



ELSEVIER

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt



On a remarkable identity in class numbers of cubic rings



Evan O'Dorney

Department of Mathematics, Princeton University, Fine Hall, Washington Rd,
Princeton, NJ 08544, United States

ARTICLE INFO

Article history:

Received 30 August 2016

Received in revised form 14

December 2016

Accepted 15 December 2016

Available online 10 February 2017

Communicated by D. Goss

MSC:

11E76

11R16

11M41

11R37

Keywords:

Shintani zeta functions

Binary cubic forms

Higher composition laws

ABSTRACT

In 1997, Y. Ohno empirically stumbled on an astoundingly simple identity relating the number of cubic rings $h(\Delta)$ of a given discriminant Δ , over the integers, to the number of cubic rings $\hat{h}(\Delta)$ of discriminant -27Δ in which every element has trace divisible by 3:

$$\hat{h}(\Delta) = \begin{cases} 3h(\Delta) & \text{if } \Delta > 0 \\ h(\Delta) & \text{if } \Delta < 0, \end{cases} \quad (1)$$

where in each case, rings are weighted by the reciprocal of their number of automorphisms. This allows the functional equations governing the analytic continuation of the Shintani zeta functions (the Dirichlet series built from the functions h and \hat{h}) to be put in self-reflective form. In 1998, J. Nakagawa verified (1). We present a new proof of (1) that uses the main ingredients of Nakagawa's proof (binary cubic forms, recursions, and class field theory), as well as one of Bhargava's celebrated higher composition laws, while aiming to stay true to the stark elegance of the identity.

© 2017 Elsevier Inc. All rights reserved.

E-mail address: eodorney@princeton.edu.

<http://dx.doi.org/10.1016/j.jnt.2016.12.002>

0022-314X/© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Great progress has been made in recent years [15] in analyzing statistics pertaining to cubic fields, ordered by discriminant. A basic analytic tool at one's disposal is the *Shintani zeta functions*, a pair of Dirichlet series that encode the number of cubic *rings* over \mathbb{Z} of each nonzero discriminant:

$$\zeta^+(s) = \sum_{\substack{C/\mathbb{Z} \text{ cubic,} \\ \text{Disc } C > 0}} \frac{(\text{Disc } C)^{-s}}{|\text{Aut } C|}$$

$$\zeta^-(s) = \sum_{\substack{C/\mathbb{Z} \text{ cubic,} \\ \text{Disc } C < 0}} \frac{(-\text{Disc } C)^{-s}}{|\text{Aut } C|}.$$

The division by the number of automorphisms is a standard trick in this discipline which ensures, among other things, that the relative weights of a ring and its subrings (some of which may be isomorphic) are in the proper ratio. Because almost all cubic fields (and rings) have trivial automorphism group, this factor has no effect in most analytic applications.

The Shintani zeta functions were introduced in 1972 by Shintani, who proved that they have meromorphic continuations to the whole complex plane satisfying a reflection formula of the form (see [10], eq. (0.1))

$$\begin{bmatrix} \zeta^+(1-s) \\ \zeta^-(1-s) \end{bmatrix} = \begin{bmatrix} c_1(s) & c_2(s) \\ c_3(s) & c_4(s) \end{bmatrix} \begin{bmatrix} \hat{\zeta}^+(s) \\ \hat{\zeta}^-(s) \end{bmatrix} \quad (2)$$

connecting them to two other Dirichlet series $\hat{\zeta}^+$ and $\hat{\zeta}^-$ (the c_i , which are certain elementary expressions involving the Γ function, need not detain us). The functions $\hat{\zeta}^+$ and $\hat{\zeta}^-$ arise as follows. Call a cubic ring *integer-matrix*, or \mathbb{Z} -*mat* for short, if the trace of each of its elements is a multiple of 3. (This name will be demystified in the next section.) The discriminant of such a ring is always divisible by 27, making the scaling of the following Dirichlet series natural:

$$\hat{\zeta}^+(s) = 3^{3s} \sum_{\substack{C/\mathbb{Z} \text{ } \mathbb{Z}\text{-mat,} \\ \text{Disc } C > 0}} \frac{(\text{Disc } C)^{-s}}{|\text{Aut } C|}$$

$$\hat{\zeta}^-(s) = 3^{3s} \sum_{\substack{C/\mathbb{Z} \text{ } \mathbb{Z}\text{-mat,} \\ \text{Disc } C < 0}} \frac{(-\text{Disc } C)^{-s}}{|\text{Aut } C|}.$$

Shintani's functional equation stood unimproved until 1997, when Y. Ohno computed the first 200 terms of each of the four zeta functions and conjectured that they are equal in pairs, up to a curiously sign-dependent scale factor [12]:

$$\hat{\zeta}^+(s) = \zeta^-(s) \quad \text{and} \quad \hat{\zeta}^-(s) = 3\zeta^+(s).$$

This implies that the Shintani zeta functions satisfy a self-reflective functional equation, just like the Riemann zeta function. This striking conjecture was verified by Nakagawa the following year. In purely algebraic form, it is the following, which will be the subject of this essay.

Theorem 1.1. *Let $h(\Delta)$ denote the number of cubic rings of discriminant Δ , each weighted by the reciprocal of its number of automorphisms. Let $\hat{h}(\Delta)$ denote the number of \mathbb{Z} -mat cubic rings of discriminant -27Δ , weighted in the same manner. Then for each integer $\Delta \neq 0$,*

$$\hat{h}(\Delta) = \begin{cases} 3h(\Delta) & \text{if } \Delta > 0 \\ h(\Delta) & \text{if } \Delta < 0. \end{cases} \quad (3)$$

Developments in number theory since 1998, specifically Bhargava's beautiful work in higher composition laws in the early 2000's, suggest revisiting this beautiful identity. (A *higher composition law*, in Bhargava's parlance, is a parametrization of interesting algebraic objects by the orbits of an algebraic group action [3]; it need not be a group operation.) In particular, one of the main steps in Nakagawa's proof relates \mathbb{Z} -mat rings of discriminant -27Δ to ideals in orders of the quadratic algebra $\mathbb{Q}(\sqrt{\Delta})$, and one of Bhargava's higher composition laws relates the same sort of objects. Can Bhargava's result be adapted as a replacement for Nakagawa's somewhat ad hoc computation? We answer this question affirmatively. We also find simple recursive formulas for $h(\Delta)$ and $\hat{h}(\Delta)$ valid when Δ has high prime power divisors (Theorem 4.1). Finally, unlike Nakagawa, we treat the cases $\Delta > 0$ and $\Delta < 0$ simultaneously, enabling us to explain the factor of 3 in the statement quite readily. It arises from the existence of a fundamental unit in $\mathbb{Q}(\sqrt{\Delta})$, except when Δ is a square, in which case it arises from the extra automorphism of order 3 belonging to cubic fields of square discriminant.

Example 1.2. The simplest case of Theorem 1.1 is when $\Delta = 1$. There is just one cubic ring of discriminant 1, namely $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$, and it has six automorphisms, so $h(1) = 1/6$. There is also just one \mathbb{Z} -mat ring of discriminant -27 , namely $\mathbb{Z}[t]/(t^3 - 1)$, and it has a single nontrivial automorphism $t \mapsto t^2$, so $\hat{h}(1) = 1/2$, in accordance with the theorem.

1.1. Outline of the proof

Our proof of Theorem 1.1 follows four main steps:

1. Construct a recursion allowing one to reduce to the case where the prime powers dividing Δ are not too high (Section 4).
2. Use Bhargava's theory of higher composition laws to relate cubic rings of discriminant -27Δ to certain ideals in orders of the quadratic algebra $\mathbb{Q}(\sqrt{\Delta})$ (Section 5).

3. Use class field theory to relate cubic fields of discriminant Δ to certain characters on the ideal group of $\mathbb{Q}(\sqrt{\Delta})$ (Section 7).
4. Combine the foregoing steps to prove the theorem (Section 8).

The first three steps are completely independent, and we have chosen to order them in a manner that places the non-elementary material last. Each of the steps culminates in a theorem that has an analogue in Nakagawa’s proof, though potentially with some conditions altered, or, in the case of step 1, a beautiful and apparently new recursive formula for $h(\Delta)$ and $\hat{h}(\Delta)$.

At first glance, the two sides of (3) are analogous, even “dual” to each other. Indeed, the space \mathbb{Q}^4 of rational binary cubic forms, which parametrize cubic rings over \mathbb{Q} with a chosen basis, has a natural $\mathrm{SL}_2\mathbb{Q}$ -invariant skew form $aa' - \frac{1}{3}bb' + \frac{1}{3}cc' - dd'$, with respect to which the lattices of integral and \mathbb{Z} -mat cubic forms are mutually dual; this duality was used by Shintani to establish the functional equation (2) in [14]. By contrast, h and \hat{h} are treated asymmetrically in Nakagawa’s proof and even more asymmetrically in the present one: we only apply class field theory to h and Bhargava’s parametrizations to \hat{h} , allowing us to minimize the amount of time spent treating the prime 3 specially.

Incidentally, there are ten lattices $L \subset \mathbb{Q}^4$ of binary cubic forms, up to scaling, which are invariant under the action of $\mathrm{SL}_2\mathbb{Z}$. Ohno and Taniguchi have verified the analogue of Theorem 1.1 for each of the five pairs of dual lattices [13]. The author aspires to elucidate the reasons behind the existence of such deep symmetries in a future paper.

2. Basic notions

Let A be a principal ideal domain (PID); quintessentially $A = \mathbb{Z}$, although we will also use $A = \mathbb{Z}_p$ in this paper. By an n -ic ring, or a ring of rank n , over A we will mean a commutative ring C with unit which is isomorphic to A^n as an A -module. Only quadratic ($n = 2$) and cubic ($n = 3$) rings concern us here.

The *discriminant* $\mathrm{Disc} C$ of an n -ic ring is, as usual, the determinant of the trace pairing matrix $[\mathrm{tr} \alpha_i \alpha_j]_{i,j=1}^n$, where $[\alpha_1, \dots, \alpha_n]$ is any A -basis for C . It is well defined up to multiplication by the square of a unit in A , so if $A = \mathbb{Z}$, the discriminant is simply an integer, while if $A = \mathbb{Z}_p$, a discriminant is determined up to a finite list of possibilities by its valuation $v_p(\mathrm{Disc} C)$. A ring C is called *nondegenerate* if its discriminant is nonzero, in which case C is an order in some product $K_1 \times \dots \times K_r$ of finite extensions of $K = \mathrm{Frac} A$.

A classical theorem due to Stickelberger states that the discriminant of a number field, and hence of any finite-rank ring over \mathbb{Z} , is congruent to 0 or 1 modulo 4. In the case of a cubic ring, we will soon give a direct proof. We mention Stickelberger’s theorem here only to motivate the following definitions. Let

$$\mathcal{D}iscs = (4\mathbb{Z} \setminus 0) \cup (1 + 4\mathbb{Z})$$

be the set of all possible discriminants for a nondegenerate \mathbb{Z} -algebra. Note that there is exactly one quadratic \mathbb{Z} -algebra of each discriminant $\Delta \in \mathcal{D}iscs$; we denote it by \mathcal{O}_Δ . Call $\Delta \in \mathcal{D}iscs$ a *fundamental discriminant* if Δ is not of the form $\Delta'k^2$, where $k > 1$ and $\Delta' \in \mathcal{D}iscs$. The fundamental discriminants are exactly those Δ such that \mathcal{O}_Δ is maximal (being either $\mathbb{Z} \times \mathbb{Z}$ or the ring of integers of a quadratic field). A general $\Delta \in \mathcal{D}iscs$ can be written uniquely in the form $\Delta_0 f^2$, where $f \geq 1$ and Δ_0 is fundamental; we have an identification $\mathcal{O}_\Delta \cong \mathbb{Z} + f\mathcal{O}_{\Delta_0}$.

Analogously, let $\mathcal{D}iscs_p$ be the set of all possible discriminants for a nondegenerate \mathbb{Z}_p -algebra, namely

$$\mathcal{D}iscs_p = \begin{cases} (4\mathbb{Z}_2 \setminus 0) \cup (1 + 4\mathbb{Z}_2) & p = 2 \\ \mathbb{Z}_p \setminus 0 & \text{otherwise.} \end{cases}$$

Call $\Delta \in \mathcal{D}iscs_p$ *fundamental* if it is not p^2 times an element of $\mathcal{D}iscs_p$. This is the same as requiring that the unique quadratic ring over \mathbb{Z}_p of discriminant Δ be maximal. One computes that the fundamental p -adic discriminants are, for $p \neq 2$, those not divisible by p^2 , and for $p = 2$, those congruent to 1 (mod 4) or to 8 or 12 (mod 16).

If K is a nondegenerate \mathbb{Q} -algebra and \mathcal{O}_K is the integral closure of \mathbb{Z} in K , then the *splitting type* of a prime p is the symbol $f_1^{e_1} \cdots f_r^{e_r}$, where the f_i and e_i are the degrees and ramification indices of the primes into which p splits in \mathcal{O}_K , or equivalently of the extensions of \mathbb{Q}_p into which the completed algebra K_p splits. The splitting type may be defined uniformly regardless of whether K itself is a field.

3. Cubic rings and binary cubic forms

Cubic rings may be studied by means of a correspondence with binary cubic forms, stated for nondegenerate rings over \mathbb{Z} by Delone and Faddeev [5] building on earlier work of F.W. Levi [8], and generalized to an arbitrary PID by Gross and Lucianovic [6]. Following Delone and Faddeev, we call these forms *index forms*, because they essentially attach to each ring element $x \in C$ the index $[C : \mathbb{Z}[x]]$ of the suborder it generates, up to sign. Bhargava ([2], pp. 868–869) discovered an attractive coordinate-free formulation which we follow here.

Theorem 3.1 ([6], Prop. 2.1). *Let A be a PID. The association to a cubic ring C of the index form*

$$\xi \mapsto 1 \wedge \xi \wedge \xi^2 : C/A \rightarrow \Lambda^3 C$$

defines a bijection between isomorphism classes of cubic rings over A and orbits of binary cubic forms

$$\phi(x, y) = ax^3 + bx^2y + cxy^2 + dy^3, \quad a, b, c, d \in A$$

under the $\mathrm{GL}_2 A$ -action

$$\left(\begin{bmatrix} p & q \\ r & s \end{bmatrix} \cdot \phi \right) (x, y) = \frac{1}{ps - qr} \phi(px + ry, qx + sy). \quad (4)$$

Moreover, the automorphism group of C , as an A -algebra, is isomorphic to the stabilizer in $\mathrm{GL}_2 A$ of the corresponding index form.

Proof. First note that for $\xi \in C$ and $n \in A$, we formally have

$$1 \wedge (\xi + n) \wedge (\xi + n)^2 = 1 \wedge \xi \wedge \xi^2,$$

so $\phi_C(x) = 1 \wedge \xi \wedge \xi^2$ really does define a cubic map from C/A to $\Lambda^3 C$. If we pick a basis $[\bar{\alpha}, \bar{\beta}]$ for C/A lifting to some basis $[1, \alpha, \beta]$ of C , then $\Lambda^3 C$ acquires a distinguished generator $1 \wedge \alpha \wedge \beta$ and $\phi : A^2 \rightarrow A$ becomes a binary cubic form.

Switching to a different basis $[\bar{\alpha}', \bar{\beta}'] = [p\alpha + q\beta, r\alpha + s\beta]$ multiplies the distinguished generator of $\Lambda^3 C$ by the determinant $ps - qr$ and thus changes the form ϕ in the manner indicated in (4).

To establish the bijection, it is enough to show that every cubic form ϕ arises from exactly one cubic ring C with distinguished basis $[\bar{\alpha}, \bar{\beta}]$ for C/A in this way. First note that the selection of basis $[\bar{\alpha}, \bar{\beta}]$ is tantamount to a selection of a *normal basis* for C , that is, a basis $[1, \alpha, \beta]$ such that $\alpha\beta \in A$: if α', β' are any lifts of $\bar{\alpha}$ and $\bar{\beta}$, then

$$\alpha'\beta' = t + u\alpha' + v\beta' \quad (t, u, v \in A),$$

and $[1, \alpha' - v, \beta' - u]$ is the unique such basis.

Now write the multiplication table of C , still undetermined, in terms of this basis:

$$\alpha^2 = \ell - b\alpha + a\beta$$

$$\alpha\beta = m$$

$$\beta^2 = n - d\alpha + c\beta,$$

where the signs and letters will be motivated momentarily. We compute

$$\begin{aligned} \phi(x, y) &= 1 \wedge (\alpha x + \beta y) \wedge (\alpha x + \beta y)^2 \\ &= 1 \wedge (\alpha x + \beta y) \wedge [(\ell - b\alpha + a\beta)x^2 + mxy + (n - d\alpha + c\beta)] \\ &= (ax^3 + bx^2y + cxy^2 + dy^3)(1 \wedge \alpha \wedge \beta). \end{aligned}$$

So the cubic form ϕ exactly carries the information of the four coefficients a, b, c , and d . Expanding out the associative laws $(\alpha^2)\beta = \alpha(\alpha\beta)$ and $(\alpha\beta)\beta = \alpha(\beta^2)$ shows that the conditions for this multiplication table to define a ring are

$$\ell = -ac, \quad m = -ad, \quad n = -bd.$$

In particular, each choice of a , b , c , and d yields precisely one ring structure, establishing the bijection.

An A -algebra automorphism σ of the ring C clearly induces an automorphism of the module C/A such that the cubic forms induced by bases $[\bar{\alpha}, \bar{\beta}]$ and $[\sigma(\bar{\alpha}), \sigma(\bar{\beta})]$ are the same (for any fixed basis $[\bar{\alpha}, \bar{\beta}]$). Conversely, if some $\sigma : C/A \rightarrow C/A$ has this property, it arises from a unique automorphism of R , namely the linear map that sends the normal basis lifting $[\bar{\alpha}, \bar{\beta}]$ to the normal basis lifting $[\sigma(\bar{\alpha}), \sigma(\bar{\beta})]$. This establishes a bijection between the automorphism groups, which is easily seen to be a group isomorphism. \square

We will have occasion to use the index form ϕ in many contexts: sometimes as a coordinate-free map $\phi : C/A \rightarrow \Lambda^3 C$, sometimes in a specific basis as a polynomial $\phi : A^2 \rightarrow A$. Sometimes we will be plugging an element of C/A into ϕ , but treating the output as a number in A ; this requires one to choose a generator ω_C of $\Lambda^3 C$, otherwise known as an *orientation* on C , and we write

$$\phi(\xi) = \frac{1}{\omega_C} 1 \wedge \xi \wedge \xi^2.$$

Happily enough, the index form corresponding to a monogenic ring $A[\xi]/(\xi^3 + b\xi^2 + c\xi + d)$ is simply the homogenized form

$$\phi(x, y) = x^3 + bx^2y + cxy^2 + dy^3$$

(take the normal basis $[1, \xi, \xi^2 + b\xi + c]$). This leads to a quick proof of the identity that the discriminant of the ring C corresponding to a form ϕ is the usual polynomial discriminant

$$\text{Disc } \phi = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd, \quad (5)$$

just by noting that both sides are homogeneous polynomials in a , b , c , and d of degree 4 that coincide when $a = 1$. Note that this immediately implies Stickelberger's theorem that (when $A = \mathbb{Z}$) $\text{Disc } C \equiv 0, 1 \pmod{4}$, since

$$\text{Disc } \phi \equiv (bc - ad)^2 \equiv 0 \text{ or } 1 \pmod{4}.$$

The index form has one other beautiful property, which was Levi's original reason for studying it for maximal cubic orders [8]: if C is a maximal cubic \mathbb{Z} -algebra, then for any prime $p \in A$, the splitting type of C at p is the same as the splitting type of ϕ modulo p . In other words, the prime ideals lying above p in C can be put in bijection with the distinct linear factors of ϕ in such a way that the inertia and ramification indices, on the one hand, equal the degrees and multiplicities on the other. This can be proved using the fact that all maximal cubic \mathbb{Z}_p -algebras are monogenic, except $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ which is directly seen to correspond to $\phi(x, y) = xy(x + y)$.

3.1. \mathbb{Z} -mat rings

Just as a quadratic form can be represented by a symmetric matrix, a binary cubic form ϕ can be represented by a triply symmetric cubical box

$$\begin{array}{ccccc}
 & a & \text{---} & b/3 & \\
 & | & \diagdown & & \diagup \\
 & & b/3 & \text{---} & c/3 \\
 b/3 & | & & | & \\
 & & c/3 & \text{---} & c/3 \\
 & \diagup & & \diagdown & \\
 & c/3 & \text{---} & d &
 \end{array} \tag{6}$$

that has integer entries exactly when $3|b$ and $3|c$, in which case we call ϕ an *integer-matrix form*, or a \mathbb{Z} -mat form for short. It is not hard to see that this property is $\text{GL}_2\mathbb{Z}$ -invariant. The following proposition shows the link with \mathbb{Z} -mat rings as we previously defined them.

Proposition 3.2. *Let C be a cubic ring. The following are equivalent:*

- (a) *The cubic form corresponding to C is \mathbb{Z} -mat;*
- (b) *The trace of every element of C is a multiple of 3;*
- (c) *$C = \mathbb{Z} \oplus C^0$, where $C^0 \subseteq C$ is the subgroup of elements having trace zero.*

Proof. The equivalence (b) \Leftrightarrow (c) is straightforward. For (a) \Leftrightarrow (b), write the multiplication table of C in terms of a normal basis:

$$\alpha^2 = -ac - b\alpha + a\beta$$

$$\alpha\beta = -ad$$

$$\beta^2 = -bd - d\alpha + c\beta.$$

The trace of α may of course be computed by adding the coefficients of x in αx as x runs over the basis $[1, \alpha, \beta]$. Since $\alpha \cdot 1$ has no constant term and $\alpha\beta$ has no β term, we get $\text{tr } \alpha = -b$, and likewise $\text{tr } \beta = c$. So the traces of all elements of C are multiples of 3 if and only if $3|b$ and $3|c$, i.e. the corresponding form is \mathbb{Z} -mat. \square

3.2. The maximal \mathbb{Z} -mat subring

It is well known that every nondegenerate cubic ring C sits in a unique maximal cubic ring, namely the integral closure of \mathbb{Z} in the corresponding \mathbb{Q} -algebra $K = C \otimes_{\mathbb{Z}} \mathbb{Q}$. The corresponding theorem for \mathbb{Z} -mat rings is also true.

Proposition 3.3. *Let C be a cubic ring. The family of \mathbb{Z} -mat rings lying in C has a unique maximal element $C^{\mathbb{Z}\text{m}}$ in which all others are contained.*

Proof. A first guess would be to let $C^{\mathbb{Z}m}$ be the set of elements of C whose trace is divisible by 3, but these do *not* in general form a ring. Instead, let

$$C^{\mathbb{Z}m} = \{x \in C \mid x^3 \in \mathbb{Z} + 3C\}.$$

We verify the three desired properties:

1. $C^{\mathbb{Z}m}$ is a ring. Clearly C contains the integers and is closed under multiplication. If $x, y \in C$, then

$$(x + y)^3 = x^3 + y^3 + 3(x^2y + xy^2) \in \mathbb{Z} + 3C$$

so $x + y \in C^{\mathbb{Z}m}$.

2. $C^{\mathbb{Z}m}$ is \mathbb{Z} -mat. Given $x \in C$, pick $n \in \mathbb{Z}$ such that $x^3 \equiv n \pmod{3C}$; then $(x - n)^3 \equiv 0 \pmod{3C}$. On the $\mathbb{Z}/3\mathbb{Z}$ -module $C/3C$, the multiplier $x - n$ acts nilpotently and thus has trace zero. Thus $3 \mid \text{tr}(x - n)$, and thus $3 \mid \text{tr } x$.
3. Any \mathbb{Z} -mat subring of C lies in $C^{\mathbb{Z}m}$. If x lies in a \mathbb{Z} -mat subring, then $3 \mid \text{tr } x$ and also $3 \mid \text{tr } x^2$. Thus the characteristic polynomial of x modulo 3 has the form $t^3 - n$, so $x^3 \equiv n \pmod{3C}$ and hence $x \in C^{\mathbb{Z}m}$. \square

If C is any nondegenerate \mathbb{Z} -mat ring, then there is a largest \mathbb{Z} -mat ring containing C , namely $C_0^{\mathbb{Z}m}$, where C_0 is the maximal cubic ring containing C . We call $C_0^{\mathbb{Z}m}$ a *maximal \mathbb{Z} -mat ring*, to be distinguished from a \mathbb{Z} -mat maximal ring (that is, a maximal ring that is \mathbb{Z} -mat).

Although we have worked for convenience only over \mathbb{Z} , the foregoing theory of \mathbb{Z} -mat rings is applicable without change over \mathbb{Z}_3 . (Of course, if $p \neq 3$, every cubic ring over \mathbb{Z}_p is \mathbb{Z} -mat, or shall I say \mathbb{Z}_p -mat?)

4. Reducing to the case that D has no high prime powers

For the first section of our proof, we will tackle a step that occupies the last section of Nakagawa's treatment: eliminating all D with high prime power factors by means of a recursion that expresses both $h(D)$ and $\hat{h}(D)$ in terms of simpler discriminants.

The main result of this section is as follows:

Theorem 4.1. *For all $D \in \text{Discs}$ and all primes p ,*

$$h(p^6 D) = h(p^4 D) + p \cdot (h(D) - h(D/p^2)) \tag{7}$$

$$\hat{h}(p^6 D) = \hat{h}(p^4 D) + p \cdot (\hat{h}(D) - \hat{h}(D/p^2)), \tag{8}$$

using the natural convention that $h(a) = \hat{h}(a) = 0$ for all $a \notin \text{Discs}$.

Remark 4.2. If $D = D_0 f^2$ with D_0 fundamental, and if $p^3 | f$ for some prime p , then this result proves [Theorem 1.1](#) for us in the case $\Delta = D$, given the cases $\Delta = D/p^2$, D/p^6 , and D/p^8 . Inducting on $|\Delta|$ (all cases with $\Delta \notin \mathcal{Discs}$ being trivial) allows us to assume that $\Delta = \Delta_0 f^3$ with Δ_0 fundamental and f cubefree in [Theorem 1.1](#).

Proof of Theorem 4.1. We prove more strongly that for each cubic algebra C_1 over \mathbb{Z} that is maximal (resp. maximal \mathbb{Z} -mat) at p , the contributions to the left and right sides of (7) (resp. (8)) coming from subrings $C \subseteq C_1$ of p -power index are equal. Here is the first of many times that the $1/|\text{Aut } C|$ weighting in [Theorem 1.1](#) is to our advantage: since every automorphism of such a C lifts to an automorphism of C_1 , we have the identity

$$\frac{1}{|\text{Aut } C|} = \frac{|\{C' \subseteq C_1 : C \cong C'\}|}{|\text{Aut } C_1|}$$

and we can simply count subrings of C_1 without worrying whether they are isomorphic or have automorphisms. (If C_1 is \mathbb{Z} -mat, all its finite-index subrings will also be \mathbb{Z} -mat, by definition.)

The enumeration of subrings of a fixed ring is a local problem, and without further ado we will let C_1 denote a maximal (resp. maximal \mathbb{Z} -mat) nondegenerate cubic algebra over \mathbb{Z}_p and s_n the number of subrings of C_1 of index p^n . In particular $s_0 = 1$ and $\dots = s_{-2} = s_{-1} = 0$. It suffices to prove the recursion

$$s_{n+3} = s_{n+2} + p(s_n - s_{n-1}) \quad (9)$$

for all n that are big enough that

$$p^{2n} \text{Disc } C_1 \in \mathcal{Discs}_p, \quad \text{resp.} \quad -\frac{1}{27} \cdot p^{2n} \text{Disc } C_1 \in \mathcal{Discs}_p, \quad (10)$$

recalling that \mathcal{Discs}_p consists of the p -adic integers congruent to 0 or 1 mod 4 (this condition being vacuous unless $p = 2$). Clearly all $n \geq 0$ satisfy (10); we will discover that $n = -1$ and -2 sometimes do, and $n \leq -3$ never does (thankfully, as (9) is clearly false for $n = -3$).

If $C \subseteq C_1$ is a subring of index p^n , then C_1/C is a quotient group of $C/\mathbb{Z}_p \cong \mathbb{Z}_p \oplus \mathbb{Z}_p$ and thus has at most two elementary divisors. Write

$$C_1/C \cong \mathbb{Z}/p^i \mathbb{Z} \oplus \mathbb{Z}/p^j \mathbb{Z}$$

where $0 \leq i \leq j$ are integers with $i + j = n$. Using this isomorphism, we get a normal basis $[1, \alpha, \beta]$ for C_1 such that $[1, p^i \alpha, p^j \beta]$ is a basis for C , manifestly also normal. One then computes that if

$$\phi_0(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$$

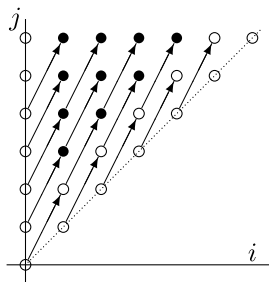


Fig. 1. Valid (i, j) pairs. Solid dots indicate where the contribution to the number of subrings can be computed by multiplying by p the number coming from $(i - 1, j - 2)$.

is the index form attached to C_1 in the basis $[1, \alpha, \beta]$, then the corresponding index form attached to C is

$$\phi_C(x, y) = ap^{2i-j}x^3 + bp^ix^2y + cp^jxy^2 + dp^{2j-i}. \quad (11)$$

In particular, if $2i \geq j$, then this form has integer coefficients and so C will be a ring no matter what normal basis $[1, \alpha, \beta]$ we pick. Otherwise we must impose the condition that $a = \phi(\bar{\alpha})$ is divisible by p^{j-2i} .

Of course, different normal bases $[1, \alpha, \beta]$, or equivalently, different bases $[\bar{\alpha}, \bar{\beta}]$ for the lattice $L_1 = C_1/\mathbb{Z}_p$, may yield the same ring C , which is determined by the lattice

$$L_C = p^{-i}(C/\mathbb{Z}_p) = \langle \bar{\alpha}, p^{j-i}\bar{\beta} \rangle = p^{j-i}L_1 + \langle \bar{\alpha} \rangle.$$

In particular, the vector β is immaterial, and α may range over all vectors of L_1 not divisible by p , up to translation by $p^{j-i}L_1$ and scaling by units. In other words, the parameter space for α is the finite projective line $\mathbb{P}^1(\mathbb{Z}/p^{j-i}\mathbb{Z})$, and s_n is the total number of solutions to

$$\phi(x, y) \equiv 0 \pmod{p^{\max\{j-2i, 0\}}} \quad (12)$$

for $[x : y] \in \mathbb{P}^1(\mathbb{Z}/p^{j-i}\mathbb{Z})$, where (i, j) ranges over integer pairs with $0 \leq i \leq j$ and $i + j = n$.

The key point to note is that replacing (i, j) with $(i + 1, j + 2)$ does not change the condition (12), but gives us a projective line $\mathbb{P}^1(\mathbb{Z}/p^{j-i}\mathbb{Z})$ with p points lying over every point that was there before. We get $s_n \approx ps_{n-3}$, subject to three corrective terms (compare Fig. 1):

- When $j = i + 1$, $\mathbb{P}^1(\mathbb{Z}/p\mathbb{Z})$ has $p + 1$ points instead of p , contributing an extra point for $n \geq 3$ odd;
- When $i = j$, the pair (i, i) is inaccessible by this translation and contributes 1 ($= |\mathbb{P}^1(\mathbb{Z}/1\mathbb{Z})|$) extra point for n even;

- The pair $(i, j) = (0, n)$ is also inaccessible by this translation and contributes r_n points, where r_n is the number of solutions to $\phi(x, y) \equiv 0 \pmod{p^n}$ in $\mathbb{P}^1(\mathbb{Z}/p^n\mathbb{Z})$.

Thus, for $n \geq 2$,

$$s_n = ps_{n-3} + 1 + r_n,$$

and in particular, for $n \geq 0$,

$$s_{n+3} = s_{n+2} - p(s_n - s_{n-1}) + r_{n+3} - r_{n+2}.$$

Thus proving the desired recursion (9) for $n \geq 0$ is equivalent to showing that r_m is constant for $m \geq 2$. For large m this follows from a suitably strong version of Hensel's lemma; in our situation, some remarkable circumstances converge to give the results for the n that we desire.

We also have $s_2 = 1 + r_2$ and $s_1 = r_1$ by a direct determination of the (i, j) pairs involved. Hence

$$(9) \text{ holds for } n = -1 \iff r_2 = r_1 - 1 \quad (13)$$

$$(9) \text{ holds for } n = -2 \iff r_1 = 1. \quad (14)$$

Suppose first that C_1 is maximal. Let

$$\phi_0(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$$

be its associated cubic form. Suppose that we are given a root of ϕ_0 in $\mathbb{P}^1(\mathbb{Z}/p^2\mathbb{Z})$; choose our basis $[1, \alpha, \beta]$ of C_1 so that it is at $[1 : 0]$, so $p^2|a$. If $p|b$, then applying the formula (11) with $i = -1$ and $j = 0$ shows that $\langle 1, p^{-1}\alpha, \beta \rangle$ is a ring, contradicting the maximality of C . So $[1 : 0]$ is a simple root and thus has a unique lift mod all p^m by Hensel's lemma, proving (9) for $n \geq 0$.

The cases for $n < 0$ only pop up when $p|\text{Disc } C_1$, that is, C_1 is ramified. This can happen either when $C_1 \cong \mathbb{Z}_p \times \mathcal{O}_{K_2}$, where K_2 is a ramified quadratic extension of \mathbb{Z}_p , or $C_1 = \mathcal{O}_{K_3}$ where K_3 is a totally ramified cubic extension of \mathbb{Z}_p . But in the former case, $\text{Disc } C_1 = \text{Disc } K_2$ is fundamental, so we still only have to prove $n \geq 0$.

In the totally ramified cubic case, we have $\text{Disc } C_1 \leq 5$ by the Dedekind–Hensel bound (which in general says that for L/K an extension of local fields, $v_K(\text{Disc}_K L) \leq e - 1 + ev_K(e)$ where $e = e_{L/K}$ is the ramification index). So $n \geq -2$. Mod p , ϕ_0 has a single root of multiplicity 3 (because the splitting type of C is 1^3); mod p^2 , ϕ_0 has no roots, or else C_1 would be non-maximal as was just shown. So (13) and (14) both hold, which shows (9) for $n = -1$ and -2 .

This completes the proof of (9) for C_1 maximal, and thus also the proof of (7). There remains the case that $p = 3$ and C_1 is the maximal \mathbb{Z} -mat subring in a maximal ring C_0

that is not \mathbb{Z} -mat. Note that we are now proving (8), so n is governed by the stronger inequality

$$-\frac{1}{27} \cdot 3^{2n} \text{Disc } C_1 \in \mathbb{Z}_3,$$

that is,

$$2n - 3 + v_3(\text{Disc } C_1) \geq 0.$$

Note also that $[C_0 : C_1]$ is either 3 or 9 since $\mathbb{Z}_3 + 3C_0 \subseteq C_1$.

Consider first the case that $[C_0 : C_1] = 9$, that is, $C_1 = \mathbb{Z}_3 + 3C_0$. Note that C_0 must be unramified since otherwise there is an element $\xi \notin \mathbb{Z} + 3C_0$ whose cube lies in $3C_0$, contradicting the construction of the maximal \mathbb{Z} -mat subring. So $p \nmid \text{Disc } C_0$, yielding $v_3(\text{Disc } C_1) = 4$ and $n \geq 0$. Now the form ϕ_0 corresponding to C_1 is 3 times the form ϕ_1 corresponding to C_0 (by (11) with $i = j = 1$) and so r_m ($m \geq 2$) is simply 3 times the number of roots of

$$\phi_1(x, y) \equiv 0 \pmod{3^{m-1}}$$

on $\mathbb{P}^1(\mathbb{Z}/3^{m-1}\mathbb{Z})$, which is constant for $m \geq 2$ by Hensel's lemma.

In the case $[C_0 : C_1] = 3$, the relationship between the corresponding forms ϕ_0 and ϕ_1 is governed by (11) with $i = 1$ and $j = 0$, so we can write

$$\phi_0(x, y) = 9a'x^3 + 3b'x^2y + 3c'xy^2 + dy^3$$

$$\phi_1(x, y) = a'x^3 + b'x^2y + 3c'xy^2 + 3dy^3$$

where $a', b', c', d \in \mathbb{Z}_3$. Note that $3 \nmid d$ since C_0 is maximal. So the only root of $\phi_0 \pmod{3}$ is $[1 : 0]$, and the roots $\pmod{3^m}$ must be expressible in the form $[1 : 3y']$, $y' \in \mathbb{Z}/3^{m-1}\mathbb{Z}$. Note that

$$\phi_0(1, 3y') \equiv 0 \pmod{3^m} \iff \phi_1(1, y') \equiv 0 \pmod{3^{m-1}}.$$

Now $[0 : 1]$ is a root of $\phi_1 \pmod{3}$, with multiplicity exactly 2: we have $3 \nmid b'$ since C_0 is not \mathbb{Z} -mat. So there is a single simple root of the form $[1 : y]$ modulo 3. By Hensel's lemma, there is a single root of this form modulo all higher powers of 3, yielding 3 roots of ϕ_0 modulo 3^m for all $m \geq 2$. This proves (9) for all $n \geq 0$, which is all that is needed: for we have shown that C_0 has splitting type 1^21 and so $v_3(\text{disc } C_1) = 3$. \square

Remark 4.3. This proof also shows that, if C_1 is maximal, the initial terms s_1, s_2 of the recursion can be computed using only the splitting type σ of ϕ_0 at p : s_1 is the number of roots \pmod{p} , and s_2 is 1 plus the number of *simple* roots \pmod{p} , as these are the only ones that lift to $\pmod{p^2}$. The values of these numbers are tabulated below for future reference.

$$\begin{array}{c|ccccc}
 \sigma & 111 & 12 & 3 & 1^2 1 & 1^3 \\
 \hline
 s_1 & 3 & 1 & 0 & 2 & 1 \\
 \hline
 s_2 & 4 & 2 & 1 & 2 & 1
 \end{array} \tag{15}$$

Together with $s_0 = 1$ and $s_{-1} = 0$, they enable the computation of the number of subrings of any index of a maximal cubic ring over \mathbb{Z}_p (or, indeed, over \mathbb{Z}).

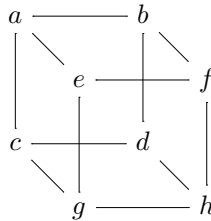
Incidentally, the recursion (9) can be solved explicitly to get a formula

$$s_n = \frac{p^{\lfloor \frac{n+3}{3} \rfloor} - 1 + (s_1 - 1) \left(p^{\lfloor \frac{n+2}{3} \rfloor} - 1 \right) + (s_2 - s_1) \left(p^{\lfloor \frac{n+1}{3} \rfloor} - 1 \right)}{p - 1}$$

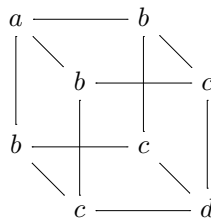
(cf. [10], Lemma 3.7), but this will be less useful to us.

5. \hat{h} and self-balanced ideals

Many readers will no doubt have seen Bhargava’s dazzling reinterpretation of Gauss’s 200-year-old composition law on binary quadratic forms [1]: a cube



corresponds to a triple of quadratic forms whose Gauss composite is 0, or more generally to three fractional ideals of a quadratic order that are “balanced,” meaning that their product is nearly the unit ideal in a suitably defined sense. Here, our focus is on the triply symmetric cubes



which we have already mentioned as the natural pictorial avatars of \mathbb{Z} -mat cubic forms. Due to the symmetry, these cubes correspond in Bhargava’s bijection to “balanced” triples consisting of three ideals in the *same* class; only this latter bijection need be described in detail here.

Definition 5.1. A *self-balanced triple* is a triple (\mathcal{O}, I, γ) , where \mathcal{O} is an order in a quadratic \mathbb{Q} -algebra K_2 , I is a fractional ideal of \mathcal{O} , and $\gamma \in K_2^\times$ is a scalar, satisfying the two conditions

$$\gamma I^3 \subseteq \mathcal{O} \quad (16)$$

$$|N(\gamma)| \cdot N(I)^3 = 1. \quad (17)$$

Also define an equivalence relation on self-balanced triples by

$$(\mathcal{O}, I, \gamma) \sim (\mathcal{O}, \lambda I, \lambda^{-3} \gamma)$$

for every $\lambda \in K_2^\times$. (It is immediate that the second triple is self-balanced if the first is.)

Recall that an *oriented* cubic ring C is one with a distinguished generator $\omega_C \in \Lambda^3 C$, enabling us to view its index form $\phi : C/\mathbb{Z} \rightarrow \Lambda^3 C$ as taking values in \mathbb{Z} . A cubic ring C can be oriented in two ways, which are isomorphic if and only if C has an orientation-reversing automorphism; thus there are precisely $2\hat{h}(\Delta)$ oriented \mathbb{Z} -mat cubic rings of discriminant -27Δ , if we weight by the reciprocal of the number of *oriented* automorphisms.

We are now ready to state the pertinent bijection.

Theorem 5.2 (cf. [1], Theorem 13). *Oriented \mathbb{Z} -mat cubic rings C of discriminant $-27\Delta \neq 0$ are in bijection with equivalence classes of self-balanced triples of the quadratic order \mathcal{O}_Δ of discriminant Δ . Also, those C having a nontrivial oriented automorphism, necessarily of order 3, correspond to those equivalence classes having a representative*

$$(\mathcal{O}_\Delta, \mathbb{Z}[\omega], \gamma),$$

where $\mathbb{Z}[\omega]$ is the unit ideal in the ring generated by a primitive 3rd root of unity (clearly Δ must be -3 times a square for this to happen).

Proof. For a hands-on proof (that also works when $\text{Disc } C = 0$), see [1]. Here we present a new proof based on that most ancient nexus between quadratic and cubic number fields: the Tartaglia–Cardano cubic formula.

Let C be a nondegenerate oriented \mathbb{Z} -mat cubic ring. By Proposition 3.2(c), $C = \mathbb{Z} \oplus C^0$, where C^0 is the sublattice of elements of trace 0.

The following observation is worth recording: For any $\alpha \in C^0$, the characteristic polynomial of α has the form $\alpha^3 + 3t\alpha + u$, where

$$t = \frac{1}{6} \text{tr}(\alpha^2) \in \mathbb{Z}. \quad (18)$$

This is because, on the one hand, $3t \in \mathbb{Z}$ by the integrality of the characteristic polynomial, but also $2t = \text{tr}(\alpha^2)/3 \in \mathbb{Z}$ because C is \mathbb{Z} -mat.

Pick a generic element $\alpha \in C^0$; specifically, we should have that

- $1 \wedge \alpha \wedge \alpha^2 \neq 0$, that is, $[1, \alpha, \alpha^2]$ is a \mathbb{Q} -basis of $C \otimes \mathbb{Q}$; and
- $\text{tr}(\alpha^2) \neq 0$, for reasons that will soon be clear.

Using the nondegeneracy of C , these conditions are not hard to fulfill. They are also homogeneous, and there is no harm in taking α a *primitive* element, that is, one such that $\mathbb{Q}\alpha \cap C = \mathbb{Z}\alpha$.

Then α has characteristic polynomial $\alpha^3 + 3t\alpha + u = 0$ for some $t, u \in \mathbb{Z}$. We can now “solve” for α using the Tartaglia–Cardano formula:

$$\alpha = \sqrt[3]{\gamma} + \sqrt[3]{\bar{\gamma}}, \quad (19)$$

where

$$\gamma = \frac{-u + \sqrt{u^2 + 4t^3}}{2} \quad \text{and} \quad \bar{\gamma} = \frac{-u - \sqrt{u^2 + 4t^3}}{2}.$$

If C admits an embedding into \mathbb{C} , this is literally true, provided that we choose the cube roots such that their product is t . In general, we can interpret the expression as follows. First note that the polynomial $x^3 + 3tx + u$ has discriminant $-27(u^2 + 4t^3)$, whence

$$-27(u^2 + 4t^3) = \text{Disc } \mathbb{Z}[\alpha] = [C : \mathbb{Z}[\alpha]]^2 \cdot \text{Disc } C = \phi(\alpha)^2 \cdot (-27\Delta),$$

where $\phi(\xi) = \frac{1}{\omega_C} 1 \wedge \xi \wedge \xi^2$ is the index form of C . Thus we can view $\sqrt{u^2 + 4t^3} = \phi(\alpha)\sqrt{\Delta}$, and hence γ and $\bar{\gamma}$, as elements of the nondegenerate quadratic algebra $K_2 = \mathbb{Q}[\sqrt{\Delta}]$ canonically associated to C . Then in the sextic algebra $K_6 = K_2[\sqrt[3]{\gamma}]$, $\sqrt[3]{\gamma}$ is invertible (because $\gamma\bar{\gamma} = t^3$ is invertible) and the element $\sqrt[3]{\gamma} = t/\sqrt[3]{\bar{\gamma}}$ is a cube root of $\bar{\gamma}$. Then, by the usual derivation of the cubic formula, $\alpha \mapsto \sqrt[3]{\gamma} + \sqrt[3]{\bar{\gamma}}$ identifies C with a cubic subring of K_6 .

We have

$$(\sqrt[3]{\gamma})^2 = \frac{1}{t} (\sqrt[3]{\gamma})^3 \sqrt[3]{\gamma} = \frac{1}{t} \gamma \sqrt[3]{\gamma},$$

so

$$\alpha^2 = (\sqrt[3]{\gamma})^2 + 2\sqrt[3]{\gamma}\sqrt[3]{\bar{\gamma}} + \left(\sqrt[3]{\bar{\gamma}}\right)^2 = 2t + \frac{1}{t}(\bar{\gamma}\sqrt[3]{\gamma} + \gamma\sqrt[3]{\bar{\gamma}}),$$

and since $[1, \alpha, \alpha^2]$ is a \mathbb{Q} -basis of $C \otimes_{\mathbb{Z}} \mathbb{Q}$, we see that

$$C \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q} \oplus \{\xi \sqrt[3]{\gamma} + \bar{\xi} \sqrt[3]{\bar{\gamma}} \mid \xi \in K_2\}$$

and hence

$$C = \mathbb{Z} \oplus C^0 \cong \mathbb{Z} \oplus \{\xi \sqrt[3]{\gamma} + \bar{\xi} \sqrt[3]{\bar{\gamma}} \mid \xi \in I\} \quad (20)$$

for some lattice $I \subset K_2$. We claim that these choices of I and γ make $(\mathcal{O}_\Delta, I, \gamma)$ a self-balanced triple.

For brevity we write $c(\xi) = \xi \sqrt[3]{\gamma} + \bar{\xi} \sqrt[3]{\bar{\gamma}}$, so $c : K_2 \rightarrow C^0 \otimes \mathbb{Q}$ is an isomorphism of \mathbb{Q} -vector spaces.

Note that 1 is a primitive vector in I , so I has a basis $[1, \tau]$ where

$$\tau = \frac{s + \sqrt{\Delta}}{q}$$

for some $s, q \in \mathbb{Q}$, and C has a basis $[1, c(1), c(\tau)] = [1, \alpha, c(\tau)]$. Let us choose the sign of τ such that the distinguished generator $1 \wedge \alpha \wedge c(\tau)$ of $\Lambda^3 C$ is the given ω_C . Then

$$\begin{aligned} \phi(\alpha) &= \frac{1}{\omega_C} 1 \wedge \alpha \wedge \alpha^2 \\ &= \frac{1}{\omega_C} 1 \wedge \alpha \wedge \left(2t + c\left(\frac{\bar{\gamma}}{t}\right) \right) \\ &= \frac{1}{\omega_C} 1 \wedge \alpha \wedge c\left(\frac{-\sqrt{u^2 + 4t^3}}{2t}\right) \\ &= \frac{1 \wedge \alpha \wedge \frac{-\phi(\alpha)c(\sqrt{\Delta})}{2t}}{1 \wedge \alpha \wedge \frac{c(\sqrt{\Delta})}{q}} \\ &= \frac{-q\phi(\alpha)}{2t}, \end{aligned}$$

that is, $q = -2t$.

The multiplication law on C is given by

$$c(\xi)c(\eta) = t(\xi\bar{\eta} + \bar{\xi}\eta) + c\left(\frac{\bar{\gamma}\xi\bar{\eta}}{t}\right); \quad (21)$$

hence the conditions for C to be a ring are that

$$t(\xi\bar{\eta} + \bar{\xi}\eta) \in \mathbb{Z} \quad (22)$$

and

$$\frac{1}{t}\bar{\xi}\bar{\eta}\bar{\gamma} \in I \quad (23)$$

for all $\xi, \eta \in I$. Plugging $\xi = 1, \eta = \tau$ in (22) yields $s \in \mathbb{Z}$. Plugging $\xi = \eta = \tau$ in (22) would yield $2t\tau\bar{\tau} \in \mathbb{Z}$, but we can get a stronger result by applying the trace relation (18) to the element $c(\tau)$, yielding that

$$r := t\tau\bar{\tau} = \frac{s^2 - \Delta}{4t} \in \mathbb{Z}.$$

Consequently $s \equiv \Delta \pmod{2}$, and the multiplier $(s - \Delta)/2$, which generates the order \mathcal{O}_Δ , takes τ to an integer $-r$. This shows that I is an ideal of \mathcal{O}_Δ .

Condition (17) is immediate, as I has norm $1/|t|$.

We must now prove (16), namely that $\gamma I^3 \subseteq \mathcal{O}_\Delta$. Since $\gamma I^2 \subseteq t\bar{I}$ by (23), it suffices to prove that $tI\bar{I} \subseteq \mathcal{O}_\Delta$. But using the known \mathbb{Z} -basis,

$$\begin{aligned} tI\bar{I} &= \langle t, t\tau, t\bar{\tau}, t\tau\bar{\tau} \rangle \\ &= \left\langle t, \frac{s + \sqrt{\Delta}}{-2}, \frac{s - \sqrt{\Delta}}{-2}, \frac{s^2 - \Delta}{2t} \right\rangle \\ &= \left\langle t, s, r, \frac{s + \sqrt{\Delta}}{2} \right\rangle \end{aligned}$$

which clearly lies in \mathcal{O}_Δ . This completes the construction of a self-balanced triple $(\mathcal{O}_\Delta, I, \gamma)$ corresponding to C .

We must still show that the choice of α made at the outset does not change the self-balanced triple, up to equivalence, derived from a cubic ring C . Suppose $(\mathcal{O}_\Delta, I, \gamma)$ and $(\mathcal{O}_\Delta, I', \gamma')$ both arose from this method, which also provides identifications of oriented \mathbb{Z} -modules $c : I \rightarrow C^0$, $c' : I' \rightarrow C^0$ and, in particular, an isomorphism $\psi = c'^{-1} \circ c : I \rightarrow I'$.

Plugging $\eta = \xi$ into (21), we see that

$$\frac{1}{6} \operatorname{tr}(c(\xi)^2) = t\xi\bar{\xi} = tN(\xi).$$

Thus $N(\psi(\xi))/N(\xi)$ is a constant t'/t for all ξ (where t' is the value of t corresponding to placing α' in place of α). In particular, $\psi(1)$ is invertible in K_2 , and the normalized map

$$\tilde{\psi}(\xi) = \frac{\psi(\xi)}{\psi(1)}$$

extends linearly to a \mathbb{Q} -linear self-map of K_2 that preserves 1 and norms. There are only two such, the identity and conjugation, and the latter is ruled out by the fact that ψ respects orientation. So ψ is a scaling $\xi \mapsto \lambda\xi$ for all ξ , and using the full multiplication law (21), it is easy to see that $\gamma' = \lambda^{-3}\gamma$ so the two self-balanced triples are equivalent.

By the same argument, a nontrivial oriented automorphism of C arises if and only if the associated balanced triple (\mathcal{O}, I, γ) is equivalent to itself via scaling by some multiplier $\lambda \neq 1$. To leave γ fixed, we must have $\lambda^3 = 1$ and so I is an ideal of the order $\mathbb{Z}[\omega]$. Since $\mathbb{Z}[\omega]$ is a PID, this implies that $I = \mathbb{Z}[\omega]$ up to scaling, as stated. Conversely, if $I = \mathbb{Z}[\omega]$, the map $c(\xi) \mapsto c(\omega\xi)$ clearly defines a nontrivial automorphism of C .

Conversely, given a self-balanced triple $(\mathcal{O}_\Delta, I, \gamma)$, we scale I so that it contains 1 as a primitive element (and scale γ appropriately). Let $t \in \mathbb{Z}$ be determined by $N(I) = 1/|t|$ and $\text{sgn } t = \text{sgn } N(\gamma)$. Then I has a basis

$$\left[1, \tau = \frac{s + \sqrt{\Delta}}{-2t} \right]$$

for some $s \in \mathbb{Z}$ of the same parity as Δ . We get from (20) a cubic ring C with a distinguished element $\alpha = c(1)$ for which the foregoing process returns the given triple $(\mathcal{O}_\Delta, I, \gamma)$, if we can prove that (22) and (23) hold. The verification of (22) devolves into the fact that the norm of an element of I is a multiple of $N(I) = 1/|t|$, since

$$t(\xi\bar{\eta} + \eta\bar{\xi}) = t[N(\xi + \eta) - N(\xi) - N(\eta)].$$

As for (23), it is convenient to use the identity

$$\frac{\xi \wedge \eta}{1 \wedge \tau} = \frac{\bar{\xi}\eta - \xi\bar{\eta}}{\tau - \bar{\tau}},$$

which may be proved merely by noting that $\bar{\xi}\eta - \xi\bar{\eta}$ is a \mathbb{Q} -linear, $\mathbb{Q}\cdot\sqrt{\Delta}$ -valued alternating 2-form on K_2 . Note that an element $\xi \in K_2$ belongs to I if and only if $\xi \wedge \eta \in \Lambda^2 I$ for every $\eta \in I$. Now for every $\xi, \eta, \zeta \in I$,

$$\frac{\bar{\gamma}\bar{\xi}\bar{\eta}}{1 \wedge \tau} \wedge \zeta = \frac{\gamma\xi\eta\zeta - \gamma\bar{\xi}\bar{\eta}\zeta}{t(\tau - \bar{\tau})} = \frac{\gamma\xi\eta\zeta - \gamma\bar{\xi}\bar{\eta}\zeta}{-\sqrt{\Delta}} \in \mathbb{Z}$$

since $\gamma\xi\eta\zeta \in \mathcal{O}_\Delta$, proving (23).

To show that the C corresponding to a self-balanced triple $(\mathcal{O}_\Delta, I, \gamma)$ is unique, it suffices to express the index form of C in terms of the triple, which is not difficult:

$$\begin{aligned} \phi(c(\xi)) &= \frac{1}{\omega_C} 1 \wedge c(\xi) \wedge c(\xi^2) \\ &= \frac{1}{\omega_C} 1 \wedge c(\xi) \wedge c\left(\frac{\bar{\xi}^2\gamma}{t}\right) \\ &= \frac{\xi \wedge \frac{\bar{\xi}^2\gamma}{t}}{1 \wedge \tau} \\ &= \frac{\xi^3\gamma - \bar{\xi}^3\gamma}{\sqrt{\Delta}}. \quad \square \end{aligned} \tag{24}$$

Here ends our proof of Bhargava's Theorem 13, but for our purposes, a slightly transformed description of the parametrization is preferable. The ideal I may or may not be invertible in \mathcal{O}_Δ . Indeed, with respect to a basis

$$I = \left\langle 1, \frac{s + \sqrt{\Delta}}{2t} \right\rangle,$$

we found that

$$tI\bar{I} = \left\langle t, s, r, \frac{s + \sqrt{\Delta}}{2} \right\rangle,$$

where $r = (s^2 - \Delta)/(4t)$ is an integer. If t , s , and r are relatively prime (incidentally, they are the coefficients of the *quadratic form* $tx^2 + sxy + ry^2$ of discriminant Δ associated to the class of I), then $tI\bar{I} = \mathcal{O}_\Delta$ and so I is invertible. However, in general, there may be a common factor $g = \gcd(t, s, r)$, and then one verifies that I is an ideal of the order $\mathcal{O}_{\Delta'}$, $\Delta' = \Delta/g^2$, with inverse $\bar{I}/(gt)$. Note that γI^3 is an $\mathcal{O}_{\Delta'}$ -ideal contained in \mathcal{O}_Δ . We need a little lemma about such ideals:

Lemma 5.3. *Let $\Delta \in \mathcal{Discs}$ and $g \geq 1$. An ideal I of \mathcal{O}_Δ that is contained in $\mathcal{O}_{\Delta g^2}$ is actually contained in $g\mathcal{O}_\Delta$.*

Proof. Let $\mathcal{O}_\Delta = \mathbb{Z}[\xi]$, so $\mathcal{O}_{\Delta g^2} = \mathbb{Z}[g\xi]$. Suppose $\eta = a + bg\xi \in \mathcal{O}_{\Delta g^2}$ is an element of I . Then multiplying by the conjugate $\bar{\xi} = \text{tr } \xi - \xi \in \mathcal{O}_\Delta$, we get that

$$\bar{\xi}\eta = a\bar{\xi} + bgN(\xi) = [bgN(\xi) + a \text{tr } \xi] - a\xi$$

belongs to I , and hence to $\mathcal{O}_{\Delta g^2}$. So $g|a$, and thus $\eta \in g\mathcal{O}_\Delta$. \square

Thus

$$J = \frac{\gamma I^3}{g}$$

is an invertible integral ideal of $\mathcal{O}_{\Delta'}$ of norm

$$\begin{aligned} N(J) &= \frac{|N(\gamma)|N_{\mathcal{O}_{\Delta'}}(I)^3}{g^2} \\ &= \frac{|N(\gamma)|(g \cdot N_{\mathcal{O}_\Delta}(I))^3}{g^2} \\ &= g. \end{aligned}$$

Conversely, if J is an invertible ideal of $\mathcal{O}_{\Delta'}$ of norm g whose class in $\text{Pic } \mathcal{O}'_\Delta$ is a cube (a clearly necessary condition), then J will in general correspond to a number of self-balanced triples $(\mathcal{O}_\Delta, I, \gamma)$, where $\Delta = \Delta' g^2$. There are $|\text{Pic}(\mathcal{O}_\Delta)[3]|$ possibilities for the class of I , and for each I , the value of γ is determined only up to units, whereas we have $(\mathcal{O}_\Delta, I, \gamma) \sim (\mathcal{O}_\Delta, I, \gamma')$ only when γ/γ' is the *cube* of a unit, yielding a further

$|\mathcal{O}_{\Delta'}^\times/(\mathcal{O}_{\Delta'}^\times)^3|$ possibilities. An appeal to the structure of the unit groups of quadratic fields shows that

$$|\mathcal{O}_{\Delta'}^\times/(\mathcal{O}_{\Delta'}^\times)^3| = \begin{cases} 3 & \text{if } \Delta' = -3 \text{ or } \Delta' \text{ is a positive non-square} \\ 1 & \text{otherwise.} \end{cases}$$

The exceptional status of $\Delta' = -3$ is welcome, since these are precisely the cases where we must count the corresponding rings with weight $1/3$ owing to the nontrivial automorphism. We also get an extra factor of 3 for Δ' a positive non-square, in other words, for Δ a positive non-square. We summarize our findings as follows.

Theorem 5.4 ([10], Theorem 2.6 is the case $\Delta < 0$). *Let $w_\Delta = 3$ if Δ is a square, 1 otherwise. Also let $\sigma_\Delta = 1/3$ if Δ is positive, 1 if Δ is negative. The following quantities are equal:*

- $2w_\Delta\sigma_\Delta\hat{h}(\Delta)$;
- The number of invertible ideals J of norm g whose class is a cube in orders $\mathcal{O}_{\Delta'}$ for integers $g > 0$, Δ' satisfying $\Delta'g^2 = \Delta$, each counted with weight

$$|\text{Pic}(\mathcal{O}_{\Delta'})[3]|.$$

Remark 5.5. In Nakagawa's work [10], a more computational proof is given that centers on the fact that the quadratic form $tx^2 + sxy + ry^2$ attached to I is actually the *Hessian* of the cubic form attached to C , that is, the determinant of second partial derivatives, up to scaling. The same form appears in Mantilla-Soler [9] as the *trace form* of C , that is, the value of $\text{tr}(\alpha^2)/6$ as $\alpha \in C^0$ varies.

6. Interlude: Links with class field theory

We pause for a moment to consider how Theorem 1.1 is transformed under the elementary tools developed so far, and how in certain special cases one is led to the founding concerns of class field theory. We already have Theorem 5.4, which relates $\hat{h}(\Delta)$ to ideals in quadratic orders. Although it will not be used in the sequel, a comparable description of $h(\Delta)$ is not so hard to come by. For simplicity we treat only the case $3 \nmid \Delta$.

Proposition 6.1. *Let $3 \nmid \Delta$. To compute $6w_{-3\Delta}\sigma_{-3\Delta}h(\Delta)$, add the contributions to $2w_{-27\Delta}\sigma_{-3\Delta}\hat{h}(-27\Delta)$ in Theorem 5.4 for which $3 \nmid g$.*

Proof. We can make any cubic form \mathbb{Z} -mat by multiplying it by 3, that is, passing from the associated cubic ring C to the subring $\mathbb{Z} + 3C$. We now want to count \mathbb{Z} -mat cubic forms of discriminant 81Δ satisfying the additional condition $3|a$, $3|d$. Following this condition through the bijection of Theorem 5.2 shows that $2w_{-3\Delta}h(\Delta)$ is the number of inequivalent balanced triples $(\mathcal{O}_{-3\Delta}, I, \gamma)$ such that

$$\gamma\alpha^3 \in \mathcal{O}_{-27\Delta} \quad (25)$$

for each $\alpha \in I$.

Suppose $(\mathcal{O}_{-27\Delta}, I', \gamma)$ is a balanced triple such that I' is *not* an ideal of $\mathcal{O}_{-3\Delta}$, that is, the corresponding J in [Theorem 5.4](#) has $3 \nmid g$. Then $I = I'\mathcal{O}_{-3\Delta}$ is an ideal of index 3 over I . The triple $(\mathcal{O}_{-3\Delta}, I, \gamma)$ is clearly balanced, and any element $\alpha \in I$ can be written as $\kappa + \lambda\xi$, where $\kappa, \lambda \in I'$ and $\xi = \frac{-3\Delta + \sqrt{-3\Delta}}{2}$ is a generator of $\mathcal{O}_{-3\Delta}$; one checks that $\xi^3 \in 3\mathcal{O}_{-3\Delta} \subseteq \mathcal{O}_{-27\Delta}$, and thus

$$\gamma\alpha^3 = \gamma\kappa^3 + 3(\gamma\kappa^2\lambda\xi + \gamma\kappa\lambda^2\xi^2) + \gamma\lambda^3\xi^3 \in \mathcal{O}_{-27\Delta},$$

verifying (25). Conversely, if $(\mathcal{O}_{-3\Delta}, I, \gamma)$ is balanced and satisfies (25), then I has four sublattices I' of index 3, one of which is $\mathfrak{p}I$ (using that $3 = \mathfrak{p}^2$ ramifies in $\mathcal{O}_{-3\Delta}$). The other three are ideals of $\mathcal{O}_{-27\Delta}$ but not of $\mathcal{O}_{-3\Delta}$. Thus they yield triples $(\mathcal{O}_{-27\Delta}, I', \gamma)$ which are balanced since we can write $I' = 3I + \mathbb{Z}\alpha_0$ and get

$$\gamma I'^3 \subseteq \gamma\alpha_0^3\mathbb{Z} + 3\gamma II'^2 \subseteq \mathcal{O}_{-27\Delta} + 3\mathcal{O}_{-3\Delta} = \mathcal{O}_{-27\Delta}.$$

So we have a 3-to-1 correspondence between the balanced triples involved, establishing the desired identity. \square

We now present two examples showing the sorts of problems we encounter when tackling [Theorem 1.1](#) with both sides interpreted in this ideal-theoretic way.

Example 6.2. If $\Delta = \Delta_0$ is a fundamental discriminant, then only the terms with $g = 1$ count on either side, and [Theorem 1.1](#) devolves into

$$\frac{|\mathrm{Pic}(\mathcal{O}_{-27\Delta})[3]|}{|\mathrm{Pic}(\mathcal{O}_{\Delta})[3]|} = \begin{cases} 3, & \Delta > 1 \\ 1, & \Delta \leq 1. \end{cases}$$

Since there is a surjection $\mathrm{Pic}(\mathcal{O}_{-27\Delta}) \rightarrow \mathrm{Pic}(\mathcal{O}_{-3\Delta})$ whose kernel has size 1 or 3, we get a corollary concerning the class groups of quadratic number fields:

$$\frac{|\mathrm{Cl}(\mathbb{Q}(\sqrt{-3\Delta}))[3]|}{|\mathrm{Cl}(\mathbb{Q}(\sqrt{\Delta}))[3]|} = \begin{cases} 3 \text{ or } 1, & \Delta > 1 \\ 1 \text{ or } \frac{1}{3}, & \Delta < 0. \end{cases}$$

This is the *Scholz reflection principle*, proved by Scholz in 1932 as a stunning application of class field theory.

Example 6.3. For an example of a different flavor, take $\Delta = p^2q^2$, where p and q are primes with $p \equiv 1 \pmod{3}$, $q \equiv 2 \pmod{3}$. Then verifying [Theorem 1.1](#) reduces to counting ideals of various norms in suborders of $\mathbb{Z} \times \mathbb{Z}$ and $\mathbb{Z}[\omega]$ (where ω is a primitive cube

root of unity) and checking the cubicality of their classes in the Picard group, which is governed by the familiar exact sequence

$$0 \rightarrow \mathcal{O}^\times \rightarrow \mathcal{O}_K^\times \rightarrow \frac{(\mathcal{O}/\mathfrak{f})^\times}{(\mathcal{O}_K/\mathfrak{f})^\times} \rightarrow \text{Pic } \mathcal{O} \rightarrow \text{Pic } \mathcal{O}_K \rightarrow 0, \quad (26)$$

where K is any number field, \mathcal{O}_K its ring of integers, \mathcal{O} an order in K , and $\mathfrak{f} = \{x \in \mathcal{O}_K \mid x\mathcal{O}_K \subseteq \mathcal{O}\}$ its conductor ([11], Propositions I.1.9 and I.1.11; the proof extends readily to the case that K is a finite product of number fields). We present the outcomes here.

For $2\hat{h}(p^2q^2)$, we count:

- ideals of norm 1 in $\mathbb{Z} + pq(\mathbb{Z} \times \mathbb{Z})$: 1, with weight 3.
- ideals of norm pq in $\mathbb{Z} \times \mathbb{Z}$: 4, with weight 1.
- ideals of norm p in $\mathbb{Z} + q(\mathbb{Z} \times \mathbb{Z})$: 2, with weight 1.
- ideals of norm q in $\mathbb{Z} + p(\mathbb{Z} \times \mathbb{Z})$: here things become interesting. The Picard group is $(\mathbb{Z}/p\mathbb{Z})^\times$; there are 2 such ideals (with weight 3) if q is a cube mod p , and none otherwise.

For $6h(p^2q^2)$, we count:

- ideals of norm 1 in $\mathbb{Z}[3pq\omega]$: 1, with weight 9.
- ideals of norm pq in $\mathbb{Z}[3\omega]$: none.
- ideals of norm q in $\mathbb{Z}[3p\omega]$: none.
- ideals of norm p in $\mathbb{Z}[3q\omega]$: here again things become interesting. There are exactly 2 ideals of $\mathcal{O} = \mathbb{Z}[3q\omega]$ of norm p , namely the intersections with \mathcal{O} of the two ideals $\alpha\mathbb{Z}[\omega]$, $\bar{\alpha}\mathbb{Z}[\omega]$ into which p splits in $\mathbb{Z}[\omega]$. They are cubes in the class group $\text{Pic}(\mathbb{Z}[3q\omega]) \cong \mathbb{F}_{q^2}^\times/\mathbb{F}_q^\times$ if and only if α (equivalently $\bar{\alpha}$) or one of its associates $\omega\alpha$, $\omega\alpha^2$ is a cube modulo $3q$ (or an integer times a cube, but all integers mod $3q$ are cubes). The mod 3 condition requires we pick the unique associate (up to sign) with $\alpha \in \mathbb{Z}[3\omega]$, that is, α is *primary* in the classical terminology; and then we get a contribution of 2 ideals (with weight 3) or 0 according as this α is a cube or not modulo q .

So verifying Theorem 1.1 in this case amounts to proving a case of cubic reciprocity: that α is a cube mod q if and only if q is a cube mod α . Similar analysis of the case $p \equiv q \equiv 1 \pmod{3}$ forces us to invoke cubic reciprocity on two generic elements $\alpha, \beta \in \mathbb{Z}[3\omega]$. Although elementary proofs of cubic reciprocity are known, we can then proceed to the case $\Delta = -p^2q^2$, which leads us to an exotic cubic reciprocity law linking the fields $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt{3})$. The quest to systematize such reciprocity laws was, of course, one of the founding aims of class field theory.

7. h and class field theory

We now return to the general case and seek to interpret the cubic rings counted by $h(\Delta)$ via class field theory. Consider first the most generic case, in which our given cubic ring C sits in a cubic field K_3 which is not Galois over \mathbb{Q} (so Δ is not a square). Then the normal closure K_6 of K_3 has Galois group S_3 ; it contains a single quadratic subfield $K_2 = \mathbb{Q}(\sqrt{\Delta})$ of discriminant Δ_0 , the fundamental discriminant arising from decomposing $\Delta = \Delta_0 d^2$. The key insight regarding this network of fields

$$\begin{array}{ccc}
 & K_6 & \\
 & \swarrow \quad \searrow & \\
 K_3 & & K_2 \\
 & \swarrow \quad \searrow & \\
 & \mathbb{Q} &
 \end{array} \tag{27}$$

is the following theorem of Hasse. Recall that the *conductor* of an abelian extension L/K of number fields is the minimal modulus that the Artin symbol $\chi_{L/K} = \left(\frac{L}{K} \right)$ admits: it is a product of the ramified primes, appearing to exponents that may be computed using ramification groups. We denote it by $\text{cond}(L/K)$ or $\text{cond}(\chi_{L/K})$.

Lemma 7.1 ([10], Lemma 1.3; [7]). *The conductor of the extension K_6/K_2 is the principal ideal $(d) \subseteq \mathcal{O}_{K_2}$.*

Proof. First, we are asserting that K_6 is unramified at infinity, which is automatic for a Galois extension of odd degree.

The finite part of the conductor is related to the discriminant via the theory of ramification groups, which in the simple case of an extension L/K of prime order p , simply yields

$$\text{Disc}(L/K) = \text{cond}(L/K)^{p-1}$$

or, in our case,

$$\text{Disc}(K_6/K_2) = \text{cond}(K_6/K_2)^2.$$

Next, we have the “Brauer relation”

$$\text{Disc}(K_6) = \text{Disc}(K_3)^2 \cdot \text{Disc}(K_2)$$

(see [4], equation (2.7) for a more general discussion). This can be proved by writing the discriminants in terms of the ramification groups of the extensions K_6/K_3 , K_6/K_2 , and K_6/\mathbb{Q} . From this we get

$$N_{K_2/\mathbb{Q}}(\text{Disc}(K_6/K_2)) = \frac{\text{Disc } K_6}{(\text{Disc } K_2)^3} = (d^4).$$

However, by symmetry, $\text{Disc}(K_6/K_2)$ equals its conjugate, so $\text{Disc}(K_6/K_2) = (d^2)$ and thus $\text{cond}(K_6/K_2) = (d)$. \square

Thus there is associated to each non-Galois cubic field K_3 of discriminant $\Delta = \Delta_0 m^2$ an Artin map

$$\chi : I_{K_2}(m)/I_{K_2}(m, 1) \twoheadrightarrow \mu_3 \quad (28)$$

from the ray class group mod m onto a cyclic group of order 3, uniquely defined up to conjugation. (Here, as usual, $I_{K_2}(m)$ denotes the ideals prime to m and $I_{K_2}(m, 1)$ the principal ideals generated by elements congruent to 1 mod (m) .)

Conversely, given such a map χ , class field theory gives a cyclic extension K_6/K_2 of conductor dividing m of which it is the Artin map. However, not all of these extensions K_6 will be S_3 -Galois over \mathbb{Q} . The maps we want to consider are as follows:

Lemma 7.2. *If K_6/K_2 is a cubic Galois extension of a quadratic field K_2 , then the following are equivalent:*

- (a) K_6/\mathbb{Q} is Galois with $\text{Gal}(K_6/\mathbb{Q}) \cong S_3$;
- (b) $\chi_{K_6/K_2}(a) = 1$ for all $a \in \mathbb{Z}$ prime to the conductor $\mathfrak{d} = \text{cond}(K_6/K_2)$.

Proof. If (a) holds, the element $\tau \in \text{Gal}(K_6/\mathbb{Q})$ corresponding to a transposition $(12) \in S_3$ has the properties that

- (i) τ extends conjugation $\bar{\cdot} \in \text{Gal}(K_2/\mathbb{Q})$;
- (ii) Conjugation by τ interchanges the two nontrivial elements of $\text{Gal}(K_6/K_2)$.

From this symmetry, we derive that the Artin map $\chi = \chi_{K_6/K_2}$ satisfies

$$\chi(\bar{\mathfrak{a}}) = \chi(\mathfrak{a})^{-1} \quad (29)$$

for all $\mathfrak{a} \in I(K_2, \mathfrak{d})$, and plugging $\mathfrak{a} = (a)$ gives $\chi(a) = 1$.

Conversely, if (b) holds, then for any $\mathfrak{a} \in I(K_2, \mathfrak{d})$, the norm $N(\mathfrak{a}) = \mathfrak{a}\bar{\mathfrak{a}}$ is the quotient of two positive integers prime to \mathfrak{d} , and thus

$$\chi(\bar{\mathfrak{a}}) = \frac{\chi(N(\mathfrak{a}))}{\chi(\mathfrak{a})} = \chi(\mathfrak{a})^{-1},$$

verifying (29). Denote by \bar{K}_6 the field K_6 with K_2 embedded in it in the conjugate of the natural way. Then (29) can be written as

$$\chi_{\bar{K}_6/K_2} = \chi_{K_6/K_2}^{-1}.$$

By the uniqueness theorem of class field theory, K_6 and \bar{K}_6 are isomorphic as extensions of K_2 , but with the induced isomorphism of Galois groups being the negative of the natural one. In other words, K_6 has an automorphism τ having properties (i) and (ii) above. Now K_6 has six distinct automorphisms over \mathbb{Q} , so it is Galois, and it is easy to see that its Galois group must be S_3 . \square

So we only need to study Artin maps that factor through the quotient

$$I_{K_2}(m)/I_{K_2}(m, \mathbb{Z})$$

where $I_{K_2}(m, \mathbb{Z})$ is the subgroup of principal ideals generated by an element congruent to some integer (necessarily coprime to m) modulo (m) . This is a familiar quotient group: it is the ring class group $\text{Pic}(\mathcal{O}_\Delta)$ of the quadratic order $\mathcal{O}_\Delta = \mathbb{Z} + m\mathcal{O}_{\Delta_0}$ ([10], Lemma 1.9; also an easy consequence of (26) above).

Any $\chi : \text{Pic } \mathcal{O}_\Delta \rightarrow \mu_3$ yields an S_3 -Galois field K_6/\mathbb{Q} , and hence a non-Galois cubic field K_3/\mathbb{Q} . The discriminant of K_3 will be $\Delta = \Delta_0 m^2$ unless χ vanishes on a larger subgroup $I_{K_2}(d, \mathbb{Z}) \cap I_{K_2}(m)$, in which case χ has conductor d (for the smallest such d) and $\text{Disc}(K_3) = \Delta_0 d^2$.

Say that an integer x *squarely divides* an integer y if y/x is the square of an integer. We have just proved:

Lemma 7.3. *If $\Delta = \Delta_0 m^2$ is a non-square integer, the Artin map provides a bijection between cubic fields whose discriminant squarely divides Δ and group epimorphisms*

$$\chi : \text{Pic}(\mathcal{O}_\Delta) \rightarrow \mu_3$$

up to conjugation.

Remark 7.4. In particular, we have shown that for any such χ , the conductor $\text{cond}(\chi)$ is a principal ideal generated by an integer. An elementary proof of this fact is also possible; the details are left to the reader.

The case that $\Delta = m^2$ is a square, that is, $\Delta_0 = 1$, is similar but simpler, as we need only apply class field theory to the Galois extension K_3/\mathbb{Q} itself. The method of Lemma 7.1 shows that K_3/\mathbb{Q} has conductor m , yielding an Artin map

$$\chi_1 : I_{\mathbb{Q}}(m)/I_{\mathbb{Q}}(m, 1) \rightarrow \mu_3.$$

In the interest of conformity with the preceding, we use the bijection $\chi_1 \mapsto (\chi_1, \chi_1^{-1})$ to put these in bijection with maps

$$\chi : I_{\mathbb{Q} \times \mathbb{Q}}(m)/I_{\mathbb{Q} \times \mathbb{Q}}(m, \mathbb{Z}) = \text{Pic}(\mathcal{O}_{m^2}) \rightarrow \mu_3,$$

yielding the following uniform parametrization.

Lemma 7.5. *Let $\Delta \in \text{Discs}$, and let \mathcal{O}_Δ denote the quadratic ring of discriminant Δ . The Artin map provides a bijection between cubic fields whose discriminant squarely divides Δ and group epimorphisms*

$$\chi : \text{Pic}(\mathcal{O}_\Delta) \twoheadrightarrow \mu_3$$

up to conjugation.

7.1. Splitting types

Suppose we wish to compute $h(\Delta)$ for some $\Delta = \Delta_0 m^2$. We can list all of the fields K whose discriminant $\Delta_0 d^2$ squarely divides Δ ; then we must count orders in K of index m/d . By Remark 4.3, we can compute this knowing the splitting types of K at each of the primes dividing m/d . The following proposition (which should also be credited to Hasse: see the table in [7], p. 568) gives a simple way to find these splitting types given the corresponding Artin map χ .

Proposition 7.6 (Hasse). *Let $\chi : \text{Pic}(\mathcal{O}_{\Delta_0 d^2}) \rightarrow \mu_3$ be a primitive character (i.e. one that does not factor through any $\text{Pic}(\mathcal{O}_{\Delta_0 d'^2})$, $d' | d$). Let $p \in \mathbb{Z}$ be a prime. The splitting type of p in the \mathbb{Q} -algebra K corresponding via Lemma 7.5 to χ is*

- 1^3 if $p | d$,
- $1^2 1$ if $p \nmid d$ but $p | \Delta_0$,
- 12 if $p \nmid d$ and p is inert in \mathcal{O}_{Δ_0} (i.e. the Kronecker symbol $\left(\frac{\Delta_0 d^2}{p}\right)$ has the value -1),
- 111 if $p \nmid d$ and $p = \mathfrak{p}\bar{\mathfrak{p}}$ in $\mathcal{O}_{\Delta_0 d^2}$ with $\chi(\mathfrak{p}) = 1$;
- 3 if $p \nmid d$ and $p = \mathfrak{p}\bar{\mathfrak{p}}$ in $\mathcal{O}_{\Delta_0 d^2}$ with $\chi(\mathfrak{p}) \neq 1$.

In particular, all splitting types can be told apart merely by reference to the discriminant $\Delta = \Delta_0 d^2$, except for 111 and 3 .

Proof. If $\Delta_0 = 1$, then K , being Galois, can only have splitting type 111 , 3 , or 1^3 , and class field theory readily implies that these cases occur exactly under the conditions claimed. So assume that $\Delta_0 > 1$.

The primitivity of χ implies that

$$\text{Disc } K = \Delta = \Delta_0 d^2.$$

We immediately see that K is ramified if and only if $p | \Delta$. If K has splitting type $1^2 1$, then the completion K_p has the form $\mathbb{Q}_p \times Q$, where Q/\mathbb{Q}_p is a ramified quadratic extension. Bearing in mind that the discriminant of a p -adic field is well-defined only up to the square of a p -adic unit, we have

$$\text{Disc } Q = \text{Disc } K_p = \text{Disc } K = \Delta_0 d^2.$$

But a quadratic ring over \mathbb{Z}_p , as over \mathbb{Z} , is determined by its discriminant, so

$$\mathcal{O}_Q = (\mathcal{O}_{\Delta_0 d^2})_p = \mathbb{Z}_p + d(\mathcal{O}_{\Delta_0})_p.$$

Since \mathcal{O}_Q is maximal, we get $p \nmid d$, and hence $p \mid \Delta_0$.

If K has splitting type 1^3 , then K_6/K_2 is ramified at some prime $\mathfrak{p}_2 \mid p$ (otherwise the ramification index of p in K_6 could be at most 2), implying that \mathfrak{p}_2 , and hence p , divides the conductor $\text{cond}(K_6/K_2) = d$.

In the case that p is unramified, there are just three cases: the splitting types 111, 12, 3 are also the cycle types of Frob_p as an element of $\text{Gal}(K_6/\mathbb{Q}) \cong S_3$. Note that cycle type 12, being the only odd permutation, corresponds exactly to the case that the discriminant field $\mathbb{Q}(\sqrt{\Delta})$ is inert at p . The other two cases can be told apart via class field theory: here p splits as a product $p = \mathfrak{p}_2 \bar{\mathfrak{p}}_2$ in \mathcal{O}_{Δ_0} and hence as a product $\mathfrak{p} \bar{\mathfrak{p}}$ in \mathcal{O}_Δ (as $p \nmid d$). The Artin symbol $\chi(\mathfrak{p}) = \chi(\mathfrak{p}_2)$ vanishes if and only if \mathfrak{p}_2 splits completely in K_6 , which happens exactly when p splits completely in K_3 . \square

In our computation of $h(\Delta)$, we are still missing the contribution of the nondomains, which are subrings of $K_s = \mathbb{Q} \times \mathcal{O}_{\Delta_0}$. The splitting type of K_s at every prime p is either 111, 12, or $1^2 1$, and one finds that applying [Proposition 7.6](#) to the trivial character $\chi = 1$, of conductor $d = 1$, yields the right answer. So it makes sense to define the Artin map of a nondomain to be identically 1.

Note that K_s has twice as many automorphisms as the other \mathbb{Q} -algebras whose discriminants squarely divide Δ (6 vs. 3 if $\Delta_0 = 1$, and 2 vs. 1 otherwise). On the other hand, if we sum up over all maps $\chi : \text{Pic } \mathcal{O}_\Delta \rightarrow \mu_3$, then the Artin maps corresponding to fields K appear twice, due to the conjugation ambiguity, but the trivial Artin map appears only once. So, counting the automorphisms carefully, we arrive at the following result.

Theorem 7.7. *Let $\Delta = \Delta_0 m^2 \in \mathcal{D}\text{iscs}$. Let $w_\Delta = 3$ if Δ is a square, 1 otherwise. The following quantities are equal:*

- $2w_\Delta h(\Delta)$;
- The sum, over all characters $\chi : \text{Pic } \mathcal{O}_\Delta \rightarrow \mu_3$, of the number of subrings of index $m/\text{cond } \chi$ in the cubic ring C whose local splitting types are determined by $\text{Disc } C = \Delta_0(\text{cond } \chi)^2$ and χ .

8. Finishing the proof of [Theorem 1.1](#)

To finish the proof for a value $\Delta = \Delta_0 m^2$, we would like to equate the expression for $2w_\Delta h(\Delta)$ in [Theorem 7.7](#) with the expression for $2w_\Delta \sigma_\Delta \hat{h}(\Delta)$ in [Theorem 5.4](#), which is reproduced here for convenience:

- $2w_{\Delta}\sigma_{\Delta}\hat{h}(\Delta)$ is the number of invertible ideals I of norm g whose class is a cube in orders $\mathcal{O}_{\Delta'}$ for integers $g > 0$, Δ' satisfying $\Delta'g^2 = \Delta$, each counted with weight

$$|\text{Pic}(\mathcal{O}_{\Delta'})[3]|.$$

It is not hard to turn $2w_{\Delta}\sigma_{\Delta}\hat{h}(\Delta)$ into a character sum, as follows. If I is an invertible ideal in $\mathcal{O}_{\Delta'}$, then

$$\sum_{\chi: \text{Pic } \mathcal{O}_{\Delta'} \rightarrow \mu_3} \chi(I) = \begin{cases} |\text{Hom}(\text{Pic } \mathcal{O}_{\Delta'}, \mu_3)| = |\text{Pic}(\mathcal{O}_{\Delta'})[3]| & \text{if } [I] \in \text{Pic}(\mathcal{O}_{\Delta'})^3 \\ 0 & \text{otherwise.} \end{cases}$$

So

$$\begin{aligned} w_{\Delta}\sigma_{\Delta}\hat{h}(\Delta) &= \sum_{cf=m} \sum_{\substack{I \subseteq \mathcal{O}_{\Delta_0 c^2} \\ \text{invertible, norm } f}} \sum_{\chi: \text{Pic } \mathcal{O}_{\Delta_0 c^2} \rightarrow \mu_3} \chi(I) \\ &= \sum_{\chi: \text{Pic } \mathcal{O}_{\Delta} \rightarrow \mu_3} \sum_{\substack{cf=m, \\ \text{cond}(\chi)|c}} \sum_{\substack{I \subseteq \mathcal{O}_{\Delta_0 c^2} \\ \text{invertible, norm } f}} \chi(I). \end{aligned}$$

It suffices to prove that, at least for cubefree m (in view of [Remark 4.2](#)), the contribution of each χ to $2w_{\Delta}h(\Delta)$ and $2w_{\Delta}\sigma_{\Delta}\hat{h}(\Delta)$ is the same. In other words, fix a χ ; let $c_1 = \text{cond } \chi$, $\Delta_1 = \Delta_0 c_1^2$, $m_1 = m/c_1$. We will prove that the number of subrings of the corresponding \mathbb{Q} -algebra K_{χ} of index m_1 is equal to

$$h(m_1, \chi) = \sum_{c'f=m_1} \sum_{\substack{I \subseteq \mathcal{O}_{\Delta_1 c'^2} \\ \text{invertible, norm } f}} \chi(I). \quad (30)$$

We first observe that the number of subrings is a multiplicative function of the index m_1 and claim that $h(m_1, \chi)$ is also. If $m_1 = m_2 m_3$ with $\gcd(m_2, m_3) = 1$, then we get corresponding decompositions $c' = c'_2 c'_3$ and $f = f_2 f_3$. An invertible ideal I of norm $f_2 f_3$ in $\mathcal{O}_{\Delta_0(c_0 c'_2 c'_3)^2}$ can be decomposed uniquely as a product $I_2 I_3$, where $I_i = f_{5-i}^{-1} I \cap \mathcal{O}_{\Delta_1(c'_2 c'_3)^2}$ is an invertible ideal of norm f_i ; since f_i is prime to c'_{5-i} , invertible ideals of norm f_i in the orders $\mathcal{O}_{\Delta_1(c'_2 c'_3)^2}$ and $\mathcal{O}_{\Delta_1 c'_i^2}$ are in bijection. The sum then factors readily, using the multiplicativity of χ .

Thus we can assume that $m_1 = p^k$ is a prime power. We once again have a local problem. There are several cases. The case $k = 0$ is trivial, so we have $k = 1$ or $k = 2$. The following table shows the types of invertible ideals on which we must evaluate χ and sum:

	$c' = 1$	$c' = p$	$c' = p^2$
$k = 1$	norm p in \mathcal{O}_{Δ_1}	unit ideal in $\mathcal{O}_{\Delta_1 p^2}$	
$k = 2$	norm p^2 in \mathcal{O}_{Δ_1}	[norm p in $\mathcal{O}_{\Delta_1 p^2}$]	unit ideal in $\mathcal{O}_{\Delta_1 p^4}$

The bottom middle entry has been placed in brackets because no such ideals exist. Suppose to the contrary that we had a map $\phi : \mathcal{O}_{\Delta_1 p^2} \rightarrow \mathbb{F}_p$ with kernel an invertible ideal. Let ξ be a generator of \mathcal{O}_{Δ_1} , so $\mathcal{O}_{\Delta_1 p^2} = \mathbb{Z}[p\xi]$. We have $\phi(p\xi)^2 = p \cdot \phi(p\xi^2) = 0$, so $\phi(p\xi) = 0$ and hence $\ker \phi = \mathbb{Z}\langle p, p\xi \rangle = p\mathcal{O}_{\Delta_1}$, which is not an invertible ideal.

If $p|c_1$, then $k = 1$, and by the same argument, \mathcal{O}_{Δ_1} has no invertible ideals of index p , so the value of $h(p, \chi)$ is simply 1, coming from the unit ideal in $\mathcal{O}_{\Delta_1 p^2}$. This accords with the number 1 of subrings of index p in a maximal ring of splitting type 1^3 , as tabulated in (15).

In the remaining cases, $p \nmid c_1$, so p has the same splitting type in \mathcal{O}_{Δ_1} as in \mathcal{O}_{Δ_0} . We only have to sum χ over ideals of norm p^k in \mathcal{O}_{Δ_1} , all of which will be invertible, and add the contribution 1 coming from the unit ideal in $\mathcal{O}_{\Delta_1 p^{2k}}$.

If p is inert, then \mathcal{O}_{Δ_1} has no ideals of norm p and one ideal of norm p^2 , namely (p) , with $\chi((p)) = 1$. So the total $h(p^k, \chi)$ is 1 for $k = 1$ and 2 for $k = 2$, in accordance with (15) for K_χ having splitting type 12.

If $p = \mathfrak{p}^2$ ramifies in \mathcal{O}_{Δ_1} , then \mathcal{O}_Δ has one ideal each of norm p and p^2 . Note that $\chi(\mathfrak{p}) = 1$ since $\mathfrak{p}^2 = (p)$ is principal. So the total $h(p^k, \chi)$ is 2 for both $k = 1$ and $k = 2$, in accordance with (15) for splitting type $1^2 1$.

Finally, if $p = \mathfrak{p}\bar{\mathfrak{p}}$ is split in \mathcal{O}_{Δ_1} , then \mathcal{O}_{Δ_1} has two ideals of norm p (\mathfrak{p} and $\bar{\mathfrak{p}}$) and three ideals of norm p^2 (\mathfrak{p}^2 , $\bar{\mathfrak{p}}^2$, and $\mathfrak{p}\bar{\mathfrak{p}} = (p)$). We know that $\chi(\bar{\mathfrak{p}}) = \chi(\mathfrak{p})^{-1}$. Adding up χ on the relevant ideals in the two cases $\chi(\mathfrak{p}) = 1$, $\chi(\mathfrak{p}) \neq 1$ matches the four entries of (15) for splitting types 111 and 3, finishing the proof. \square

Remark 8.1. In the splitting types 111, 12, and $1^2 1$, we can assume without loss of generality that $K_3 = K_2 \times \mathbb{Q}$, and the number of suborders of index p^k is seen to be $h(p^k, 1)$ via the bijection

$$I \longleftrightarrow \mathbb{Z}_p + (I_p \times 0).$$

It is unclear whether such a bijective proof is to be hoped for when $\chi \neq 1$.

Acknowledgments

I thank Manjul Bhargava for communicating this beautiful problem to me on a visit to Princeton (during which Bhargava was supposed to be presenting at Harvard, but his flight was felicitously canceled). I thank my fellow Cambridge students for encouraging me to propose the problem as an essay topic and write a Part III essay on it. I thank Jack Thorne for reading it as essay assessor.

This research was completed on a Churchill Scholarship funded by Mario Gabelli through the Winston Churchill Foundation of the United States.

References

- [1] Manjul Bhargava, Higher composition laws. I. A new view on Gauss composition, and quadratic generalizations, *Ann. of Math. (2)* 159 (1) (2004) 217–250.
- [2] Manjul Bhargava, Higher composition laws. II. On cubic analogues of Gauss composition, *Ann. of Math. (2)* 159 (2) (2004) 865–886.
- [3] Manjul Bhargava, Higher composition laws and applications, in: *International Congress of Mathematicians*, vol. II, Eur. Math. Soc., Zürich, 2006, pp. 271–294.
- [4] Henri Cohen, Simon Rubinstein-Salzedo, Frank Thorne, Identities for field extensions generalizing the Ohno–Nakagawa relations, *Compos. Math.* 151 (11) (2015) 2059–2075.
- [5] B.N. Delone, D.K. Faddeev, *The Theory of Irrationalities of the Third Degree*, *Translations of Mathematical Monographs*, vol. 10, American Mathematical Society, Providence, R.I., 1964.
- [6] Benedict H. Gross, Mark W. Lucianovic, On cubic rings and quaternion rings, *J. Number Theory* 129 (6) (2009) 1468–1478.
- [7] Helmut Hasse, *Arithmetische Theorie der kubischen Zahlkörper auf klassenkörpertheoretischer Grundlage*, *Math. Z.* 31 (1) (1930) 565–582.
- [8] Friedrich Wilhelm Levi, *Kubische Zahlkörper und binäre kubische Formenklassen (Cubic number fields and cubic form classes)*, *Leipz. Ber.* 66 (1914) 26–37.
- [9] Guillermo Mantilla-Soler, Integral trace forms associated to cubic extensions, *Algebra Number Theory* 4 (6) (2010) 681–699.
- [10] Jin Nakagawa, On the relations among the class numbers of binary cubic forms, *Invent. Math.* 134 (1) (1998) 101–138.
- [11] Jürgen Neukirch, *Algebraic Number Theory*, *Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences)*, vol. 322, Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.
- [12] Yasuo Ohno, A conjecture on coincidence among the zeta functions associated with the space of binary cubic forms, *Amer. J. Math.* 119 (5) (1997) 1083–1094.
- [13] Yasuo Ohno, Takashi Taniguchi, Relations among Dirichlet series whose coefficients are class numbers of binary cubic forms II, *Math. Res. Lett.* 21 (2) (2014) 363–378.
- [14] Takuro Shintani, On Dirichlet series whose coefficients are class numbers of integral binary cubic forms, *J. Math. Soc. Japan* 24 (1972) 132–188.
- [15] Frank Thorne, Four perspectives on secondary terms in the Davenport–Heilbronn theorems, in: *Integers*, 12B, *Proceedings of the Integers Conference 2011*: Paper No. A5, 23 pp., 2012/13.