗ 132	2	$\chi_3\ \chi_{11}$	[1, 0, −33]	+ +
			[33, 0, −1]	− −
133	2	$\chi_7\ \chi_{19}$	[1, 1, −33]	+ +
			[33, 1, −1]	− −
136	4	$\chi_8\ \chi_{17}$	[1, 0, −34]	+ +
			[34, 0, −1]	+ +
			[3, ±2, −11]	− −

Δ	h_Δ	Char.	Forms	
137	1	χ_{137}	[1, 1, −34]	+
140	2·2	$\chi_4\ \chi_5\ \chi_7$	[1, 0, −35]	+ + +
			[35, 0, −1]	− + −
			[2, 2, −17]	− − +
			[17, 2, −2]	+ − −
141	2	$\chi_3\ \chi_{47}$	[1, 1, −35]	+ +
			[35, 1, −1]	− −
145	4	$\chi_5\ \chi_{29}$	[1, 1, −36]	+ +
			[4, 1, −9]	+ +
			[2, ±1, −18]	− −
∗ 148	3	χ_{37}	[1, 0, −37]	+
			[3, ±2, −12]	+
149	1	χ_{149}	[1, 1, −37]	+
152	2	$\chi_8'\ \chi_{19}$	[1, 0, −38]	+ +
			[38, 0, −1]	− −
∗ 153	2	$\chi_3\ \chi_{17}$	[1, 1, −38]	+ +
			[38, 1, −1]	− +
156	2·2	$\chi_4\ \chi_3\ \chi_{13}$	[1, 0, −39]	+ + +
			[39, 0, −1]	− − +
			[2, 2, −19]	+ − −
			[19, 2, −2]	− + −
157	1	χ_{157}	[1, 1, −39]	+
∗ 160	2·2	$\chi_4\ \chi_8\ \chi_5$	[1, 0, −40]	+ + +
			[40, 0, −1]	− + +
			[3, 2, −13]	− − −
			[13, 2, −3]	+ − −
161	2	$\chi_7\ \chi_{23}$	[1, 1, −40]	+ +
			[40, 1, −1]	− −
∗ 164	1	χ_{41}	[1, 0, −41]	+
165	2·2	$\chi_3\ \chi_5\ \chi_{11}$	[1, 1, −41]	+ + +
			[41, 1, −1]	− + −
			[3, 3, −13]	− − +
			[13, 3, −3]	+ − −
168	2·2	$\chi_8\ \chi_3\ \chi_7$	[1, 0, −42]	+ + +
			[42, 0, −1]	+ − −
			[2, 0, −21]	− − +
			[21, 0, −2]	− + −
172	2	$\chi_4\ \chi_{43}$	[1, 0, −43]	+ +
			[43, 0, −1]	− −
173	1	χ_{173}	[1, 1, −43]	+

4. Periodic Separator Lines

The dotted vertical lines are lines of mirror symmetry and the heavy dots along the separator lines are points of rotational skew symmetry.

Δ	Q	
5	$[1,1,-1]$	
8	$[1,0,-2]$	
12	$[1,0,-3]$	
	$[3,0,-1]$	
13	$[1,1,-3]$	
17	$[1,1,-4]$	
20	$[1,0,-5]$	
21	$[1,1,-5]$	
	$[5,1,-1]$	
24	$[1,0,-6]$	
	$[6,0,-1]$	

Δ	Q	
28	$[1, 0, -7]$	
	$[7, 0, -1]$	
29	$[1, 1, -7]$	
32	$[1, 0, -8]$	
	$[8, 0, -1]$	
33	$[1, 1, -8]$	
	$[8, 1, -1]$	
37	$[1, 1, -9]$	
40	$[1, 0, -10]$	
	$[2, 0, -5]$	

Nonstandard Terminology

In a few instances we have chosen not to use standard terminology for certain concepts, usually because the traditional names seem somewhat awkward in the context of this book, or not as suggestive of the meaning as they could be. Here is a short summary of the main instances where translation may be needed when reading other sources.

Quadratic Forms. These are usually divided into three types, but for our purposes it is useful to split one of the three types into two for a total of four types as defined at the beginning of Chapter 5. Here are the traditional names with our equivalents:

- definite = elliptic
- indefinite = hyperbolic or 0-hyperbolic
- semidefinite = parabolic

Besides the convenience of having separate names for hyperbolic and 0-hyperbolic forms, the other motivation for the change is that the ordinary meanings of "definite" and "indefinite" do not seem to convey very well their mathematical meanings.

What we call a symmetry of a quadratic form is more often called an automorph or automorphism of the form, although the latter terms are sometimes reserved just for orientation-preserving symmetries. We call a form having an orientation-reversing symmetry a mirror symmetric form, or a form with mirror symmetry, whereas classically such forms are called ambiguous, a term that has suffered somewhat in the translation from Gauss's original Latin.

Representing Numbers by Quadratic Forms. The traditional terminology is to say that a quadratic form $Q(x,y)$ represents a number n when there exist integers x and y such that $Q(x,y) = n$. However in this book we are almost always interested only in the case that x and y are coprime, so to avoid extra words to specify this every time, we take the word "represent" always to mean "represent with coprime integers x and y".

Primes. There is uniform agreement about what a prime number is when one is talking about positive integers, namely a number greater than 1 that is divisible only by itself and 1. For the sake of consistency we use the natural extension of this definition to other sorts of numbers considered in the last chapter of the book, namely Gaussian integers and their analogues in quadratic fields $\mathbb{Q}(\sqrt{d})$. Thus we call such numbers prime if the only way they factor is with one factor a unit (and they are not units or 0 themselves). Over the years it has become more usual to call numbers with this property irreducible rather than prime, using the term prime for numbers with the property that if they divide a product, then they must divide one of the factors. For example in the ring $\mathbb{Z}[\sqrt{-5}]$ the number 2 is prime according to our definition but not according to the standard definition since 2 divides $6 = (1+\sqrt{-5})(1-\sqrt{-5})$ but does not divide either factor $1 \pm \sqrt{-5}$.

We make a similar divergence from standard terminology when we define prime ideals in Chapter 8.

Topographs. Of much more recent origin is Conway's notion of the topograph of a quadratic form. Here we do not always follow Conway's picturesque terminology. What we call a separator line he called a river, and our source vertices and edges are his simple and double wells. He called a region with label 0 a lake but we call this just a 0 region.

Bibliography

The first group of eight books can be recommended for the quality of their exposition. They are arranged in approximate order of increasing sophistication and mathematical background required.

J. H. CONWAY and R. K. GUY, *The Book of Numbers*, Springer-Verlag, 1996.

- *A delightful collection of the wonders of numbers.*

J. H. CONWAY, *The Sensual Quadratic Form*, MAA, 1997.

- *Where topographs first appeared. Very enjoyable reading.*

H. DAVENPORT, *The Higher Arithmetic*, Cambridge U. Press, fifth ed. 1982 (orig. 1952).

- *A classical and accessible introduction to number theory.*

M. H. WEISSMAN, *An Illustrated Theory of Numbers*, AMS, 2017.

- *Many illuminating pictures, with chapters on topographs and quadratic forms.*

J. STILLWELL, *Numbers and Geometry*, Springer, 1998.

- *A pleasing intermingling of algebra and geometry.*

A. WEIL, *Number Theory: An Approach Through History*, Birkhäuser, 1984.

- *A scholarly historical study by one of the 20th century greats.*

J. H. SILVERMAN and J. TATE, *Rational Points on Elliptic Curves*, Springer-Verlag, 1992.

- *A natural next step after the present book.*

J.-P. SERRE, *A Course in Arithmetic*, Springer-Verlag, 1973 (French orig. 1970).

- *A master expositor writing at the graduate level, in spite of the title.*

Other books with coverage of quadratic forms, organized in the same way as the previous list.

D. E. FLATH, *Introduction to Number Theory*, Wiley, 1989. AMS Chelsea 2018.

H. COHN, *Advanced Number Theory*, Dover, 1980.

H. E. ROSE, *A Course in Number Theory*, Clarendon Press 1994.

J. L. LEHMAN, *Quadratic Number Theory*, AMS, 2019.

F. HALTER-KOCH, *Quadratic Irrationals*, CRC Press, 2013.

D. A. BUELL, *Binary Quadratic Forms*, Springer-Verlag, 1989.

D. A. COX, *Primes of the form $x^2 + ny^2$*, Wiley, 1989.

A few historical references, in chronological order.

C. F. GAUSS, *Disquisitiones Arithmeticae, English trans. Springer-Verlag, 1986 (Latin orig. 1801).*

- *The first book about quadratic forms.*

P. G. L. DIRICHLET, *Lectures on Number Theory, English trans. AMS, 1999 (German orig. 1863).*

- *With supplements by R. Dedekind.*

A. HURWITZ, *Über die Reduktion der binären quadratischen Formen, Math. Annalen 45 (1894), 85–117.*

- *This article (in German) is where the Farey diagram first appeared.*

Index