MATHEMATICS OF COMPUTATION
Volume 72, Number 244, Pages 1935–1946
S 0025-5718(03)01518-7
Article electronically published on April 29, 2003

A NOTE ON NUCOMP

ALFRED J. VAN DER POORTEN

Abstract. This note is a detailed explanation of Shanks–Atkin NUCOMP—composition and reduction carried out "simultaneously"—for all quadratic fields, that is, including real quadratic fields. That explanation incidentally deals with various "exercises" left for confirmation by the reader in standard texts. Extensive testing in both the numerical and function field cases by Michael J Jacobson, Jr, reported elsewhere, confirms that NUCOMP as here described is in fact efficient for composition both of indefinite and of definite forms once the parameters are large enough to compensate for NUCOMP's extra overhead. In the numerical indefinite case that efficiency is a near doubling in speed already exhibited for discriminants as small as 10^7.

1. Introduction

This note comprises an introduction to and gloss on the notorious Banff talks [10] of Daniel Shanks on "Gauss and Composition". In those lectures, Shanks makes it plain that he has definite forms in mind. The same holds for the implementation [1] subsequently proposed by Oliver Atkin and detailed by Henri Cohen [3], §5.4.2. However, the remarks were intended in general, and of course they do hold in general. I notice that implementations intended for the definite case are equally appropriate in the indefinite case, and emphasise the indefinite case in the latter part of my remarks. An upshot is a detailed explanation of Shanks–Atkin NUCOMP—composition and reduction carried out "simultaneously"—for all quadratic fields. This explanation incidentally deals with various "exercises" left for confirmation by the reader in [3], §5.

Extensive testing in both the numerical and function field cases by Michael J Jacobson, Jr, reported elsewhere [7], confirms that NUCOMP as here described is in fact efficient for composition both of indefinite and of definite forms once the parameters are large enough to compensate for NUCOMP's extra overhead. In the numerical indefinite case that efficiency is a near doubling in speed already exhibited for discriminants as small as 10^7.

2. Composition

As usual, denote the quadratic form $uX^2 + vXY + wY^2$ by (u, v, w). Daniel Shanks points out in his notorious lecture [9] that a computationally efficient rule for

Received by the editor January 10, 2002.
2000 *Mathematics Subject Classification.* Primary 11Y40, 11E16, 11R11.
Key words and phrases. Binary quadratic form, composition.
The author was supported in part by a grant from the Australian Research Council.

composing forms $\varphi_1 = (u_1, v_1, w_1)$ and $\varphi_2 = (u_2, v_2, w_2)$ of the same discriminant D is provided by way of the "magic" matrix

$$M = \begin{bmatrix} A_x & B_x & C_x & D_x \\ A_y & B_y & C_y & D_y \end{bmatrix} \tag{1}$$

given by

$$G = \gcd\left(u_1, u_2, \tfrac{1}{2}(v_1 + v_2)\right), \tag{2}$$

$$M(\varphi_1, \varphi_2) = \begin{bmatrix} A_x = G & B_x & C_x & D_x \\ A_y = 0 & B_y = u_1/G & C_y = u_2/G & D_y = s/G \end{bmatrix},$$

and the conditions

$$\Delta_{B,C} = B_x C_y - B_y C_x = m, \tag{3a}$$

$$\Delta_{C,D} = C_x D_y - C_y D_x = w_1, \tag{3b}$$

$$\Delta_{B,D} = B_x D_y - B_y D_x = w_2\,. \tag{3c}$$

For brevity, we write $s = \frac{1}{2}(v_1 + v_2)$, and $m = -\frac{1}{2}(v_1 - v_2)$. Here, and in the sequel, roman letters denote rational integers; the "M" above is an upper case μ.

3. The Plücker relation

It is not obvious that the three equations (3) can be solved for the *three* integer unknowns B_x, C_x and D_x. However, a 2×4 array (1) naturally presents itself as the linear subspace of three-dimensional projective space $\mathbb{P}^3$ defined by the two planes given by the rows, dually—in this example, self-dually—as the line defined by the two points given by those rows. The six *Grassmann coordinates* defining the linear subspace are precisely the determinants $\Delta_{A,B} = A_x B_y - A_y B_x = u_1$, $\Delta_{A,C} = A_x C_y - A_y C_x = u_2$, $\Delta_{A,D} = A_x D_y - A_y D_x = \frac{1}{2}(v_1 + v_2) = s$, and the three determinants of (3). More generally, an m by n array, with $m \le n$, defines a linear subspace of $\mathbb{P}^{n-1}$ via its $\binom{n}{m}$ maximal minors.

Now, if two linear subspaces of $\mathbb{P}^{n-1}$ have the same coordinates, then they coincide. But it is not the case that every vector of $\binom{n}{m}$ coordinates defines a linear subspace. Those coordinates must satisfy certain *Plücker relations*, to wit, those generated by certain quadratic relations on the determinants.

A congenial reminder of those matters can be found in the relatively elementary technical source [6]. For the basics, see Klein's remarks [4]; one might even risk Grassmann's work directly [5]. Whatever, one can learn in particular that the six determinants of our example (more to the point: of any 2×4 example) are not independent. There is the one primitive Plücker relation

$$0 = \Delta_{A,B}\Delta_{C,D} - \Delta_{A,C}\Delta_{B,D} + \Delta_{A,D}\Delta_{B,C}\,. \tag{4}$$

Happily, it asserts precisely that the two forms φ_1 and φ_2 must have the same discriminant. They do, so a solution to two of the equations (3) must satisfy the third equation.

4. The composite form

Theorem 1. *The form* $\varphi_3 = (u_3, v_3, w_3)$ *obtained from* $M(\varphi_1, \varphi_2)$ *by*

$$\begin{aligned} u_3 &= B_y C_y - A_y D_y, \\ v_3 &= (A_x D_y + A_y D_x) - (B_x C_y + B_y C_x), \\ w_3 &= B_x C_x - A_x D_x \end{aligned}$$

is a compound of φ_1 *and* φ_2.

Proof. For $i = 1$, 2, and 3, set $\varphi_i(X_i, Y_i) = u_i X_i^2 + v_i X_i Y_i + w_i Y_i^2$. To see that φ_3 *is* a compound of the forms φ_1 and φ_2, it suffices to return to the first principles definition of composition. First, note that

$$\begin{aligned} \varphi_1(X_1, Y_1) &= \Delta_{A,B} X_1^2 + (\Delta_{A,D} - \Delta_{B,C}) X_1 Y_1 + \Delta_{C,D} Y_1^2, \\ \varphi_2(X_2, Y_2) &= \Delta_{A,C} X_2^2 + (\Delta_{A,D} + \Delta_{B,C}) X_2 Y_2 + \Delta_{B,D} Y_2^2 \,. \end{aligned} \tag{5}$$

Next, set

$$\begin{aligned} X_3 &= A_x X_1 X_2 + B_x X_1 Y_2 + C_x Y_1 X_2 + D_x Y_1 Y_2, \\ Y_3 &= A_y X_1 X_2 + B_y X_1 Y_2 + C_y Y_1 X_2 + D_y Y_1 Y_2 \,. \end{aligned} \tag{6}$$

Finally, by brute force, verify that, indeed,

$$\varphi_3(X_3, Y_3) = \varphi_1(X_1, Y_1)\varphi_2(X_2, Y_2). \tag{7}$$

□

To see why these facts are so, and to obtain a significantly less brutal argument, notice first that (6) is

$$\begin{aligned} X_3 &= (A_x X_2 + B_x Y_2) X_1 + (C_x X_2 + D_x Y_2) Y_1, \\ Y_3 &= (A_y X_2 + B_y Y_2) X_1 + (C_y X_2 + D_y Y_2) Y_1, \end{aligned} \tag{8}$$

or, just so,

$$\begin{aligned} X_3 &= (A_x X_1 + C_x Y_1) X_2 + (B_x X_1 + D_x Y_1) Y_2, \\ Y_3 &= (A_y X_1 + C_y Y_1) X_2 + (B_y X_1 + D_y Y_1) Y_2 \,. \end{aligned} \tag{9}$$

Thus the definition $\varphi_3(X_3, Y_3) = \varphi_1(X_1, Y_1)\varphi_2(X_2, Y_2)$, viewed as an identity in the variables X_1 and Y_1, promptly reveals on comparing discriminants that

$$\varphi_2(X_2, Y_2) = \begin{vmatrix} A_x X_2 + B_x Y_2 & C_x X_2 + D_x Y_2 \\ A_y X_2 + B_y Y_2 & C_y X_2 + D_y Y_2 \end{vmatrix}$$

and, on similarly emphasising the other pair of variables,

$$\varphi_1(X_1, Y_1) = \begin{vmatrix} A_x X_1 + C_x Y_1 & A_y X_1 + C_y Y_1 \\ B_x X_1 + D_x Y_1 & B_y X_1 + D_y Y_1 \end{vmatrix}.$$

These are the remark (5).

I am indebted to Renate Scheidler for being led to notice that one can now fairly readily "discover" φ_3 from

$$\begin{aligned} &\varphi_3(x, y)\varphi_2(x', y') \\ &\quad = \varphi_1\left((C_y x' + D_y y')x - (C_x x' + D_x y')y, -(A_y x' + B_y y')x + (A_x x' + B_x y')y\right), \end{aligned}$$

carefully evaluated as a determinant. For example, its coefficient of $x^2 x'^2$ is

$$\begin{vmatrix} A_x C_y - C_x A_y & A_y C_y - A_y C_y \\ B_x C_y - A_y D_x & B_y C_y - A_y D_y \end{vmatrix} = (B_y C_y - A_y D_y)\Delta_{A,C} \,.$$

5. Computing the magic matrix

To see that the entries of $M(\varphi_1, \varphi_2)$ can be chosen to be integers, consider the choice

$$B_x = \tfrac{1}{2}G(v_2 - v_3)/u_2\,, \quad C_x = \tfrac{1}{2}G(v_1 - v_3)/u_1\,, \quad D_x = G(B_xC_x - w_3)\,,$$

noticing, because $u_3 = u_1u_2/G^2$, and all three forms have discriminant D, that this *is* a solution to (3). However, it is more constructive to compute the quantities.

In brief, we use $\Delta_{B,C} = m$ and, if necessary, also $\Delta_{B,D} = w_2$. On the way, we explain the choice (2) of $A_x = G$. Namely, (a) use Euclid's extended algorithm to compute (b, c, F) such that $bu_2 + cu_1 = F = \gcd(u_1, u_2)$. It is mostly the case that $F = 1$; in any case, if $F \mid s$ set $A_x \leftarrow G = F$, $B_x \leftarrow mb$.

However, if $F \nmid s$, (b) use Euclid's extended algorithm again, to compute (x, y, G) such that $xF + ys = G = \gcd(F, s)$. This is the interesting case, cf. [2], where $F/G = H > 1$, so that $\Delta_{B,C} = m$ only yields $B_x \equiv bm/H \pmod{u_1/GH}$, rather than modulo u_1/G. However, we do have the Plücker relation (4), confirming that

$$\tfrac{1}{4}(v_1 + v_2)(v_1 - v_2) = -sm = u_1w_1 - u_2w_2.$$

Furthermore, $GH = bu_2 + cu_1$ and $G = xGH + ys$ remind us that

$$1 = bu_2/F + cu_1/F \quad \text{and} \quad 1 \equiv ys/G \pmod{H}\,.$$

We determine $B_x \pmod{B_y = u_1/G}$ by also using (3c):

$$w_2 = \Delta_{B,D} \equiv B_xs/G \pmod{u_1/G}\,,$$

and noticing that

$$-s \cdot bm/F = w_1bu_1/F - w_2bu_2/F = -w_2 + (bw_1 + cw_2) \cdot u_1/F\,.$$

Accordingly, if $H = F/G \neq 1$, (c) compute $l \leftarrow y(bw_1 + cw_2) \pmod{H}$. Set $A_x \leftarrow G = F/H$, $B_x \leftarrow bm/H + l \cdot u_1/F$, and note that one can now readily backtrack also to obtain C_x and D_x as integers.

6. Notes

6.1. The matrix $M(\varphi_1, \varphi_2)$ is just one of an equivalence class

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} A_x & B_x & C_x & D_x \\ A_y & B_y & C_y & D_y \end{bmatrix}, \quad \text{where } a,\ b,\ c \text{ and } d \in \mathbb{Z} \text{ satisfy } ad - bc = 1,$$

of 2×4 matrices with Grassmann coordinates defined by the pair of forms (5); the corresponding compounds are $\varphi_3(dX - bY, -cX + aY)$, where φ_3 is detailed in Theorem 1.

6.2. The choice of M recommended by (Gauss and) Shanks, and adopted here, is not arbitrary. Indeed, $M(\varphi_1, \varphi_2)$ "solves" the product

$$\begin{aligned} &G(X_1 - \alpha_1Y_1)(X_2 - \alpha_2Y_2) \\ &\qquad = GX_1X_2 - G\alpha_2X_1Y_2 - G\alpha_1Y_1X_2 + G\alpha_1\alpha_2Y_1Y_2 = X_3 - \alpha_3Y_3 \\ &= (A_x - \alpha_3A_y)X_1X_2 + (B_x - \alpha_3B_y)X_1Y_2 + (C_x - \alpha_3C_y)Y_1X_2 + (D_x - \alpha_3D_y)Y_1Y_2\,, \end{aligned}$$

where, for $i = 1$, 2, and 3, we have set $\alpha_i = (-v_i + \sqrt{D}\,)/2u_i$.

Because $4u_iw_i = v_i^2 - D$, it follows that each $\mathbb{Z}$-module $\langle u_i, \frac{1}{2}(-v_i + \sqrt{D}\,)\rangle$ is an ideal in the ring of integers of $\mathbb{Q}(\sqrt{D})$. Thus

$$\begin{aligned}(u_1X_1 - \tfrac{1}{2}(-v_1 + \sqrt{D}\,)Y_1)(u_2X_2 - \tfrac{1}{2}(-v_2 + \sqrt{D}\,)Y_2)\\ = G(u_3X_3 - \tfrac{1}{2}(-v_3 + \sqrt{D}\,)Y_3)\end{aligned}$$

details multiplication of fractional ideals

$$\langle u_1, \tfrac{1}{2}(-v_1 + \sqrt{D}\,)\rangle \cdot \langle u_2, \tfrac{1}{2}(-v_2 + \sqrt{D}\,)\rangle = G\langle u_3, \tfrac{1}{2}(-v_3 + \sqrt{D}\,)\rangle\,. \tag{10}$$

6.3. Note that *composition*, as detailed by Theorem 1 and the construction of §5, is defined up to u_3, $v_3 \pmod{2u_3}$, and $4u_3w_3 = D - v_3^2$; and thus it well defines multiplication of ideals.

6.4. We get a useful alternative emphasis as follows. Denote by δ a quadratic irrational integer with norm n and trace t; so $\delta^2 - t\delta + n = 0$. Let P and Q be integers and consider elements $\alpha = (\delta + P)/Q$, where Q divides the norm $n + tP + P^2$ of $\delta + P$. Then the $\mathbb{Z}$-module $\langle Q, \delta + P\rangle$ is an ideal in the order $\mathbb{Z}[\delta]$ which corresponds in the sense just hinted at to the quadratic form $\varphi(X, Y) = Q(X - \alpha Y)(X - \overline{\alpha}Y)$. We have $u = Q$, $v = -(t + 2P)$, $w = (n + tP + P^2)/Q$. More to the point, perhaps, we have $D = t^2 - 4n$, $-\frac{1}{2}(v_1 + v_2) = -s = P_1 + P_2 + t$, and $-\frac{1}{2}(v_1 - v_2) = m = P_1 - P_2$.

6.5. It is now clear that our remarks generalise readily to quadratic forms defined over function fields of arbitrary characteristic, *including* characteristic 2.

7. Reduction

The coefficients of a reduced form of discriminant D are of size $O(|D|^{1/2})$. However, while the coefficients of a compound of two reduced forms generally are of size $O(|D|)$, the entries of the magic matrix are just of size $O(|D|^{1/2})$. The essence of NUCOMP [9] is to reduce M, indirectly reducing the composite φ_3 by working with four-tuples of half-sized integers. Atkin's refinement [1] reduces the bit-complexity of this computation.

Atkin tells of computing just G and B_x. He then does an extended partial Euclidean algorithm on B_x and B_y until the smaller[1] is at most $L = \lfloor\, |D/4|^{1/4}\,\rfloor$. He now, in effect, has the first two columns of

$$\mu = \begin{bmatrix} a_x & b_x & c_x & d_x \\ a_y & b_y & c_y & d_y \end{bmatrix},$$

and then finds the other two columns, say by noting that each of the 4-tuples

$$\begin{bmatrix} m & -u_2 & u_1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} w_2 & -s & 0 & u_1 \end{bmatrix}$$

is orthogonal to both rows of M. Specifically, the equations

$$\begin{aligned} c_xu_1/G &= b_xu_2/G - ma_x/G, \\ c_yb_x &= b_yc_x + m, \\ d_xu_1/G &= b_xs/G - w_2a_x/G, \\ d_ya_x &= d_xa_y + s \end{aligned} \tag{11}$$

sequentially provide the unknowns from exact division of integers not likely to be larger than $O(|D|^{3/4})$ by integers of expected size $O(|D|^{1/2})$. In case $b_x = 0$ it helps to notice also that $\begin{bmatrix} w_1 & 0 & -s & u_2 \end{bmatrix}$ is orthogonal to both rows of M.

[1] We remark in [7] that it is more efficient to take one more step so that *both* are at most L.

"'Conceptually', we reduce M to μ by row = row − constant · other row, alternately, least absolute remainders. We get a form 'very close' to reduced." [1].

8. Distance

8.1. The continued fraction expansion $\beta = [\, a_0 \,,\, a_1 \,,\, a_2 \,,\, \dots\,]$ of a real irrational is defined by appropriate conditions on the "remainders" ρ_h, say $-1 < \rho_h < 0$, and

$$\beta_0 = \beta\,; \quad \beta_h = a_h - \rho_h\,; \quad \beta_{h+1} = -1/\rho_h\,.$$

The expansion yields the convergents x_h/y_h to β by the rule

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_h & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_h & x_{h-1} \\ y_h & y_{h-1} \end{pmatrix}, \quad h = -1,\, 0,\, 1,\, 2,\, \dots,$$

and for those h it entails

$$\rho_0\rho_1\rho_2\cdots\rho_h = x_h - y_h\beta\,. \tag{12}$$

These formulas are formal. They do not depend on the rules selecting the remainders and, hence, do not depend on the "choice" of partial quotients a_h.

8.2. If two ideals, $\mathfrak{a}$ and $\mathfrak{a}'$, are in the same class (of ideals modulo principal ideals), then there is a principal fractional ideal (λ) so that $\mathfrak{a}' = (\lambda)\mathfrak{a}$. Suppose $D > 0$. One calls $\delta(\mathfrak{a}, \mathfrak{a}') = \log|\lambda|$ the Shanks *distance* from $\mathfrak{a}$ to $\mathfrak{a}'$. Because λ is only defined up to units of $\mathbb{Q}(\sqrt{D}\,)$, it follows that the distance is defined only modulo the regulator $\mathcal{R}$ of $\mathbb{Q}(\sqrt{D}\,)$ and should be viewed as lying in $\mathbb{R}/\mathcal{R}\mathbb{Z}$. It will be convenient to speak of the element λ as giving the *Distance* from $\mathfrak{a}$ to $\mathfrak{a}'$.

8.3. If two elements γ and γ' are equivalent (that is, there are integers a, b, c, and d satisfying $ad - bc = \pm 1$ so that $\gamma = (a\gamma' + b)/(c\gamma' + d)$), then there are integers $a_0, a_1, \dots, a_n$, say, so that $\gamma = [\, a_0 \,,\, a_1 \,,\, \dots \,,\, a_n \,,\, \gamma'\,]$ relates γ to γ' by a continued fraction expansion. In the notation of §8.1 we have $a = x_n$, $b = x_{n-1}$, $c = y_n$, $d = y_{n-1}$; and $\gamma' = \gamma_{n+1}$.

8.4. Moreover, elements, forms, and ideals correspond one to the other, though not in altogether an obviously well-defined manner. Specifically, each triple (u, v, w) of integers satisfying $v^2 - 4uw = D$ variously provides an element $\gamma = (v - \sqrt{D}\,)/2u$, a form $uX^2 + vXY + wY^2$, and a $\mathbb{Z}$-module, thus a fractional ideal, $\langle u, \frac{1}{2}(-v + \sqrt{D}\,)\rangle$. However, if the ideal is deemed to have a sign attached, to wit the sign of u, this is a well-defined correspondence, up to v only being given modulo $2u$.

Theorem 2. *Suppose* $\gamma = (v - \sqrt{D}\,)/2u = [a_0, a_1, \dots, a_n, (v' - \sqrt{D}\,)/2u']$. *Set*

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_n & x_{n-1} \\ y_n & y_{n-1} \end{pmatrix}.$$

Then $x_n - \overline{\gamma} y_n$ *provides the Distance from the form* $uX^2 + vXY + wY^2$ *to the form* $u'X^2 + v'XY + w'Y^2$, *or equivalently from the signed ideal* $\langle u, \frac{1}{2}(-v + \sqrt{D}\,)\rangle$ *to the signed ideal* $\langle u', \frac{1}{2}(-v' + \sqrt{D}\,)\rangle$.

Remark. It is easy to see that $x_n - \overline{\gamma} y_n$ is a suitable candidate for Distance, in that it is a unit if $\gamma' = \gamma$, but the point of the claim is that it alleges that $x_n - \overline{\gamma} y_n$ is indeed a value of λ as given in §8.2. We explain that below in §9.2 in the context of our example. Note here also my use of the capitalised word "Distance" to distinguish it from its logarithm, distance proper (compare the near universal use of H for naïve height as contrasted to more sophisticated, invariably logarithmic, versions of height denoted by h).

8.5. Suppose now that a magic matrix M yields a not necessarily reduced form φ, whereas $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) M$ provides the near reduced form $\varphi'(x, y) = \varphi(dx - by, -cx + ay)$. If φ corresponds, in the sense of §8.4, to the element γ, then φ' similarly corresponds to the element γ', where $\gamma = (d\gamma' + b))/(c\gamma' + a)$. Among other things, it follows that the Distance from φ to φ' is $d - c\overline{\gamma}$. However, the point of NUCOMP is not to have to compute the unreduced form φ, nor therefore the element γ. So we note that

$$\gamma' = (a\gamma - b)/(-c\gamma + d)$$

and that the Distance from φ to φ' is $(a + c\overline{\gamma}')^{-1}$

8.6. Set $d = \lfloor D \rfloor$. Suppose $\gamma = \frac{1}{2}(d - \sqrt{D}\,)$ (that is, $u = 1$ and $v \equiv d \pmod 2$), so that the ideal $\langle +1, \frac{1}{2}(-d + \sqrt{D}\,)\rangle$, or, more precisely, the form $\mathbf{1} = \big(1, d, \frac{1}{4}(d^2 - D)\big)$, is the identity with respect to composition. Then $\mathcal{R}' = \log|x_n - y_n\overline{\gamma}|$ is *the* Shanks distance $\delta(\mathbf{1}, \varphi')$ of the form $\varphi' = \Big(u', v', \frac{1}{4}({v'}^2 - D)/u'\Big)$.

8.7. Shanks distance as just now described is computationally congenial but unnatural philosophically. Thus (10) provides $\mathcal{R}_1 + \mathcal{R}_2 = \mathcal{R}_3 + \log G$. Hendrik Lenstra [8] proposes

$$\partial(\mathbf{1}, \varphi') = \frac{1}{2}\log|(x_n - y_n\overline{\gamma})/(x_n - y_n\gamma)|$$

as the preferred definition. Then, more decently, $\partial(\mathbf{1}, \varphi_1) + \partial(\mathbf{1}, \varphi_2) = \partial(\mathbf{1}, \varphi_3)$.

8.8. In the function field case, where the x_n and y_n will be polynomials in some variable z, say, and γ is a formal Laurent series in z^{-1}, the logarithm of the absolute value becomes the degree in z of the series. Thus, in the function field case, distance takes discrete values.

9. A toy example

9.1. It is convenient to take $\omega = \frac{1}{2}\sqrt{D}$ if $D \equiv 0 \pmod 4$ and $\omega = \frac{1}{2}(1 + \sqrt{D})$ if $D \equiv 1 \pmod 4$; also set $\omega\overline{\omega} = n$ and $\omega + \overline{\omega} = t$. Then a typical step in the continued fraction expansion of ω is

$$(\omega + P_h)/Q_h = a_h - (\overline{\omega} + P_{h+1})/Q_h,$$

so

$$P_h + P_{h+1} + t = Q_h a_h;$$

and

$$-Q_h Q_{h+1} = (\omega + P_{h+1})(\overline{\omega} + P_{h+1}) = n + tP_{h+1} + P_{h+1}^2.$$

Take $D = 10209$. Then $\omega = \frac{1}{2}(1 + \sqrt{10209}\,)$, where $51 < \omega < 52$, $\omega\overline{\omega} = -2552$, $\omega + \overline{\omega} = 1$. We have

0	$(\omega + 50)/1 = 101 - (\overline{\omega} + 50)/1$	14	$(\omega + 22)/33 = 2 - (\overline{\omega} + 43)/33$
1	$(\omega + 50)/2 = 50 - (\overline{\omega} + 49)/2$	15	$(\omega + 43)/20 = 4 - (\overline{\omega} + 36)/20$
2	$(\omega + 49)/51 = 1 - (\overline{\omega} + 1)/51$	16	$(\omega + 36)/61 = 1 - (\overline{\omega} + 24)/61$
3	$(\omega + 1)/50 = 1 - (\overline{\omega} + 48)/50$	17	$(\omega + 24)/32 = 2 - (\overline{\omega} + 39)/32$
4	$(\omega + 48)/4 = 24 - (\overline{\omega} + 47)/4$	18	$(\omega + 39)/31 = 2 - (\overline{\omega} + 22)/31$
5	$(\omega + 47)/74 = 1 - (\overline{\omega} + 26)/74$	19	$(\omega + 22)/66 = 1 - (\overline{\omega} + 43)/66$
6	$(\omega + 26)/25 = 3 - (\overline{\omega} + 48)/25$	20	$(\omega + 43)/10 = 9 - (\overline{\omega} + 46)/10$
7	$(\omega + 48)/8 = 12 - (\overline{\omega} + 47)/8$	21	$(\omega + 46)/39 = 2 - (\overline{\omega} + 31)/39$
8	$(\omega + 47)/37 = 2 - (\overline{\omega} + 26)/37$	22	$(\omega + 31)/40 = 2 - (\overline{\omega} + 48)/40$
9	$(\omega + 26)/50 = 1 - (\overline{\omega} + 23)/50$	23	$(\omega + 48)/5 = 19 - (\overline{\omega} + 46)/5$
10	$(\omega + 23)/40 = 1 - (\overline{\omega} + 16)/40$	24	$(\omega + 46)/78 = 1 - (\overline{\omega} + 31)/78$
11	$(\omega + 16)/57 = 1 - (\overline{\omega} + 40)/57$	25	$(\omega + 31)/20 = 4 - (\overline{\omega} + 48)/20$
12	$(\omega + 40)/16 = 5 - (\overline{\omega} + 39)/16$	26	$(\omega + 48)/10 = 9 - (\overline{\omega} + 41)/10$
13	$(\omega + 39)/62 = 1 - (\overline{\omega} + 22)/62$	27	$(\omega + 41)/83 = 1 - (\overline{\omega} + 41)/83$

The symmetry at line 27 signals the midpoint of the period. The remaining 26 lines of the period are the conjugates of lines 26 to 1. Notice that $Q_{27} = 83$ in the symmetric line entails that 83 divides the discriminant of $\mathbb{Z}[\omega]$. Indeed, $10209 = 83 \cdot 3 \cdot 41$, displaying a nontrivial composite, 3403.

9.2. Line h of the array, namely $(\omega + P_h)/Q_h = a_h - (\overline{\omega} + P_{h+1})/Q_h$, corresponds to the form $(-1)^h Q_h x^2 + (2P_{h+1} + t)xy + (-1)^{h+1} Q_{h+1} y^2$, to the signed ideal $\langle (-1)^{h+1} Q_h\,, -(\overline{\omega} + P_{h+1})\rangle$, and to the element $(-1)^{h+1}(\overline{\omega} + P_{h+1})/Q_h$. Note that $\langle (-1)^{h+1} Q_h\,, (\omega + P_h)\rangle$ is the same ideal.

Proof of Theorem 2. It suffices to notice the evident identity

$$\begin{aligned}(\omega + P_{h+1})/Q_h \cdot \langle Q_h, -(\overline{\omega} + P_{h+1})\rangle &= \langle (\omega + P_{h+1}), Q_{h+1}\rangle \\ &= \langle Q_{h+1}, -(\overline{\omega} + P_{h+2})\rangle\,,\end{aligned}$$

the fact that $(\overline{\omega} + P_{h+1})/Q_h$ is a remainder in the sense of §8.1, and (12). □

9.3. With $D = 10209$ take $f = (-2, 99, 51)$. We obtain the *duplicate* $f * f =: f_2$ from the magic matrix

$$\begin{bmatrix} 1 & -1 & -1 & 75 \\ 0 & -2 & -2 & 99 \end{bmatrix} \longrightarrow (4, 95, -74) = g_4\,.$$

Next, $f_4 := f_2 * f_2$ comes from

$$\begin{bmatrix} 1 & 2 & 2 & 66 \\ 0 & 4 & 4 & 95 \end{bmatrix} \longrightarrow (16, 79, -62) = g_{12}\,,$$

and $f_8 := f_4 * f_4$ from

$$\begin{bmatrix} 1 & -2 & -2 & -6 \\ 0 & 16 & 16 & 79 \end{bmatrix} \longrightarrow (256, 143, 10)$$

and

$$\begin{pmatrix} 12 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} 1 & -2 & -2 & -6 \\ 0 & 16 & 16 & 79 \end{bmatrix} = \begin{bmatrix} 12 & -8 & -8 & 7 \\ -1 & 2 & 2 & 6 \end{bmatrix} \longrightarrow (10, 97, -20) = g_{28}\,.$$

The remarks "$= g_h$" refer to the line h of the continued fraction expansion of §9.1 on which the form appears.

9.4. In practice one computes the Distance as a real number $\tau \times 2^{Ne}$ with τ of appropriately limited accuracy. Here, of course, and in fact in general, we can readily keep track of the precise distance. The form f_1 has Distance $\omega + 50$, so up to the adjustment corresponding to the reduction step, f_8 has Distance $(\omega + 50)^8$. The reduction has matrix $\left(\begin{smallmatrix} 12 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, so one multiplies by $(\omega + 71)/256$, or by the reciprocal of $12 - 1(\omega + 48)/10 = (\overline{\omega} + 71)/10$, to make the adjustment.

We obtain $f_6 = f_2 * f_4$ from $\left[\begin{smallmatrix} 1 & -2 & -6 & -28 \\ 0 & 4 & 16 & 87 \end{smallmatrix}\right] \longrightarrow (64, 143, 40)$ and

$$\begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} 1 & -2 & -6 & -28 \\ 0 & 4 & 16 & 87 \end{bmatrix} = \begin{bmatrix} 3 & -2 & -2 & 3 \\ -1 & 2 & 6 & 28 \end{bmatrix} \longrightarrow (40, 97, -5) = g_{22},$$

so its Distance $(\omega + 50)^6$ is to be adjusted by $(\omega + 71)/64$, or the reciprocal of $3 - (\omega + 48)/40 = (\overline{\omega} + 71)/40$. Similarly, we obtain $f_{14} = f_6 * f_8$ from

$$\begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} 1 & 20 & 5 & 49 \\ 0 & 40 & 10 & 97 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & -1 \\ -1 & -20 & -5 & -49 \end{bmatrix} \longrightarrow (51, 99, -2) = g_{52}\,,$$

its Distance requiring adjustment by $(-2 - (\omega + 49)/51)^{-1} = ((\overline{\omega} - 152)/51)^{-1}$. That checks with the unreduced composite being $(400, -303, 51)$. It follows that the Distance to f_{14} is

$$(\omega + 50)^{14} \, (12 - 1(\omega + 48)/10)^{-1} \, (3 - (\omega + 48)/40)^{-1} \, (-2 - (\omega + 49)/51)^{-1}\,.$$

Finally, $f_{14} * f_1$ is given by

$$\begin{pmatrix} 100 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 51 & -2 & 99 \end{bmatrix} = \begin{bmatrix} 100 & 51 & -2 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow (1, 101, -2) \equiv g_0\,,$$

multiplying the Distance to f_{14} by $(\omega + 50) \times (100 - (\omega + 50)/1)^{-1}$. Thus the fundamental unit of $\mathbb{Z}[\omega]$ is given by

$$\begin{aligned} &(\omega + 50)^{15} \times (12 - 1(\omega + 48)/10)^{-1} \\ &\qquad \times (3 - (\omega + 48)/40)^{-1} \, (-2 - (\omega + 49)/51)^{-1} \, (100 - (\omega + 50))^{-1} \\ &\quad = 1296732767317670450014672368 19 + 259243902732643631595188 3912\omega. \end{aligned}$$

Note that the suffixes h on the f_h have no well-defined meaning and depend on the reduction steps performed at each composition. For example, we might have computed "f_8" as

$$\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{bmatrix} 1 & -2 & -2 & -6 \\ 0 & 16 & 16 & 79 \end{bmatrix} = \begin{bmatrix} 1 & -2 & -2 & -6 \\ 3 & 10 & 10 & 61 \end{bmatrix} \longrightarrow (-83, 83, 10) = g_{27}\,,$$

with Distance $(\omega + 50)^8$ adjusted by the reciprocal of $1 + 3(\omega + 41)/(-83)$, thus by $((3\omega + 40)/83)^{-1}$. Obviously, the duplicate of the ambiguous form $(-83, 83, 10)$ yields the identity; indeed, with $G = 83$,

$$\begin{bmatrix} 83 & 9 & 9 & 1 \\ 0 & -1 & -1 & 1 \end{bmatrix} \longrightarrow (1, 101, -2) = g_{54}\,,$$

so the fundamental unit is also given by

$$(\omega + 50)^{16} \cdot ((3\omega + 40)/83)^{-2} \cdot 1/83\,.$$

Compare also Henri Cohen's computation with this example in [3], §5.8.3. Note that this truly is a toy example in that the magic matrices M produced already are near reduced; the subsequent reductions, when necessary, are quite artificial.

10. NUCOMP for indefinite forms

The following algorithm in effect combines the content of algorithms 5.8.6 and 5.4.9 of [3] in using NUCOMP to compose quadratic forms while retaining a record of their Distance. NUCOMP supposes that the forms to be composed are distinct; if not, use NUDUPL, below.

Algorithm 3 (NUCOMP with Distance)**.** For quadratic forms $\varphi_1 = (u_1, v_1, w_1)$ and $\varphi_2 = (u_2, v_2, w_2)$ with the same discriminant D and at Distances $\tau_1 \times 2^{Ne_1}$ and $\tau_2 \times 2^{Ne_2}$ respectively, this algorithm computes a near reduced composite $\varphi_3 = (u_3, v_3, w_3)$ of φ_1 and φ_2 and obtains its Distance $\tau_3 \times 2^{Ne_3}$. It is supposed that $L = \lfloor |D/4|^{1/4} \rfloor$ has been precomputed. The first four steps construct the following part of the magic matrix:

$$M = \begin{bmatrix} A_x = G & B_x & — & — \\ 0 & B_y = u_1/G & C_y = u_2/G & D_y = s/G \end{bmatrix}.$$

The "near reduction" of step 5 yields

$$\mu = \begin{pmatrix} x & \dots \\ y & \dots \end{pmatrix} M = \begin{bmatrix} a_x = Gx & b_x & — & — \\ a_y = Gy & b_y & — & — \end{bmatrix},$$

and the remainder of the algorithm reports the evaluation of the missing entries of μ and the consequent evaluation of the near reduced composite and its Distance.

(1) [Initialisation] Set $s \leftarrow \frac{1}{2}(v_1 + v_2)$; then $m \leftarrow v_2 - s$.

(2) Use Euclid's extended algorithm to compute (b, c, F) such that $bu_2 + cu_1 = F = \gcd(u_1, u_2)$. It is mostly the case that $F = 1$. In any case, if $F \mid s$, so in particular if $F = 1$, set $A_x \leftarrow G = F$, $B_x \leftarrow (mb \mod B_y)$, $B_y \leftarrow u_1/F$, $C_y \leftarrow u_2/F$, $D_y = s/F$, and go to step 5.

(3) [If $F \nmid s$] However, if $F \nmid s$, use Euclid's extended algorithm again to compute (x, y, G) so that $xF + ys = G = \gcd(F, s)$, and set $H \leftarrow F/G$. Also set $B_y \leftarrow u_1/G$, $C_y \leftarrow u_2/G$, $D_y \leftarrow s/G$.

(4) Compute $l \leftarrow y(bw_1 + cw_2) \mod H$ by first reducing w_1 and w_2 (which are large) modulo H (which is small), doing the operation, and reducing again. Set $B_x \leftarrow bm/H + l \cdot B_y/H \pmod{B_y}$.

(5) [A Subalgorithm] Set $b_x \leftarrow (B_x \mod B_y)$ and $b_y \leftarrow B_y$. Then execute a partial Euclidean algorithm on b_x, b_y:
   (a) Set $x \leftarrow 1$, $y \leftarrow 0$; and set $z \leftarrow 0$.
   (b) If $|b_x| > L$ go to substep 5c. Otherwise, if z is odd set $b_y \leftarrow -b_y$, $y \leftarrow -y$. Then set $a_x \leftarrow Gx$, $a_y \leftarrow Gy$. Terminate this subalgorithm.
   (c) Let $q \leftarrow \lfloor b_y/b_x \rfloor$ and simultaneously $t \leftarrow b_y \mod b_x$. Now set $b_y \leftarrow b_x$ and $b_x \leftarrow t$. Then set $t \leftarrow y - qx$, followed by $y \leftarrow x$ and $x \leftarrow t$. Finally let $z \leftarrow z + 1$ and go back to substep 5b.

(6) [Computation of Near Reduced Composite] Set $c_x \leftarrow (b_x C_y - mx)/B_y$, and compute $c_y \leftarrow (b_y c_x + m)/b_x$. If $b_x = 0$, set $c_y \leftarrow (u_2 b_y - ym)/u_1$. Similarly set $d_x \leftarrow (b_x D_y - w_2 x)/B_y$ and $d_y \leftarrow (d_x y + D_y)/x$. Then the near reduced composite form φ_3 is given by $u_3 \leftarrow b_y c_y - a_y d_y$, $w_3 \leftarrow b_x c_x - a_x d_x$ and $u_3 \leftarrow (a_x d_y + a_y d_x) - (b_x c_y + b_y c_x)$. Its Distance $\tau_3 \times 2^{Ne_3}$ is $(\tau_1 \tau_2/G) \times \left(x + y(v_3 + \sqrt{D})/2u_3\right)^{-1} \times 2^{N(e_1+e_2)}$. Sequentially adjust when $\tau_3 > 2^N$ by setting $\tau_3 \leftarrow \tau_3/2^N$ and $e_3 \leftarrow e_3 + 1$.

Notes. (*i*) The divisions in step 6 are exact and can therefore be done at the prevailing precision $O(|D|^{1/2})$. Specifically, make these floating point rather than integer computations, at that precision.

(*ii*) If done naïvely, certain of the steps, for example $B_x \leftarrow mb \pmod{B_y}$, will involve quantities greater than the "prevailing" precision. One should either not do them naïvely or treat them as multiprecision steps.

(*iii*) Notice that the quantities making up v_3 will usually have been computed in the just preceding calculations.

(*iv*) Cohen [3], §5.4.9, singles out the case $z = 0$—thus $x = 1$, $y = 0$—for special mention. He has the equivalent of

> [Special Case: $z = 0$] Set $c_x \leftarrow (b_x C_y - m)/B_y$, and $c_y \leftarrow (b_y c_x + m)/b_x$. If $b_x = 0$, set $c_y = (b_y a_2)/a_1$. Similarly set $d_x \leftarrow (b_x D_y - w_2)/B_y$ and $d_y \leftarrow D_y$. Then φ_3 is given by $u_3 \leftarrow b_y c_y$ and $w_3 \leftarrow b_x c_x - G d_x$, while $v_3 \leftarrow G d_y - (b_x c_y + b_y c_x)$. Its Distance $\tau_3 \times 2^{N e_3}$ is $(\tau_1 \tau_2 / G) \times 2^{N(e_1+e_2)}$. Sequentially adjust when $\tau_3 > 2^N$ by setting $\tau_3 \leftarrow \tau_3/2^N$ and $e_3 \leftarrow e_3 + 1$.

as intermediate step between step 5 and step 6.

(*v*) Moreover, it is desirable to single out the case $\varphi_1 = \varphi_2$, and we do that below.

Algorithm 4 (NUDUPL with Distance)**.** Given a quadratic form $\varphi = (u, v, w)$ of discriminant D and at Distance $\tau \times 2^{Ne}$, this algorithm computes a near reduced duplicate $\varphi_3 = (u_3, v_3, w_3)$ of φ and obtains its Distance $\tau_3 \times 2^{N e_3}$. It is supposed that $L = \lfloor |D/4|^{1/4} \rfloor$ has been precomputed.

1. Use Euclid's extended algorithm to compute (x, y, G) so that $xu + yv = G = \gcd(u, v)$, and set $A_x \leftarrow G$, $B_y \leftarrow u/G$, $D_y \leftarrow v/G$.
2. Compute $B_x \leftarrow (yw \mod B_y)$.
3. [A Subalgorithm] Set $b_x \leftarrow B_x$, $b_y \leftarrow B_y$. Then execute a partial Euclidean algorithm on b_x, b_y:
   (a) Set $x \leftarrow 1$, $y \leftarrow 0$; and set $z \leftarrow 0$.
   (b) If $|b_x| > L$ go to substep 3c. Otherwise, if z is odd set $b_y \leftarrow -b_y$, $y \leftarrow -y$. Then set $a_x \leftarrow Gx$, $a_y \leftarrow Gy$. Terminate this subalgorithm.
   (c) Let $q \leftarrow \lfloor b_y/b_x \rfloor$ and simultaneously $t \leftarrow b_y \mod b_x$. Now set $b_y \leftarrow b_x$ and $b_x \leftarrow t$. Then set $t \leftarrow y - qx$, followed by $y \leftarrow x$ and $x \leftarrow t$. Finally let $z \leftarrow z + 1$ and go back to substep 3b.
4. [Computation of Near Reduced Composite] Set $d_x \leftarrow (b_x D_y - wx)/B_y$ and $d_y \leftarrow (d_x y + D_y)/x$. Then the near reduced composite form φ_3 is given by $u_3 \leftarrow b_y^2 - a_y d_y$, $w_3 \leftarrow b_x^2 - a_x d_x$ and $v_3 \leftarrow (a_x d_y + a_y d_x) - 2 b_x b_y$. Its Distance $\tau_3 \times 2^{N e_3}$ is $(\tau^2/G) \times \left(x + y(v_3 + \sqrt{D})/2u_3\right)^{-1} \times 2^{2Ne}$. Sequentially adjust when $\tau_3 > 2^N$ by setting $\tau_3 \leftarrow \tau_3/2^N$ and $e_3 \leftarrow e_3 + 1$.

Note. Our previous notes apply, *mutatis mutandis*.

One now reduces the composite φ_3 by a standard algorithm, a process that should not take more than a step or two, and adjusts the Distance appropriately.

References

[1] A. O. L. Atkin, Letter to Dan Shanks on the programs NUDUPL and NUCOMP, 12 December 1988; from the Nachlaß of D. Shanks and made available to me by Hugh C. Williams.

[2] Duncan A. Buell, "*Binary quadratic forms. Classical theory and modern computations*", Springer-Verlag, New York, 1989, x+247pp. MR **92b:**11021

[3] Henri Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics, 138, Springer-Verlag, Berlin, 1993, xii+534pp. MR **94i:**11105

[4] Felix Klein, *Elementary mathematics from an advanced standpoint*: *Geometry*, reprint (New York: Dover, 1939); see §II*ff*.

[5] Hermann Grassmann, *A New Branch of Mathematics*, Open Court: Chicago and La Salle, Illinois: 1995. MR **99e:**01015

[6] W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry* (Cambridge: Cambridge University Press, 1953); see Vol.1, Chapter VII. MR **95d:**14002a

[7] Michael J Jacobson Jr and Alfred J van der Poorten, "Computational aspects of NUCOMP", to appear in Claus Fieker and David Kohel eds, *Algorithmic Number Theory* (Proc. Fifth International Symposium, ANTS-V, Sydney, NSW, Australia July 2002), Springer Lecture Notes in Computer Science **2369** (2002), 120–133.

[8] H. W. Lenstra Jr., "On the calculation of regulators and class numbers of quadratic fields", in J. V. Armitage ed., "*Journées Arithmétiques 1980*" LMS Lecture Notes **56**, Cambridge 1982, 123–151. MR **86g:**11080

[9] D. Shanks, "Class number, a theory of factorization, and genera", in *Proc. Symp. Pure Math.* **20** (1969 Institute on Number Theory), Amer. Math. Soc., Providence 1971, 415–440; see also "The infrastructure of a real quadratic field and its applications", *Proc. Number Theory Conference*, Univ. of Colorado, Boulder, CO, 1972, 217-224. MR **47:**4932; MR **52:**10762

[10] Daniel Shanks, "On Gauss and composition", in *Number Theory and Applications*, ed. Richard A. Mollin, (NATO–Advanced Study Institute, Banff, 1988) (Kluwer Academic Publishers Dordrecht, 1989), 163–204. MR **92e:**11150

ceNTRe for Number Theory Research, 1 Bimbil Pl. Killara, New South Wales 2071, Australia

E-mail address: alf@math.mq.edu.au