MATHEMATICS OF COMPUTATION
VOLUME 63, NUMBER 207
JULY 1994, PAGES 329–334

HEURISTICS ON CLASS GROUPS: SOME GOOD PRIMES ARE NOT TOO GOOD

HENRI COHEN AND JACQUES MARTINET

ABSTRACT. We correct some too optimistic predictions in an earlier paper of ours.

1

When generalizing to arbitrary relative extensions of number fields the Cohen and Lenstra heuristics of [1], it appeared that the natural hypothesis, namely not to include the p-components of the class groups for those p which divide the degree of a Galois closure, could be somewhat weakened, and we were led to the notion of a "*good prime*" (cf. §2 below). Primes which do not divide the degree of the Galois closure are good, and we added some more primes to this list. In particular, 2 was considered to be a good prime for all cubic fields, though it divides the degree of the Galois closure when the field is not Galois.

It was clear to us as soon as we began our work that these additional good primes could produce special difficulties [2, §3, p. 136]. However, the recent computations of G. Fung and H. Williams on nonreal cubic fields showed clearly that our recipes cannot be applied in a naive manner to the extra good primes.

2

Recall briefly some of the basic ideas of our heuristics. Let K_0 be a fixed number field. We first consider (within a given algebraic closure of $\mathbb{Q}$ containing K_0) a set of finite Galois extensions K/K_0 with a given Galois group Γ (up to isomorphism) and for which the conjugacy classes of the infinite Frobenius substitutions attached to the real places of K_0 are given. (In particular, the behavior in K/K_0 of the real primes of K_0 is uniquely determined.) Let χ be a character of Γ with rational values, and let e_χ be the corresponding central idempotent of $\mathbb{Q}[\Gamma]$ [3, §6]. The formula for e_χ shows that, if p is a prime which does not divide $[K : K_0]$, we can consider $e_\chi G$ for any finite $\mathbb{Z}[\Gamma]$-module G. Our heuristics try to predict the behavior of these $e_\chi G$ when G is the relative class group Cl_{K/K_0} with the prime divisors of $[K : K_0]$ removed.

To handle Cl_{K/K_0} for K/K_0 non-Galois, we simply apply the above considerations to the Galois closure of K/K_0, taking for χ the augmentation character

Received by the editor February 5, 1991 and, in revised form, September 27, 1991.
1991 *Mathematics Subject Classification.* Primary 11R29; Secondary 11Y40.
Key words and phrases. Class group, class number, number field.

(permutation character − unit character). For example, when K/K_0 is a non-cyclic cubic, the augmentation character χ is the irreducible degree-2 character of S_3.

Our heuristics rely on combinatorial calculations on finite modules over local maximal orders in semisimple algebras over $\mathbb{Q}$ [3, §2]. We must thus restrict ourselves to p-components for which e_χ is p-integral and $\mathbb{Z}[\Gamma]e_\chi$ is a maximal order locally at p in the algebra $\mathbb{Q}[\Gamma]$. These conditions are fulfilled when p does not divide the order of Γ. Now, we define a prime p to be *good for a character* χ when:

(i) $e_\chi \in \mathbb{Z}_p[\Gamma]$;
(ii) $\mathbb{Z}_p[\Gamma]e_\chi$ is a maximal order.

With this definition, 2 is a good prime for cubic extensions. (Let $S_3 = \langle \sigma, \tau ; \sigma^3 = \tau^2 = 1, \tau\sigma\tau^{-1} = \sigma^{-1} \rangle$; then, $e_\chi = \frac{2-\sigma-\sigma^2}{3}$ and $\mathbb{Z}_2[\Gamma]e_\chi \simeq M_2(\mathbb{Z}_2)$.)

3

Our heuristics predict that the proportion of extensions of discriminant $\leq x$ for which the class group is of order prime to a given (good) p, or has a p-component of order p or a cyclic one, or contains a subgroup of type (p^2, p, p), or is of order a power of p,..., has a limit for $x \to \infty$, and moreover we give a conjectural value for this limit. Our predictions fit well with the known extensive tables as far as only primes which do not divide the degree are involved.

This is the case for quadratic fields, both real and imaginary. Furthermore, the accordance with the tables is still good *if we restrict ourselves to prime discriminants or discriminants in a given (admissible) congruence class modulo some integer, or both.*

4

Let us now turn to cubic fields, and first to pure cubic fields $K = \mathbb{Q}(\sqrt[3]{p})$, $p \equiv -1 \bmod 3$ a prime (and $K_0 = \mathbb{Q}$), for which long tables exist (Shanks, Tennenhouse, and Williams, ref. 13 and 15 of [2]). Note that the hypothesis $p \equiv -1 \bmod 3$ is just to ensure that the class number is prime to 3, but that we could otherwise simply look at the prime to 3 component of the class group. The tables handle p's up to 10^6, with corresponding discriminants sometimes lower than -10^{13} (one has $d_K = -3p^2$ for $p \equiv -1 \bmod 9$ and $d_K = -27p^2$ otherwise). Then, a discrepancy appears between the $p \equiv -1 \bmod 9$ and the others ($\equiv 2$ or $5 \bmod 9$). The tables indicate that the proportion of fields with class number 1 **is not** the predicted one in these congruence classes, and it is only by averaging the three classes and taking into account the different values of the discriminant that we obtain the compatibility with our formula [2, §3, p. 136].

An inspection of the tables [4, p. 567] indicates that it is the prime 2 which is responsible for this abnormal behavior.

5

Let us now look at general complex cubic fields on the basis of Fung and Williams's paper [6], and in particular Table 5.6. As in [6], denote by h_0 the prime to 3 part of the class number. Although the number of fields examined

is not large (182417), it seems that the proportion of fields with an odd class number will remain significantly larger than our predictions (for example, for $h_0 = 1$ the proportion decreases from 73.5% to 67.5% when the discriminant bound goes from -10^5 to -10^6, while we predict 51.9%). This could of course be due to the fact that the convergence is extremely slow, and in fact Fung and Williams have noted that even the Davenport-Heilbronn *theorem* which gives $1/(4\zeta(3))$ as the limit as $x \to \infty$ of the number of complex cubic fields K such that $|d_K| < x$ divided by x, seems quite far off the observed quantities.

However, the relative discrepancy in this case is only about 10% (18.2% observed at $x = 10^6$, versus 20.8% proved by Davenport-Heilbronn) while the discrepancy for class number 1 is about 30%.

A more convincing argument against our predictions is to look at the cases of $h_0 = 5$ and $h_0 = 7$. In all the systematic tables involving class numbers that we have seen, the proportions of fields having a given property (for example $h_0 = 5$) is always a *monotonic* function of the discriminant bound, on average of course, since the class number behaves erratically. Although the reason for this behavior is not understood, it gives us confidence in saying that the observations of Fung and Williams concerning $h_0 = 5$ and $h_0 = 7$ are in contradiction with our predictions. For $h_0 = 5$, the observed proportions increase regularly from 2.77% for $x = 10^5$ to 2.945% for $x = 10^6$, while we predict 2.59%. Similarly for $h_0 = 7$, the observed proportions increase regularly from 1.21% for $x = 10^5$ to 1.38% for $x = 10^6$, while we predict 1.23%.

Although these percentages correspond to only a few thousand fields, the monotonicity of the average behavior is for us a strong indication that the prediction is incorrect. The explanation that we give for this (which at present is admittedly not supported by enough numerical evidence) is that the (supposedly) good prime 2 should be excluded.

In fact, another way of looking at the data of Fung and Williams which better shows the local behavior at each prime is, instead of looking at the proportion of fields with h_0 equal to a given number m, to look at the proportion of fields with h_0 *divisible* by a given prime p (since then the other prime numbers play no role). Then, although the convergence is still very slow, no contradiction appears for $m = 5$ and $m = 7$. For $m = 5$ (resp. $m = 7$), the proportion of fields with $m \mid h_0$ increases slowly from 2.2% to 3.7% (resp. from 1% to 1.7%), which in both cases is about 25% under the predicted limit at $x = 10^6$.

On the other hand, for $m = 2$ we are still 37% under the predicted limit (26.5% instead of 42.2%), which seems too large a discrepancy (note that this is based on the complete data kindly communicated to us by Fung and Williams and not on their paper alone which does not quite contain all the needed information).

If we decide to consider 2 as a bad prime, the numerical predictions of [2] have to be restated. Instead of considering the prime to 3 part of the class group, we must look at the prime to 6 part, and the conjectures must be changed accordingly (essentially by including a local factor for $p = 2$ in addition to the one for $p = 3$).

For example, if we look at the proportion of complex cubic fields whose class number has only powers of 2 and 3 (i.e., the prime to 6 part trivial), we predict 89.80%, while the observed values decrease slowly from 94.6% for $x = 10^5$ to

92.6% for $x = 10^6$, which seems quite reasonable. Thus, the contradictions noticed in the comments of Table 5.4 of [6] disappear when 2 is considered as a bad prime. In addition, one can check that the contradictions noticed above concerning the fields with class number equal to 5 or 7 also disappear.

6

In the original work by Cohen and Lenstra, who only considered cyclic fields K of some prime degree ℓ, the basic reason to exclude the prime ℓ was the existence of the theory of genera (or of invariant classes), which is clearly an obstruction to the random behavior of the ℓ-component $\mathrm{Cl}_{K,\ell}$ of Cl_K. Experimental data show that the existence of particular subgroups of index ℓ (or of order ℓ) is probably the only obstruction to a random behavior of Cl_K. For cubic fields, *there is no obstruction of this kind for the prime* 2, as shown by Jaulent's theorem [3, Theorem 7.8], a particular case of which states that, given a cubic extension K/K_0 with Galois closure N/K_0 and quadratic subextension L/K_0, then the natural map from Cl_{K/K_0} to $CL_{N/L}$ is one to one, and this shows that genus theory does not involve the prime 2, since $[N : L] = 3$.

All known obstructions to a random behavior of $\mathrm{Cl}_{K,2}$ vanish, and we are faced with a mystery.

6.1. **Problem.** To find an explanation for the too small proportion of complex cubic fields with even class number.

Note that this can be compared (in the situation considered in §4) with the behavior of the Selmer group of elliptic curves (cf. [4]). However, this does not explain the mystery, but simply says that it possesses a transcription in terms of elliptic curves.

Note also that the fields considered in §4 are exactly those with associated quadratic field $\mathbb{Q}(\sqrt{-3})$ (and also exactly one ramified prime besides 3, a harmless condition with respect to our problem). Now, let d be a quadratic discriminant, and $\mathscr{C}_d$ be the set of cubic fields whose Galois closure contains $\mathbb{Q}(\sqrt{d})$.

6.2. **Question.** Is the proportion of fields in $\mathscr{C}_d$ with, say, odd class number, or cyclic 2-component, the one we predict?

There is no problem for $d = 1$ (the case of cyclic fields). Note that a positive answer to question 6.2 would not contradict the global behavior of cubic fields, since we do not ask for any remainder term in the proportion of fields in $\mathscr{C}_d$.

7

In analogy with class groups, it would be interesting to guess the heuristic behavior of narrow class groups Cl_K^+ (or more generally Cl_{K/K_0}^+), and first of the quotient group $\mathrm{Cl}_K^+ / \mathrm{Cl}_K$. We were not able to suggest a reasonable answer (cf. the discussion in [3, §8, f]). The first nontrivial example is that of real cubic fields, for which the order of $\mathrm{Cl}_K^+ / \mathrm{Cl}_K$ (1,2 or 4) was calculated by Ennola and Turunen (ref. 6 of [2]). It *seems* that there is a definite proportion of fields for each of the 3 possibilities that can occur. Now, there are some close connections between $\mathrm{Cl}_K^+ / \mathrm{Cl}_K$ and $\mathrm{Cl}_{K,2}$ (Oriat, ref. [O] of [3]), which make it likely that the behavior of $\mathrm{Cl}_{K,2}$ and $\mathrm{Cl}_{K,2}^+$ should obey the same rules. Moreover ([3, §8, h]), such a behavior plus some other natural hypotheses would imply an asymptotic estimate $N(x) \sim c \times x$ (for a certain constant c) for the number of

S_4 quartic fields L with $d_L \leq x$, an estimate which is in good accordance with the tables of Buchmann and Ford in the totally real case (ref. [B-F] of [3]), and which has just received a nice confirmation in the totally complex case (Ford, [5]), a case where narrow class groups are involved.

For these reasons, we think that *a definite proportion of cubic fields* (in both the complex and the real case) *has a prescribed 2-component, a prescribed class-group modulo the 3-component*,..., and that a similar result should hold for Cl_K^+.

We also think that good, but not too good, primes should behave in the same manner in all cases (e.g., the prime 2 for Γ dihedral of order $2m$, m odd, or the prime 3 for $\Gamma \simeq \mathfrak{A}_4$ or $\mathfrak{S}_4$, but with densities different from the ones which were computed in [2]).

However, because of the "nonsemisimple Spiegelungssatz" which links $\mathrm{Cl}_K^+ / \mathrm{Cl}_K$ and $\mathrm{Cl}_{K,2}$, the prime 2 could be somewhat special.

8

We would like to finish this note with a few remarks on units. It is convenient to break the study in two parts.

a) What is the regulator distribution? We have no precise statement to suggest, and shall simply make one remark in the real quadratic case. As is well known, the regulators lie roughly between $\log d$ and $d^{\frac{1}{2}}$ (d is the discriminant). The heuristics on class groups together with the Brauer-Siegel theorem make us believe that, for any positive ε, most of the regulators should lie in the interval $[d^{\frac{1}{2}+\varepsilon}, d^{\frac{1}{2}-\varepsilon}]$. Is there a measure for which equidistribution holds?

b) Let us now rescale the logarithmic lattice such as to give it the minimal norm 1 (i.e., we consider this lattice up to similarity). We are now faced with the question of the *shape* of the lattice. But there is a natural measure on the similarity classes of lattices in a Euclidean space of a given dimension m, which is defined via the double classes

$$\mathbb{R}^* O_m(\mathbb{R}) \backslash \mathrm{Gl}_m(\mathbb{R}) / \mathrm{Gl}_m(\mathbb{Z}).$$

8.1. **Question.** Are the unit lattices equally distributed for this measure?

One may ask such a question whenever Euclidean lattices are involved, for example for the group of rational points modulo torsion of an elliptic curve defined on a number field when it is equipped with the Néron-Tate height. Question 8.1 was asked independently of us by Armand Brumer at the occasion of his work on the asymptotic behavior of elliptic curves.

Bibliography

1. H. Cohen and H. W. Lenstra, *Heuristics on class groups of number fields*, Lecture Notes in Math., vol. 1068, Springer-Verlag, Berlin and New York, 1984, pp. 33–62.
2. H. Cohen and J. Martinet, *Class groups of number fields: numerical heuristics*, Math. Comp. **48** (1987), 123–137.
3. ———, *Étude heuristique des groupes de classes des corps de nombres*, J. Reine Angew. Math. **404** (1990), 39–76.
4. H. Eisenbeis, G. Frey, and B. Ommerborn, *Computation of the 2-rank of pure cubic fields*, Math. Comp. **32** (1978), 559–569.

5. D. Ford, *Enumeration of totally complex quartic fields of small discriminant,* Computational Number Theory (A. Pethö, M. Pohst, H. Williams, and H. G. Zimmer, eds.), deGruyter, 1991.
6. G. Fung and H. Williams, *On the computation of a table of complex cubic fields with discriminant* $D > -10^6$, Math. Comp. **55** (1990), 313–325.

A2X, Mathématiques, 351, cours de la Libération, 33405 Talence cedex, France
E-mail address: cohen@ceremab.u-bordeaux.fr
E-mail address: martinet@ceremab.u-bordeaux.fr