

Binary quadratic forms

1 Introduction

Chia's underlying VDF performs computations within class groups of binary quadratic forms. This handout provides an introduction to binary quadratic forms and to their properties relevant to application in the Chia VDF.

The study of class groups is typically presented either in the context of fractional ideals of algebraic number fields or in the context of binary quadratic forms. The ideal class group of an algebraic number field K is the quotient group J_K/P_K , where J_K is the group of fractional ideals of the ring of integers, O_K of K , and P_K is its subgroup of principal fractional ideals. While many feel that this provides a more conceptually intuitive introduction to class groups, the representation of elements of a class group and the implementation of the group operation are best carried out through binary quadratic forms, so this handout will limit its scope to form class groups.

The theorems here will be presented without proof. Please see the References section for proofs and for additional information about class groups and binary quadratic forms.

2 Background

Definition 2.1 (Binary quadratic form). A binary quadratic form is

$$f(x, y) = ax^2 + bxy + cy^2$$

where $a, b, c \in \mathbb{R}$ and a, b, c are not all equal to zero.

We write $f = (a, b, c)$ and call f a “form”. These are the objects that we'll be working with throughout this handout.

Definition 2.2 (Integral form). An integral form is a binary quadratic form where $a, b, c \in \mathbb{Z}$.

Integral binary quadratic forms are of key importance in algebraic number theory, and they are the relevant forms to the Chia VDF. The remainder of this handout will focus solely on integral forms.

Definition 2.3 (Content of a form). Denoted by $\text{cont}(f)$, the content of a form is

$$\text{cont}(f) = \gcd(a, b, c)$$

Definition 2.4 (Primitive form). A form f is called primitive if $\text{cont}(f) = 1$.

2.1 Discriminant

Definition 2.5 (Discriminant). The discriminant of a form f is $\Delta(f) = b^2 - 4ac$.

Note that because the square of an integer is always congruent to 0 or 1 mod 4, the discriminant of an integral binary quadratic form is always congruent to 0 or 1 mod 4.

Further, any integer which is congruent to 0 or 1 mod 4 is the discriminant of a binary quadratic form. This is easy to see by considering the form

$$\left(1, d \bmod 4, \frac{(d \bmod 4) - d}{4}\right)$$

For $d \equiv 0$ or $1 \bmod 4$, observe that this form will always be integral, and the discriminant Δ of the form is equal to d .

Theorem 2.1. *Given $\Delta \in \mathbb{Z}$, there is at least one integer binary quadratic form with discriminant Δ if and only if $\Delta \equiv 0$ or $1 \bmod 4$.*

One can easily check that if $f = (a, b, c)$ is an integral form, then b and $\Delta(f)$ always share the same parity, i.e. $b \equiv \Delta(f) \bmod 2$.

Definition 2.6 (Fundamental discriminant). For $\Delta \in \mathbb{Z}$, Δ is a fundamental discriminant if and only if:

1. $\Delta \equiv 1 \bmod 4$ and Δ is square-free¹, or
2. $\Delta \equiv 0 \bmod 4$, $\frac{\Delta}{4} \equiv 2, 3 \bmod 4$, and $\frac{\Delta}{4}$ is square-free

Connecting back to algebraic number fields, fundamental discriminants are exactly those values which are discriminants of quadratic fields.

Definition 2.7 (Positive definite binary quadratic form). A binary form is called positive definite if for any $(x, y) \in \mathbb{R}^2$ with $(x, y) \neq (0, 0)$, the value $f(x, y)$ is positive, i.e. $f(x, y) > 0$.

Definition 2.8 (Negative definite binary quadratic form). A binary form is called negative definite if for any $(x, y) \in \mathbb{R}^2$ with $(x, y) \neq (0, 0)$, the value $f(x, y)$ is negative, i.e. $f(x, y) < 0$.

Definition 2.9 (Positive semi-definite binary quadratic form). A binary form is called positive semi-definite if for any $(x, y) \in \mathbb{R}^2$ with $(x, y) \neq (0, 0)$, the value $f(x, y)$ is non-negative, i.e. $f(x, y) \geq 0$.

Definition 2.10 (Negative semi-definite binary quadratic form). A binary form is called negative semi-definite if for any $(x, y) \in \mathbb{R}^2$ with $(x, y) \neq (0, 0)$, the value $f(x, y)$ is non-positive, i.e. $f(x, y) \leq 0$.

¹A square-free integer is one which isn't divisible by any perfect squares, i.e. there are no repeated factors in its prime decomposition.

Definition 2.11 (Indefinite binary quadratic form). A binary form is called indefinite if for any $(x, y) \in \mathbb{R}^2$ with $(x, y) \neq (0, 0)$, the value $f(x, y)$ takes on both positive and negative values.

The forms relevant to the Chia VDF are positive definite forms.

Theorem 2.2. A form f is positive definite if and only if $\Delta(f) < 0$ and $a > 0$.

2.1.1 A quick detour to build intuition about discriminants

The quantity $b^2 - 4ac$ should look familiar. You probably first saw it as the content under the square root in the quadratic formula,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Recall that the quadratic formula helps us solve single-variable quadratic equations that are too difficult to factor, giving the two roots of the associated single-variable quadratic expression. If we set the expression equal to y and consider the represented curve, the roots are the points in the plane where the curve touches the x-axis.

In the quadratic formula, what happens when $b^2 - 4ac = 0$? The square root reduces to zero, and we're left with $x = \frac{-b}{2a}$. Hence, there is only one (repeated) root. This is the case in which the graph of the curve grazes the x-axis at just one point. Relative to binary quadratic forms, these discriminants correspond to the extreme case of the semi-definite binary quadratic forms mentioned above.²

When $b^2 - 4ac < 0$, however, the quantity under the square root is negative and the roots are complex. Therefore, the curve has no roots on the real x-axis and therefore never crosses it. This gives a curve which is entirely positive or entirely negative. These cases analogously correspond to the definite binary quadratic forms.

Finally, when $b^2 - 4ac > 0$, two real roots exist, and the curve crosses the x-axis at two points, leading to at least one positive and at least one negative value along the curve. These cases analogously correspond with indefinite binary quadratic forms.

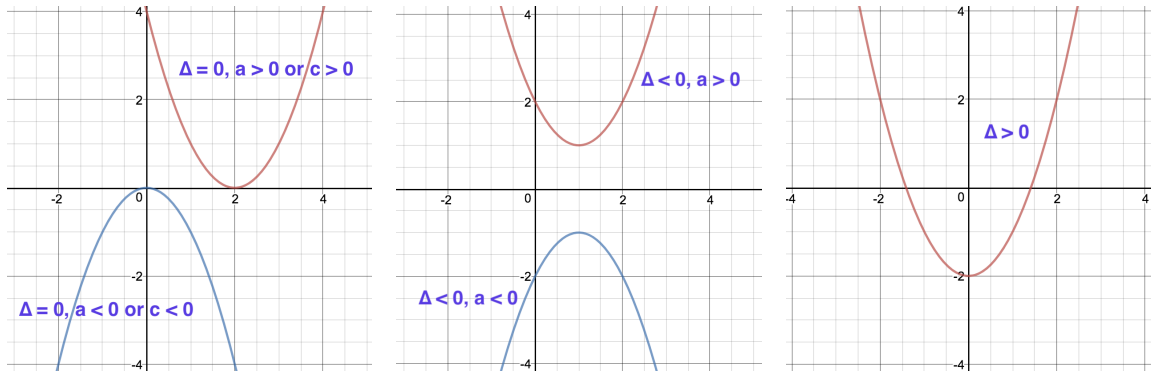


Figure 1: Analogous correspondences between binary quadratic forms (in purple) and single-variable quadratic curves (in blue and red).

²Note that the positive and negative semi-definite quadratic forms also include the positive and negative definite forms, respectively.

2.2 Matrix representations of forms

Definition 2.12 (Matrix of a form). The matrix of a form $f = (a, b, c)$ is

$$M(f) = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$$

with $\det(M(f)) = ac - \frac{b^2}{4}$.

Using its matrix form, and defining

$$X = \begin{bmatrix} x & y \end{bmatrix}$$

$f(x, y)$ can be written as

$$f(x, y) = X M(f) X^\top = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + bxy + cy^2$$

and we have

$$\Delta(f) = -4 \det(M(f)).$$

2.3 Representation of integers

Consider the equation

$$ax^2 + bxy + cy^2 = n$$

where $n \in \mathbb{Z}$ and $f = (a, b, c)$ is an integral binary quadratic form.

Definition 2.13 (A representation of n by f). A solution $(x, y) \in \mathbb{Z}^2$ to the above equation for a given f and n is called a representation of n by f .

Definition 2.14 (A proper representation of n by f). A solution $(x, y) \in \mathbb{Z}^2$ to the above equation for a given f and n , such that $\gcd(x, y) = 1$, is called a proper representation of n by f .

Notice that we can restrict our representation considerations to primitive forms. This is because if the content of f is $k \neq 1$, i.e. $\gcd(a, b, c) = k$ for some $k \neq 1$ such that, say, $a = kA$, $b = kB$, and $c = kC$, then we could factor out k from the lefthand side of the equation, which would mean that n , on the righthand side, is also divisible by k .

Theorem 2.3. *If $\Delta(f) < 0$ and if $n \in \mathbb{R}$, then the equation $ax^2 + bxy + cy^2 = n$ has only finitely many solutions, and x and y are bounded by*

$$x^2 \leq \frac{4cn}{|\Delta|} \quad \text{and} \quad y^2 \leq \frac{4an}{|\Delta|}$$

To find all of the solutions to such an equation where $\Delta(f) < 0$ and $n \in \mathbb{R}$, one could test for each solution (x, y) that obeys the constraints given above.

Example 2.1. *Solve the equation*

$$3x^2 + 2xy + 2y^2 = 28.$$

Notice that

$$\Delta(f) = b^2 - 4ac = 2^2 - 4(3)(2) = -20 < 0$$

Therefore

$$x^2 \leq \frac{4(2)(28)}{|-20|} = \frac{56}{5} = 11.2 \quad \text{and} \quad y^2 \leq \frac{4(3)(28)}{|-20|} = \frac{84}{5} = 16.8$$

And so

$$|x| \leq 3 \quad \text{and} \quad |y| \leq 4$$

Testing all pairs of (x, y) which satisfy those constraints, we find that the representations of 28 by $f = (3, 2, 2)$ are $(2, 2)$, $(-2, -2)$, $(-2, 4)$, and $(2, -4)$. \square

3 Equivalence

In this section, we'll learn what makes two forms equivalent. If two forms are equivalent then they represent the same integers, although the inverse is not necessarily true. There are two types of equivalence: wide equivalence and proper equivalence.

Definition 3.1 (Wide equivalence). Two forms $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = Ax^2 + Bxy + Cy^2$ are widely equivalent if there is an invertible change of variables

$$x' = rx + sy, \quad y' = tx + uy$$

with $r, s, t, u \in \mathbb{Z}$ and $ru - st = \pm 1$, such that

$$a(x')^2 + bx'y' + c(y')^2 = Ax^2 + Bxy + Cy^2$$

or, in other words, $g(x, y) = f(rx + sy, tx + uy)$.

Definition 3.2 (Proper equivalence). Two forms $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = Ax^2 + Bxy + Cy^2$ are properly equivalent if the wide equivalence conditions hold and also $ru - st = +1$.

Given the above definitions, a change of variables of a wide equivalence relation can be represented by a matrix³

$$T = \begin{bmatrix} r & s \\ t & u \end{bmatrix} \in GL(2, \mathbb{Z})$$

and a change of variables of a proper equivalence relation can be represented by the matrix

³See the Appendix for definitions of $GL(2, \mathbb{Z})$ and $SL(2, \mathbb{Z})$.

$$U = \begin{bmatrix} r & s \\ t & u \end{bmatrix} \in SL(2, \mathbb{Z})$$

In either case, we can write

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r & s \\ t & u \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

to describe the change of variables. The invertibility requirement of the change of variables is what imposes the constraint that $ru - st = \pm 1$ (in the case of wide equivalence, or $ru - st = +1$ in the case of proper equivalence) since only those 2×2 matrices whose determinant is equal to ± 1 are invertible.

We'll restrict our discussion to proper equivalence. What follows doesn't necessarily generalize to wide equivalence, and in particular to cases where the matrix is in $GL(2, \mathbb{Z}) \setminus SL(2, \mathbb{Z})$.

We can express form equivalence and the change of variables using the matrix representation of forms, i.e. $M(f) = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$.

Suppose that two forms $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = Ax^2 + Bxy + Cy^2$ are properly equivalent. Then we have

$$\begin{aligned} g(x, y) &= \begin{bmatrix} x & y \end{bmatrix} M(g) \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} r & t \\ s & u \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} r & s \\ t & u \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} U^\top M(f) U \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

Therefore, $M(g) = U^\top M(f) U$.

Further, observe that

$$f\left(\begin{bmatrix} r & s \\ t & u \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}\right) = f(rx + sy, tx + uy) = a(rx + sy)^2 + b(rx + sy)(tx + uy) + c(tx + uy)^2$$

For ease of notation, define $(fU)(x, y) = f\left(\begin{bmatrix} r & s \\ t & u \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}\right)$, and note that

$$M(fU) = U^\top M(f) U$$

Then we can state a second definition for proper equivalence as the following.

Definition 3.3 (Proper equivalence). Two forms f and g are properly equivalent if $g = fU$ for some $U \in SL(2, \mathbb{Z})$. The $SL(2, \mathbb{Z})$ -orbit of a form is called the proper equivalence class of that form.⁴

Notice that because $SL(2, \mathbb{Z})$ is a group (and associative under matrix multiplication), equivalence is transitive. To see this, consider three forms $f(x, y)$, $g(x, y)$, and $h(x, y)$, and let $g = fU$ and $h = gV$, where $U, V \in SL(2, \mathbb{Z})$. Then

$$M(h) = V^\top M(g) V = V^\top U^\top M(f) U V = W^\top M(f) W$$

where $W = UV \in SL(2, \mathbb{Z})$ and $W^\top = V^\top U^\top \in SL(2, \mathbb{Z})$. Therefore $h = fW$, where $W \in SL(2, \mathbb{Z})$.

4 The class group

We can now say a bit more about the connection between ideal class groups and form class groups.

Take a negative discriminant Δ and let $F(\Delta)$ be the set of primitive positive definite binary quadratic forms $ax^2 + bxy + cy^2$ whose discriminant is Δ . Define a proper equivalence relation (we'll call it \sim) on $F(\Delta)$ such that $f \sim g$ when $g = fU$ for $U \in SL(2, \mathbb{Z})$. This breaks up $F(\Delta)$ into a set of equivalence classes, $C(\Delta) = F(\Delta)/\sim$.

Now consider a quadratic extension field $K = \mathbb{Q}(\sqrt{d})$, where d is square-free. If $d \equiv 1 \pmod{4}$, then the discriminant of K is $\Delta = d$. Otherwise, the discriminant of K is $\Delta = 4d$. Let O_K be the ring of integers of K , and let J_K be the group of fractional ideals of the ring of integers O_K . P_K is the subgroup of J_K consisting of principal fractional ideals. We define the ideal class group of K as the quotient group $Cl(K) = J_K/P_K$.

The connection between the above two types of class groups — and therefore between binary quadratic forms and ideals in quadratic fields — involves bijections between special versions of each type of class group, the details of which depend on whether the discriminant in question is positive or negative. For a real quadratic field $K = \mathbb{Q}(\sqrt{d})$ with discriminant $\Delta_K > 0$, there is a bijection between the narrow ideal class group⁵ of K and the form class group of primitive integral binary quadratic forms of discriminant Δ_K . And for an imaginary quadratic field $K = \mathbb{Q}(\sqrt{d})$ with discriminant $\Delta_K < 0$, there is a bijection between the ideal class group of K and the form class group of primitive positive-definite integral binary quadratic forms of discriminant Δ_K .

5 Reduction

This section will introduce normal forms and reduced forms and give algorithms for normalization and reduction.

⁴See Appendix for discussion on orbits. This sentence basically means that the proper equivalence class of a form is the set of all forms given by computing $U^\top M(f) U$ for all of the $U \in SL(2, \mathbb{Z})$.

⁵The narrow ideal class group is such that instead of using P_K in our class group quotient group, we use P_K^+ , which is the group of totally positive principal fractional ideals of K .

5.1 Normal forms

Definition 5.1 (Normal form). A form $f = (a, b, c)$ is called normal if $-a < b \leq a$.

Definition 5.2 (Normalization operator). We define the normalization operator $\eta(f)$ as

$$\eta(f) = \eta(a, b, c) = (a, b + 2ra, ar^2 + br + c)$$

where $r = \lfloor \frac{a-b}{2a} \rfloor$.

5.1.1 Normalization algorithm

The normalization algorithm, given a form $f = (a, b, c)$, such that $\Delta < 0$ and $a > 0$ is:

1. Compute $r = \lfloor \frac{a-b}{2a} \rfloor$.
2. Compute $\eta(f)$ and update $f = \eta(f)$.
3. Return f , which is now normalized.

Example 5.1. Let's normalize the form $f = (11, 49, 55)$. Note that it is not normal because $b > a$.

First we compute r :

$$r = \lfloor \frac{a-b}{2a} \rfloor = \lfloor \frac{11-49}{2(11)} \rfloor = \lfloor -1.72 \rfloor = -2$$

Then we compute $\eta(f)$:

$$\eta(f) = \eta(a, b, c) = (a, b + 2ra, ar^2 + br + c) = (11, 49 + 2(-2)(11), 11(-2)^2 + 49(-2) + 55) = (11, 5, 1)$$

And so the normalization of $f = (11, 49, 55)$ is $f_{\text{norm}} = (11, 5, 1)$. \square

5.2 Reduced forms

Definition 5.3 (Reduced form). A positive definite form $f = (a, b, c)$ is called reduced if it is normal and $a \leq c$, and if $a = c$ then $b \geq 0$.

Reduction of forms is important to the Chia VDF because frequently reducing $f(a, b, c)$ as f is repeatedly squared keeps a , b , and c from growing too large.

For any given $\Delta < 0$, each proper equivalence class of binary quadratic forms of that discriminant contains a unique reduced representative. We can therefore know certain properties of a discriminant, such as its class number, by studying only the reduced forms in the class group of that discriminant.

For a reduced form and $\Delta < 0$, we have

$$\begin{aligned}
|\Delta| &= 4ac - |b|^2 & (\Delta < 0) \\
&\geq 4a(a) - a^2 & (-a < b \leq a, a \leq c) \\
&\geq 3a^2
\end{aligned}$$

and so

$$a \leq \sqrt{\frac{|\Delta|}{3}}$$

Therefore, for a given $\Delta < 0$, there are finitely many a , and consequently there are finitely many b and c .⁶ This means that negative discriminants have a finite number of reduced forms and therefore have a finite number of equivalence classes.

Theorem 5.1. *Each primitive positive definite form f is properly equivalent to a unique reduced form.*

Definition 5.4 (Principal form of discriminant Δ). Let Δ be a negative integer such that $\Delta \equiv 0, 1 \pmod{4}$, i.e. Δ is a negative discriminant. Let $k = \Delta \pmod{2}$. Then

$$f = \left(1, k, \frac{k^2 - \Delta}{4}\right)$$

is the unique reduced form $(1, b, c)$ of discriminant Δ . This particular reduced form is called the principal form of discriminant Δ .

Example 5.2. *Here are examples of principal forms of several discriminants.*

$$\begin{array}{ll}
\Delta = -3 & (1, 1, 1) \\
\Delta = -7 & (1, 1, 2) \\
\Delta = -11 & (1, 1, 3) \\
\Delta = -19 & (1, 1, 5)
\end{array}
\quad
\begin{array}{ll}
\Delta = -4 & (1, 0, 1) \\
\Delta = -8 & (1, 0, 2) \\
\Delta = -15 & (1, 1, 4) \\
\Delta = -20 & (1, 0, 5)
\end{array}
\quad \square$$

Definition 5.5 (Principal class of discriminant Δ). The equivalence class of the principal form of discriminant Δ is called the principal class of discriminant Δ .

The principal class of a discriminant Δ is the identity element of the form class group of Δ with respect to form composition.⁷

Example 5.3. *Let's consider the discriminant $\Delta = -23$. This discriminant has a total of three equivalence classes of positive definite binary quadratic forms. Therefore its class number is 3. In particular, the three unique reduced representatives of the equivalence classes are*

$$(1, 1, 6), \quad (2, -1, 3), \quad \text{and} \quad (2, 1, 3).$$

⁶Because $|b| \leq a$ and $c = \left(\frac{b^2 - \Delta}{4a}\right)$.

⁷Form composition is discussed in Section 6.

The principal form is $(1, 1, 6)$, and therefore this form (as a representative of the principal equivalence class) acts as the identity when composed with other forms with $\Delta = -23$. Every primitive positive definite binary quadratic form with $\Delta = -23$ is properly equivalent to one of the three forms above, and composition in the group boils down to composition between these three forms. \square

Definition 5.6 (Reduction operator). Given a form $f = (a, b, c)$, the reduction operator $\rho(f)$ is defined as

$$\rho(f) = \rho(a, b, c) = (c, -b + 2sc, cs^2 - bs + a)$$

where $s = \lfloor \frac{c+b}{2c} \rfloor$. Notice that $\rho(a, b, c)$ is equivalent to the normalization of $(c, -b, a)$.

5.2.1 Reduction algorithm

The reduction algorithm, given a form $f = (a, b, c)$ such that $\Delta < 0$ and $a > 0$ is:

1. Normalize f and update $f = f_{\text{norm}}$.
2. If f is reduced, return f ; if f is not reduced, compute s and $\rho(f)$ and update $f = \rho(f)$.
3. Repeat step 2 until a reduced form is produced.

Example 5.4. Let's reduce the form $f = (11, 49, 55)$. We first normalize it, which we did in Example 5.1, finding that $f_{\text{norm}} = (11, 5, 1)$. Solving Example 5.1 is equivalent to performing step 1 of the reduction algorithm.

But $f = f_{\text{norm}}$ is not yet reduced because $a > c$. We implement the reduction operator until we reach a reduced form.

Compute s :

$$s = \lfloor \frac{c+b}{2c} \rfloor = \lfloor \frac{1+5}{2(1)} \rfloor = \lfloor 3 \rfloor = 3$$

Compute $\rho(f)$:

$$\rho(f) = (c, -b + 2sc, cs^2 - bs + a) = (1, -5 + 2(3)(1), 1(3)^2 - 5(3) + 11) = (1, 1, 5)$$

This form is reduced and so $(1, 1, 5)$ is the reduced form of $(11, 49, 55)$. Notice that, of course, the discriminant of all three forms is -19 . \square

6 Composition

Originally developed by Gauss, composition of binary quadratic forms is a commutative operation on the class group. Others, such as Arndt, Dirichlet, and Bhargava, have contributed to the topic, and various algorithms have been developed for its implementation. In this section I motivate and outline the details of composition, and then I lay out the basic structure of the algorithm.

6.1 Explaining composition

To begin, note that we only perform composition between forms which have the same discriminant. This should make sense because we're operating within the class group of a particular discriminant, and although the objects we're operating on are integer 3-tuples, (a, b, c) , which represent binary quadratic forms, the more abstract objects that we're operating on are the equivalence classes of the class group. Analogously, when we add the rational numbers $\frac{20}{12}$ and $\frac{5}{35}$, what we're fundamentally working with are the fraction equivalence classes whose "reduced representatives" are $\frac{5}{3}$ and $\frac{1}{7}$.

To understand the basic idea behind composition, consider two binary quadratic forms

$$f_1 = ax_1^2 + bx_1y_1 + cy_1^2 \quad \text{and} \quad f_2 = \alpha x_2^2 + \beta x_2y_2 + \gamma y_2^2$$

Note that we consider (x_1, y_1) and (x_2, y_2) as independent sets of variables.

We then would like to find a form f_3 such that

$$f_1 f_2 = f_3$$

Multiplying f_1 and f_2 , we find that

$$\begin{aligned} f_1 f_2 = & a\alpha x_1^2 x_2^2 + a\beta x_1^2 x_2 y_2 + a\gamma x_1^2 y_2^2 + b\alpha x_1 y_1 x_2^2 + b\beta x_1 y_1 x_2 y_2 \\ & + b\gamma x_1 y_1 y_2^2 + c\alpha y_1^2 x_2^2 + c\beta y_1^2 x_2 y_2 + c\gamma y_1^2 y_2^2 \end{aligned}$$

We want the righthand side of the equation to resemble a binary quadratic form so that

$$f_1 f_2 = AX^2 + BXY + CY^2 = f_3$$

for some $A, B, C \in \mathbb{Z}$.

We'll do this by using a change of variables such that X and Y are linear combinations of $x_1 x_2$, $x_1 y_2$, $y_1 x_2$, and $y_1 y_2$:

$$X = jx_1x_2 + kx_1y_2 + ly_1x_2 + my_1y_2, \quad Y = rx_1x_2 + sx_1y_2 + ty_1x_2 + uy_1y_2$$

for some $j, k, l, m, r, s, t, u \in \mathbb{Z}$. The goal of our algorithm is to find A, B , and C by finding appropriate integer values for j, k, l, m, r, s, t, u .

We find j, k, l, m, r, s, t, u as follows. Define $g = \frac{1}{2}(b + \beta)$ and $h = -\frac{1}{2}(b - \beta)$. Define $w = \gcd(a, \alpha, g)$. Then define the matrix

$$M = \begin{pmatrix} j & k & l & m \\ r & s & t & u \end{pmatrix}$$

with submatrices

$$\begin{aligned} M_1 &= \begin{pmatrix} j & k \\ r & s \end{pmatrix}, \quad M_2 = \begin{pmatrix} j & l \\ r & t \end{pmatrix}, \quad M_3 = \begin{pmatrix} j & m \\ r & u \end{pmatrix}, \\ M_4 &= \begin{pmatrix} k & l \\ s & t \end{pmatrix}, \quad M_5 = \begin{pmatrix} k & m \\ s & u \end{pmatrix}, \quad M_6 = \begin{pmatrix} l & m \\ t & u \end{pmatrix} \end{aligned}$$

Set the values j, r, s, t , and u to be

$$j = w, \quad r = 0, \quad s = \frac{a}{w}, \quad t = \frac{\alpha}{w}, \quad u = \frac{g}{w}$$

and find k , l , and m using the following set of conditions:

1. $\det(M_4) = kt - ls = h$
2. $\det(M_5) = ku - ms = \gamma$
3. $\det(M_6) = lu - mt = c$

To solve this set of equations, we'll create an augmented matrix and row reduce.

$$\begin{aligned} & \left(\begin{array}{ccc|c} t & -s & 0 & h \\ u & 0 & -s & \gamma \\ 0 & u & -t & c \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \\ \leftarrow \end{array} \Rightarrow \left(\begin{array}{ccc|c} t & -s & 0 & h \\ 0 & u & -t & c \\ u & 0 & -s & \gamma \end{array} \right) \begin{array}{l} | \cdot 1/t \\ | \cdot 1/u \\ \end{array} \\ \Rightarrow & \left(\begin{array}{ccc|c} 1 & -s/t & 0 & h/t \\ 0 & 1 & -t/u & c/u \\ u & 0 & -s & \gamma \end{array} \right) \begin{array}{l} \leftarrow^+ \\ \leftarrow_{s/t} \\ \leftarrow^+ \end{array} \Rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -s/u & (hu+cs)/tu \\ 0 & 1 & -t/u & c/u \\ u & 0 & -s & \gamma \end{array} \right) \begin{array}{l} \leftarrow^{-u} \\ \leftarrow^+ \\ \leftarrow^+ \end{array} \\ \Rightarrow & \left(\begin{array}{ccc|c} 1 & 0 & -s/u & (hu+cs)/tu \\ 0 & 1 & -t/u & c/u \\ 0 & 0 & 0 & \gamma - (hu+cs)/t \end{array} \right) \end{aligned}$$

In order for this matrix to have solutions, it must be the case that $\gamma - (hu+cs)/t = 0$.⁸ Let's check to see if $\gamma - \frac{hu+cs}{t} \stackrel{?}{=} 0$:

$$\begin{aligned} \gamma - \frac{hu+cs}{t} &= \gamma - \frac{h\frac{g}{w} + c\frac{a}{w}}{\frac{\alpha}{w}} \\ &= \gamma - \frac{gh+ac}{\alpha} \\ &= \gamma - \frac{\left(\frac{1}{2}(b+\beta)\right)\left(-\frac{1}{2}(b-\beta)\right) + ac}{\alpha} \\ &= \gamma - \frac{\left(-\frac{1}{4}b^2 + \frac{1}{4}\beta^2\right) + ac}{\alpha} \\ &= \frac{1}{\alpha} \left(\alpha\gamma + \frac{1}{4}b^2 - \frac{1}{4}\beta^2 - ac \right) \\ &= \frac{1}{4\alpha} (4\alpha\gamma + b^2 - \beta^2 - 4ac) \\ &= \frac{1}{4\alpha} ((b^2 - 4ac) - (\beta^2 - 4\alpha\gamma)) \\ &= \frac{1}{4\alpha} (0) \quad [\text{because } \Delta(f_1) = \Delta(f_2)] \\ &= 0 \quad \checkmark \end{aligned}$$

Therefore, our row-reduced matrix is

$$\left(\begin{array}{ccc|c} 1 & 0 & -s/u & (hu+cs)/tu \\ 0 & 1 & -t/u & c/u \\ 0 & 0 & 0 & 0 \end{array} \right)$$

⁸Note that because the reduced binary quadratic forms relevant to Chia are positive definite forms, $a > 0$, and so $t \neq 0$.

This matrix corresponds to the following system of equations:

1. $k - \frac{s}{u}m = \frac{hu+cs}{tu}$
2. $l - \frac{t}{u}m = \frac{c}{u}$
3. $0 = 0$

We can parameterize the system, with ξ as the parameter, and we find that the system has infinite solutions of the following form:

$$k = \frac{s}{u}\xi + \frac{hu+cs}{tu}, \quad l = \frac{t}{u}\xi + \frac{c}{u}, \quad \text{and} \quad m = \xi$$

We need to choose a form of ξ such that k , l , and m are integers.

Putting ξ in terms of k , we have

$$\xi = \frac{tuk-hu-cs}{st}$$

To ensure that the fraction on the righthand side is an integer, we solve the equation

$$(tu)k = hu + cs \pmod{st}$$

which produces a set of solutions of the form $k = \mu + \nu n$, where $\mu, \nu \in \mathbb{Z}$ and n ranges over all of \mathbb{Z} . Any choice of n will yield integer values for k and m , but some n values may produce a non-integer value for l , so we now find an appropriate n value that guarantees $l \in \mathbb{Z}$. Putting l in terms of n , we have

$$l = \frac{t}{u}\xi + \frac{c}{u} = \frac{tk-h}{s} = \frac{t(\mu+\nu n)-h}{s} = \frac{(t\nu)n+(t\mu-h)}{s}$$

We want the fraction on the righthand side to equal an integer, so we solve the equation

$$(t\nu)n = h - t\mu \pmod{s}$$

which gives a set of solutions of the form $n = \lambda + \sigma n'$, where $\lambda, \sigma \in \mathbb{Z}$ and n' ranges over all of \mathbb{Z} . We choose $n' = 0$ and let $n = \lambda$.

We can now find k , l , and m by

$$k = \mu + \nu\lambda \quad l = \frac{kt-h}{s} \quad m = \frac{tuk-hu-cs}{st}$$

and these values are guaranteed to be integers. Notice that because $s = \frac{a}{w}$ and $t = \frac{\alpha}{w}$, and because a and α are never equal to 0, it will always be the case that $s, t \neq 0$, and so k , l , and m will always be defined.

Finally, $f_3 = AX^2 + BXY + CY^2$ is given by⁹

$$A = st - ru, \quad B = (ju + mr) - (kt + ls), \quad C = kl - jm$$

⁹Although $r = 0$, I'm presenting these formulae with the r terms to gesture to their inherent symmetry.

To verify that f_3 indeed equals f_1f_2 , first expand f_3 using $X = jx_1x_2 + kx_1y_2 + ly_1x_2 + my_1y_2$ and $Y = rx_1x_2 + sx_1y_2 + ty_1x_2 + uy_1y_2$. After grouping together common terms, we end up with:

$$\begin{aligned} f_3 = & (Aj^2 + Bjr + Cr^2)x_1^2x_2^2 + (2Ajk + B(js + kr) + 2Crs)x_1^2x_2y_2 \\ & + (Ak^2 + Bks + Cs^2)x_1^2y_2^2 + (2Ajl + B(jt + lr) + 2Crt)x_1y_1x_2^2 \\ & + (2A(jm + kl) + B(ju + kt + ls + mr) + 2C(ru + st))x_1y_1x_2y_2 \\ & + (2Akm + B(ku + ms) + 2Csu)x_1y_1y_2^2 + (Al^2 + Blt + Ct^2)y_1^2x_2^2 \\ & + (2Alm + B(lu + mt) + 2Ctu)y_1^2x_2y_2 + (Am^2 + Bmu + Cu^2)y_1^2y_2^2 \end{aligned}$$

We then wish to check that the above expansion of f_3 is equal to

$$\begin{aligned} f_1f_2 = & a\alpha x_1^2x_2^2 + a\beta x_1^2x_2y_2 + a\gamma x_1^2y_2^2 + b\alpha x_1y_1x_2^2 + b\beta x_1y_1x_2y_2 \\ & + b\gamma x_1y_1y_2^2 + c\alpha y_1^2x_2^2 + c\beta y_1^2x_2y_2 + c\gamma y_1^2y_2^2 \end{aligned}$$

The reader may check for coefficient equality for all terms, but here I will confirm coefficient equality for two of the terms:

1. $(Aj^2 + Bjr + Cr^2)x_1^2x_2^2 \stackrel{?}{=} a\alpha x_1^2x_2^2 \Rightarrow Aj^2 + Bjr + Cr^2 \stackrel{?}{=} a\alpha$
2. $(2Alm + B(lu + mt) + 2Ctu)y_1^2x_2y_2 \stackrel{?}{=} c\beta y_1^2x_2y_2 \Rightarrow 2Alm + B(lu + mt) + 2Ctu \stackrel{?}{=} c\beta$

We'll draw from the definitions for $A, B, C, g, h, w, j, r, s, t$, and u , and from the identities defined by the three conditions on k, l , and m .

To check the first equality, $Aj^2 + Bjr + Cr^2 \stackrel{?}{=} a\alpha$:

$$\begin{aligned} Aj^2 + Bjr + Cr^2 &= (st - ru)j^2 + ((ju + mr) - (kt + ls))jr + (kl - jm)r^2 \\ &= (st - 0)j^2 + ((ju + mr) - (kt + ls))0 + (kl - jm)0 \\ &= stj^2 \\ &= \frac{a}{w} \frac{\alpha}{w} w^2 \\ &= a\alpha \quad \checkmark \end{aligned}$$

To check the second equality, $2Alm + B(lu + mt) + 2Ctu \stackrel{?}{=} c\beta$:

$$\begin{aligned}
2Alm + B(lu + mt) + 2Ctu &= 2(st - ru)lm + ((ju + mr) - (kt + ls))(lu + mt) + 2(kl - jm)tu \\
&= 2lmst - 2lmru + jlu^2 + jmtu + lmru + m^2rt - kltu - kmt^2 \\
&\quad - l^2su - lmst + 2kltu - 2jmtu \\
&= lmst - lmru + jlu^2 + lmru + m^2rt - kmt^2 - l^2su + kltu - jmtu \\
&= lmst - 0 + jlu^2 + 0 + 0 - kmt^2 - l^2su + kltu - jmtu \\
&= lmst + jlu^2 - kmt^2 - l^2su + kltu - jmtu \\
&= (kltu - kmt^2 - l^2su + lmst) + jlu^2 - jmtu \\
&= (kt - ls)(lu - mt) + jlu^2 - jmtu \\
&= (kt - ls)(lu - mt) + ju(lu - mt) \\
&= (kt - ls)c + juc \\
&= hc + juc \\
&= hc + w\frac{g}{w}c \\
&= c(h + g) \\
&= c\left(\frac{1}{2}(b + \beta) - \frac{1}{2}(b - \beta)\right) \\
&= c\beta \quad \checkmark
\end{aligned}$$

The reader may similarly verify the equality of the other coefficients, and therefore we confirm that $f_3 = f_1f_2$.

There are several notes to make about this construction.

The first, and most important, is that the system of equations defining k, l , and m has infinite solutions. However, all of the solutions are in the same equivalence class, and they will each reduce to the same binary quadratic form.

Second, we have the observation that if we wish to reconstruct f_1 and f_2 from $g, h, w, j, k, l, m, r, s, t$, and u , we simply note that $\det(M_1)$ and $\det(M_2)$ recover a and α respectively, $\det(M_3) - \det(M_4)$ and $\det(M_3) + \det(M_4)$ recover b and β respectively, and $\det(M_6)$ and $\det(M_5)$ recover c and γ respectively.¹⁰ So f_1 and f_2 can be represented using the determinants of the six submatrices of M :

$$\begin{aligned}
f_1 &= ax_1^2 + bx_1y_1 + cy_1^2 \\
&= (js - kr)x_1^2 + ((ju - mr) - (kt - ls))x_1y_1 + (lu - mt)y_1^2 \\
&= \det(M_1)x_1^2 + (\det(M_3) - \det(M_4))x_1y_1 + \det(M_6)y_1^2 \\
f_2 &= \alpha x_2^2 + \beta x_2y_2 + \gamma y_2^2 \\
&= (jt - lr)x_2^2 + ((ju - mr) + (kt - ls))x_2y_2 + (ku - ms)y_2^2 \\
&= \det(M_2)x_2^2 + (\det(M_3) + \det(M_4))x_2y_2 + \det(M_5)y_2^2
\end{aligned}$$

¹⁰Note that recovering b, β, c , and γ uses the three conditions on k, l , and m

6.1.1 The composition algorithm

Given $f_1 = (a, b, c)$ and $f_2 = (\alpha, \beta, \gamma)$, the composition algorithm for $f_1 f_2 = f_3$ is as follows:

- Set $g = \frac{1}{2}(b + \beta)$, $h = -\frac{1}{2}(b - \beta)$, and $w = \gcd(a, \alpha, g)$.
- Set $j = w$, $r = 0$, $s = \frac{a}{w}$, $t = \frac{\alpha}{w}$, $u = \frac{g}{w}$.
- Solve $(tu)k = hu + sc \pmod{st}$, the solutions to which have the form $k = \mu + \nu n$ for all $n \in \mathbb{Z}$. Store μ and ν .
- Solve $(t\nu)n = h - t\mu \pmod{s}$, the solutions to which have the form $n = \lambda + \sigma n'$ for all $n' \in \mathbb{Z}$. Store λ .
- Set $k = \mu + \nu\lambda$, $l = \frac{kt-h}{s}$, and $m = \frac{tuk-hu-cs}{st}$.
- Set $A = st - ru$, $B = (ju + mr) - (kt + ls)$, and $C = kl - jm$. Then $f_3 = (A, B, C)$.

6.2 Composition and representation of integers

Consider a composition

$$f_1 f_2 = f_3$$

where

$$f_1 = ax_1^2 + bx_1y_1 + cy_1^2, \quad f_2 = \alpha x_2^2 + \beta x_2y_2 + \gamma y_2^2, \quad \text{and} \quad f_3 = AX^2 + BXY + CY^2$$

In Section 2.3, we saw that a binary quadratic form represents certain integers, and a representation of an integer n by a form f is a solution $(x, y) \in \mathbb{Z}^2$ to the equation

$$ax^2 + bxy + cy^2 = n$$

Recall that if two forms are equivalent then they represent the same integers. If $f_{3,nr}$ is the non-reduced form found by composing f_1 and f_2 using the algorithm in section 6.1.1, and if, for a given (x_1, y_1) and (x_2, y_2) , $f_1(x_1, y_1) = n_1$ and $f_2(x_2, y_2) = n_2$, and if X, Y are defined as they were in section 6.1, then $f_{3,nr}(X, Y) = n_1 n_2$. Note that this is not necessarily true for the reduced form $f_{3,r}$, which is found by submitting $f_{3,nr}$ to the reduction algorithm. However, $n_1 n_2$ will be representable by $f_{3,r}$ under some different set of integers (X', Y') because $f_{3,nr}$ and $f_{3,r}$ are in the same equivalence class and therefore represent the same set of integers.

7 Appendix

7.1 Special matrices

$GL(2, \mathbb{Z})$ is the set of all invertible 2×2 matrices with integer entries whose determinants are equal to ± 1 .

$SL(2, \mathbb{Z})$ is a subgroup of $GL(2, \mathbb{Z})$ defined as the set of all 2×2 matrices with integer entries whose determinants are equal to 1. The two generating matrices of $SL(2, \mathbb{Z})$ are

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

7.2 Orbits

Let S be a non-empty set and let G be a group. A *left action* of G on S is a mapping

$$G \times S \rightarrow S, \quad (g, s) \mapsto gs \in S$$

which has the following properties:

- $1_G s = s$ for all $s \in S$
- for $g, h \in G$ and $s \in S$, $g(hs) = (gh)s$

Given a left action of G on S , two elements s and t in S are equivalent if there's an element $g \in G$ such that $t = gs$. This is an equivalence relation on S , and the equivalence class $\{gs : g \in G\}$ is called the G -orbit of s .

We have an analogous definition for a *right action* of a group G on a set S . Given a right action of G on S we define the G -orbit of s in the same way as with left actions.

Orbits should bring to mind ring ideals, gesturing toward the deep connection between ideal class groups and form class groups.

References

Buchmann, J. and Vollmer, U. (2007). *Binary quadratic forms: An algorithmic approach*. (Vol. 20, Algorithms and computation in mathematics). Berlin: Springer-Verlag.

Cohen, H. (1993). *A course in computational algebraic number theory*. (Vol. 138, Graduate texts in mathematics). Berlin: Springer-Verlag.

Granville, A. (2014). Composing quadratic forms: Gauss, Dirichlet and Bhargava. *Pi in the Sky*, **18**, 3-6.

Jacobson, M. J. and van der Poorten, A. J. (2002). Computational aspects of NUCOMP. *Lecture notes in computer science, Algorithmic number theory - ANTS-V*, vol. 2369, 120-133. doi:10.1007/3-540-45455-1_10

Jenkins, E. D. (1935). On the composition of quadratic forms. *Bulletin of the American Mathematical Society*, vol. 41, number 10, 719-726.