# RESEARCH STATEMENT

SIMON RUBINSTEIN-SALZEDO

My primary research interests lie in algebraic number theory and the mathematics surrounding the inverse Galois problem. More specifically, I am interested in understanding number fields with limited ramification, which can also be interpreted in terms of a suitable étale fundamental group. In particular, I have done work on improving the Cohen-Lenstra-Martinet heuristics and on using origamis and origami curves to construct number fields ramified at only a specified set of primes.

I am also interested in other topics in arithmetic geometry, in particular point counting on curves over finite fields. Some questions of this flavor arose out of my study of origamis and branched covers of algebraic curves, but they can also be interpreted in terms of periodic points of dynamical systems, and answering such questions would teach us more about arithmetic dynamics. On the other hand, dynamical systems, and in particular postcritically finite maps, can also be used to control ramification of curves and number fields, so it is interesting that dynamical systems and ramification in number fields are subjects of mutual benefit to one another.

## 1. Cohen-Lenstra-Martinet heuristics

One project I have worked on is related to the Cohen-Lenstra-Martinet heuristics and roots of unity, as written up in my paper [RS12a]. Roughly speaking, the Cohen-Lenstra-Martinet heuristics (see [CL84], [CM87], and [CM90]) state that, if we fix a prime $p$ and an abelian $p$-group $A$, and we let $K$ vary over some suitable collection of number fields, then the proportion of $K$ so that the Sylow $p$-subgroup of the class group of $K$ is isomorphic to $A$, is inversely proportional to the size of the automorphism group of $A$ times some power of the order of $A$. The above statement must be qualified in many ways in order for it to have any hope whatsoever of being true, but here is one special case that I have investigated:

**Heuristic 1.** *Let $\ell$ be an odd prime, and let $G = C_\ell$. Let $D(X)$ be the set of $C_\ell$ number fields with discriminant less than $X$. (Such fields are necessarily totally real.) Also, let $p$ be a prime different from $\ell$ and $A$ an abelian $p$-group with the structure of a $\mathbb{Z}[\zeta_\ell]$-module. Then, if $\mathrm{Cl}_p(K)$ denotes the Sylow $p$-subgroup of the class group of $K$,*

$$\lim_{X\to\infty} \frac{\#\{K \in D(X) : \mathrm{Cl}_p(K) \cong A\}}{\#D(X)}$$

*Date*: October 26, 2012.

*exists, and is inversely proportional to* $|\operatorname{Aut}_{\mathbb{Z}[\zeta_\ell]}(A)| \cdot |A|^{\ell-1}$. *(By* $\operatorname{Aut}_{\mathbb{Z}[\zeta_\ell]}$*, we mean that we only consider automorphisms of* $A$ *which preserve the* $\mathbb{Z}[\zeta_\ell]$*-module structure.)*

Malle, in [Mal08], did extensive computations to test this heuristic and discovered that it seems to be wrong when $p = 2$, and that more generally, the predictions that the Cohen-Lenstra-Martinet heuristics make do not appear to apply for a prime $p$ when there are $p^{\text{th}}$ roots of unity either in the ground field or else in an intermediate field.

Ellenberg and Venkatesh in [VE10] proposed a correction to the Cohen-Lenstra-Martinet heuristics coming from the Schur multiplier group $H_2(A;\mathbb{Z})$. We conjecture that those fields $K$ for which the $p$-power torsion in $\operatorname{Cl}(K)$ is isomorphic to $A$ split up naturally into several classes (which we call invariants), parametrized by a certain subgroup of $H_2(A;\mathbb{Z})$, and each class occurs with the probability predicted by the Cohen-Lenstra-Martinet heuristics. The discrepancy observed in Malle's computations agrees with the number of invariants Ellenberg and Venkatesh predict, in the cases considered. Thus, in order to improve the Cohen-Lenstra-Martinet heuristics, we need to know the distribution of the invariant among the possible classes. In particular, we suspect that the invariant is equidistributed.

If we pick $p = 2$ and $A = C_2 \times C_2$, and let $K$ run over $C_3$ cubic fields, we can describe the invariants in a manner more amenable to computation and use this description to compute many examples in order to determine whether or not the invariant does indeed equidistribute. In this case, there are two possible invariants, which we call 0 and 1.

My contribution to the subject was to provide several concrete and computable interpretations of the invariant, create an algorithm to compute it relatively quickly, and then perform the computation in many cases. One of these concrete interpretations, which I actually used to do my computations, says that under some technical hypotheses, the invariant of $K$ is the class number of a certain quartic field assocated to $K$ by class field theory, modulo 2.

Using the computable description of the invariant, I computed it for 100000 fields. Of the first 100000 $C_3$ fields $K$ of prime conductor with $\operatorname{Cl}_2(K) \cong A$, nearly 54% of them have invariant 1. The discrepancy here is surprising. If the fields were to equidistribute among the two classes completely randomly, then we could model the invariants by flipping fair and independent coins. If we do that, we expect a mean of 50000 heads (invariant 1, say), with an expected deviation from the mean of around $100000^{1/2}(2\pi)^{-1/2} \approx 126$. A deviation of nearly 4000 is statistically absurd. Hence, we suspect that, while the Schur multiplier modification is a fairly good explanation for Malle's computations, this is not the end of the story by any means!

There are at least two natural explanations for the discrepancy here: one is that the invariant genuinely does not equidistribute across the two classes, and the other is that, for a large initial sequence of number fields, there is a sizable secondary term which takes a long time to fade away. The latter behavior is quite common in questions about counting number fields; perhaps most notably, the error term in the number of cubic fields is so large that many people doubted the celebrated Davenport-Heilbronn Theorem [DH71] after it had been proven. This discrepancy led to much further work, most recently by Bhargava-Shankar-Tsimerman [BST10], Hough [Hou10], and Taniguchi-Thorne [TT11]. I suspect that, if we consider $C_3$ fields $K$ for which the Sylow 2-subgroup of the class group of $K$ is $C_2 \times C_2$ with prime conductor up to $N$, then the proportion with invariant 1 is $1/2 + aN^{-1/6}$, for some constant $a$. This proposed secondary term, $aN^{-1/6}$, is both consistent with numerical data and looks familiar from the secondary term for counting cubic fields. I am interested in developing an explanation for this secondary term.

## 2. Origamis and origami curves

Roberts in [Rob04] has demonstrated that branched covers of algebraic curves can be used to construct number fields with limited ramification and large Galois group. Let us begin with an example of this phenomenon. Let

$$f(x) = -\frac{7}{221184}\frac{(343x^4 - 294x^2 - 224x + 159)^2}{x-1} = \frac{f_1(x)}{f_2(x)}.$$

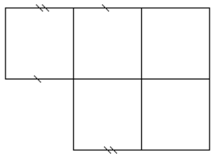
We can check that $f$ defines a degree-8 map $\mathbb{P}^1 \to \mathbb{P}^1$, branched only above 0, 1, and $\infty$. We can also use it to construct an extension $\mathbb{Q}(x)/\mathbb{Q}(t)$ of function fields, defined by $t \mapsto f(x)$, and by specialization, we can construct an extension of $\mathbb{Q}$. Let $g_t(x) = f_1(x) - tf_2(x)$. Then for most specializations of $t$, $\mathbb{Q}[x]/(g_t(x))$ is a number field with Galois group $\mathrm{PGL}_2(\mathbb{F}_7)$, ramified only above 2, 3, 7, $\infty$, and the primes dividing $t$ and $t+1$. Without the explicit map $f$, it is not immediately clear that such a number field must exist, although there are other constructions as well.

Since branched covers of $\mathbb{P}^1$ are useful for constructing number fields with limited ramification, we might expect that branched covers of other curves are similarly useful, and indeed they are. So, another one of my projects has focused on constructing branched covers of elliptic curves, and in particular, origamis. An origami is a smooth connected projective curve $C$, together with a map $f : C \to E$ to an elliptic curve, branched only above one point in $E$. If $C$ has genus $g$, we can extend $f$ to a family of maps from curves of genus $g$ to elliptic curve, i.e., we can find a curve $C(t)$ inside the product $\mathcal{M}_g \times \mathcal{M}_{1,1}$ of the moduli space of genus-$g$ curves and the moduli space of elliptic curves (or, the normalization of a possibly singular curve in $\mathcal{M}_g$), so that there is a map $f_t : C(t) \to \mathcal{M}_{1,1} = \mathbb{A}^1$ of algebraic curves, and so that $f_0 = f : C \to E$. It is then possible to compactify $C(t)$ and $\mathbb{A}^1$ to obtain a map from a projective curve inside $\overline{\mathcal{M}}_g \times \mathbb{P}^1$ to $\mathbb{P}^1$. We call such a curve in $\overline{\mathcal{M}}_g \times \mathbb{P}^1$ an origami curve.

From a topological perspective, it is quite easy to see what a (complex) origami is; indeed, a substantial amount of work has been done in this direction, especially by Herrlich and Schmithüsen, for instance in [HS09], and their students. We represent an elliptic curve $E$ as $\mathbb{C}/\Lambda$ for some lattice $\Lambda$, and we let $X \subset \mathbb{C}$ be any fundamental parallelogram for $\Lambda$. Then an origami $O$ is a finite union of translates of $X$ by vectors in $\Lambda$, together with some edge identifications so that

(1) Every left edge is identified with a unique right edge, and vice versa.
(2) Every top edge is identified with a unique bottom edge, and vice versa.
(3) The resulting space is connected.

For ease of drawing, we usually think of $E = \mathbb{C}/\mathbb{Z}[i]$, with $X$ the square with vertices $\{0, 1, i, 1+i\}$. So, the following diagram is an example of a topological origami:



We use marks on the edges to denote edge identification; unmarked edges are identified with opposite edges. From such diagrams, we can easily recover much combinatorial and topological information about the curve $C$; for example, in the pictured origami, we can quickly observe that the genus is 2, and that the monodromy group is $S_5$.

The algebraic perspective, however, is much more mysterious at the moment. Here, we are interested in determining an algebraic equation for our curve $C$, together with an algebraic morphism $f : C \to E$; by the Riemann Existence Theorem, we know that this is always possible. Although there has been much interest in writing down such algebraic equations, as is mentioned, for instance, in [Oss05], it is not known how to do this in any reasonable generality.

My work in this area has focused on finding methods for constructing these algebraic equations. Using degeneration techniques, I was able to construct an infinite family of hyperelliptic origami curves, one for each genus $g \geq 2$. In this family, the local monodromy is always maximal, so the cover is totally ramified. That is, for a genus $g$ curve, I construct a genus $g$ curve $C$ above each elliptic curve $E$ so that $f : C \to E$ has degree $2g-1$, and so that the local monodromy above the branch point is cyclic of order $2g-1$. Here is one example I came up with using a degeneration method:

**Proposition 2.** *The genus-3 curve*

$$C : y^2 = x^7 + 51x^6 + 450x^5 + 15200x^4 + 2400x^3 + 1792x^2 + 512x \tag{1}$$

*maps to the elliptic curve*

$$E : y^2 = x(x+1)(x+2)$$

*via the map*

$$(x, y) \mapsto \left( \frac{x^5}{(5x^2+20x+16)^2}, \frac{x^3(x^2+12x+16)y}{(5x^2+20x+16)^3} \right).$$

*This map is ramified only above one point, namely the point* $(0,0) \in E$*.*

I can also describe a similar genus-3 curve, as well as hyperelliptic curves of other genera, above each elliptic curve. Details can be found in [RS12c].

It is more difficult to write down an explicit equation for a non-totally ramified origami. One low-degree case to consider is that of a genus-2 curve admitting a degree-5 map to an elliptic curve branched at one point, so that there is a (unique) triple point above the branch point, as well as two unbranched points. In order to construct such a genus-2 curve, we fix an elliptic curve $E$, which I take to be the $j = 1728$ elliptic curve $y^2 = x^3 - x$. We can also make several normalization conditions on the genus-2 curve: since all genus-2 curves are hyperelliptic, we can write it in the form $C : y^2 = x(x-1)(x-a)(x-b)(x-c)$, and we normalize the map $f : C \to E$ by demanding that if $\omega = \frac{dx}{y} \in \Omega^1_E$ is an invariant differential, then $f^*\omega$ is a constant multiple of $\frac{dx}{y} \in \Omega^1_C$. (This is equivalent to putting the branch point at $\infty$ and saying that the point at $\infty$ on $C$ is the triple point.) Then, for $\gamma \in H_1(C, \mathbb{Z})$, we have

$$\int_\gamma f^*\omega = \int_{f_*\gamma} \omega.$$

From the origami diagram, we can determine the images of the homology classes of $C$ under $f$, so it is necessary to find $a$, $b$, and $c$ so that the periods satisfy certain relations.

By computing the periods to high precision, we can determine $a$, $b$, and $c$ to high precision as well. Since we know *a priori* that $C$ is defined over $\overline{\mathbb{Q}}$, we can use the LLL algorithm to conjecture their exact values as algebraic numbers, and then it is a straightforward check to see if we are right.

In this manner, I was able to deduce the following in [RS12b]:

**Theorem 3.** *the genus-2 curve*

$$C : y^2 = x(x-1)(x-\alpha)(x-2\alpha+1)(x-2\alpha), \qquad \alpha = 81 + 36\sqrt{5}$$

*admits a degree-5 map to the $j = 1728$ elliptic curve with monodromy type a 3-cycle.*

I have also used the same method to compute several more genus-2 curves that admit maps to elliptic curves, with various degrees and ramification types.

We would also like to understand the full origami curve which contains $f$. That is, we can ask for equations of the corresponding genus-2 curves which map to elliptic curves with other $j$-invariants, and we can attempt to answer that question by a degeneration procedure similar to that of the totally ramified case. Now, we start with the one $f : C \to E$ we already know and attempt to deform it into a map $f_t : C_t \to E_t$, where $f_0 : C_0 \to E_0$ is $f : C \to E$. By computing derivatives of the periods to high precision and then algebraizing as before, we can work out several terms of the power series expansions for the coefficients of $C_t$ and $f_t$. So far, I have been able to compute the expansions of the coefficients of $C_t$ up to the quadratic term, and I would like to compute enough terms to algebraize the entire family, but at the moment, the numerical computations for the derivatives of the periods is prohibitively slow. Thus, I am attempting to optimize the procedure.

However, even without computing explicit equations, I have an algorithm to determine some aspects of the global structure of the origami curve, such as the degree of the structure map to $\mathbb{P}^1$ only by looking at at the origami diagram. In this case, the structure map has degree 9. I believe that my algorithm actually detects more structure, and I am attempting to learn how to access it.

Using origamis and origami curves, we can construct number fields of limited ramification, in one of several ways. One method is to construct a Belyĭ function out of it: an origami curve gives a branched cover of $\mathbb{P}^1$, and in fact, it is unramified away from 0, 1728, and $\infty$. By work of Beckmann (see [Bec89]) and Roberts (see [Rob04]), we can construct number fields unramified away from 2, 3, and the primes dividing the order of the monodromy group of this cover, at worst. Recent work of Roberts and Venkatesh [RV12] along these lines has yielded several number fields of this type, including over 10000 fields with Galois group $S_{25}$ or $A_{25}$ unramified away from $\{2, 3, 5, \infty\}$.

## 3. Permutation rational functions

I have recently become interested in point counting on curves over finite fields. In particular, I am studying the following question:

**Question 4.** *Let $f : C \to D$ be a map between curves over $\mathbb{Q}$. We can reduce $f$, $C$, and $D$ modulo a prime $p$ (for all but finitely many primes $p$) to obtain a map $\widetilde{f} : \widetilde{C} \to \widetilde{D}$ over $\mathbb{F}_p$. Under what conditions does $\widetilde{f}$ give a bijection between the $\mathbb{F}_p$-valued points of $\widetilde{C}$ and $\widetilde{D}$?*

This question is of interest to me due to some surprising answers in the case in which $D = E$ is the elliptic curve $y^2 = x(x + 1)(x + 2)$ and $C$ is the genus-3 curve in (1). More generally, the infinite family of examples I constructed in a similar manner gives rise to numerous such examples.

If $C$ and $D$ are both $\mathbb{P}^1$ and $f \in \mathbb{Q}(x)$, then the problem has been studied due to its potential applications to cryptography: a rational function $f : \mathbb{P}^1(\mathbb{F}_p) \to \mathbb{P}^1(\mathbb{F}_p)$ is called a permutation rational function if $f$ gives a bijection of $\mathbb{P}^1(\mathbb{F}_p)$ to itself. Kayal in [Kay05] presents an algorithm to determine if a given rational function is a permutation rational function.

The question of whether a rational function over $\mathbb{F}_p$ is a permutation rational function has an interpretation in terms of dynamical systems, namely that a rational function $f$ defined over $\mathbb{F}_p$ is a permutation rational function if and only if every point in $\mathbb{P}^1(\mathbb{F}_p)$ is a periodic point of $f$.

Combining my interest in Question 4 and in permutation rational functions, I have been led to consider the following:

**Question 5.** *Given a rational function $f : \mathbb{P}^1(\mathbb{Q}) \to \mathbb{P}^1(\mathbb{Q})$ which is injective on $\mathbb{P}^1(\mathbb{Q})$, is it true that $f$ modulo $p$ becomes a permutation rational function for infinitely many primes $p$?*

The answer to Question 5 is no in general, so we need further criteria to detect rational functions over $\mathbb{Q}$ which produce many permutation rational functions modulo different primes $p$. I am curious to find which other obstructions exist that prevent an injective function over $\mathbb{Q}$ from being a permutation rational function modulo $p$, and I am currently working on finding them.

## References

[Bec89] Sybilla Beckmann. Ramified primes in the field of moduli of branched coverings of curves. *J. Algebra*, 125(1):236–255, 1989.

[BST10] Manjul Bhargava, Arul Shankar, and Jacob Tsimerman. On the Davenport-Heilbronn theorem and second order terms. 2010. `http://arxiv.org/abs/1005.0672`.

[CL84] H. Cohen and H. W. Lenstra, Jr. Heuristics on class groups of number fields. In *Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983)*, volume 1068 of *Lecture Notes in Math.*, pages 33–62. Springer, Berlin, 1984.

[CM87] H. Cohen and J. Martinet. Class groups of number fields: numerical heuristics. *Math. Comp.*, 48(177):123–137, 1987.

[CM90] Henri Cohen and Jacques Martinet. Étude heuristique des groupes de classes des corps de nombres. *J. Reine Angew. Math.*, 404:39–76, 1990.

[DH71] H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London Ser. A*, 322(1551):405–420, 1971.

[Hou10] Bob Hough. Average equidistribution of Heegner points associated to the 3-part of the class group of imaginary quadratic fields. 2010. `http://arxiv.org/abs/1005.1458`.

[HS09] Frank Herrlich and Gabriela Schmithüsen. Dessins d'enfants and origami curves. In *Handbook of Teichmüller theory. Vol. II*, volume 13 of *IRMA Lect. Math. Theor. Phys.*, pages 767–809. Eur. Math. Soc., Zürich, 2009.

[Kay05] Neeraj Kayal. Recognizing permutation functions in polynomial time. *Electronic Colloquium on Computational Complexity (ECCC)*, (008), 2005.

[Mal08] Gunter Malle. Cohen-Lenstra heuristic and roots of unity. *J. Number Theory*, 128(10):2823–2835, 2008.

[Oss05] Brian Osserman. Two degeneration techniques for maps of curves. In *Snowbird lectures in algebraic geometry*, volume 388 of *Contemp. Math.*, pages 137–143. Amer. Math. Soc., Providence, RI, 2005.

[Rob04] David P. Roberts. An *ABC* construction of number fields. In *Number theory*, volume 36 of *CRM Proc. Lecture Notes*, pages 237–267. Amer. Math. Soc., Providence, RI, 2004.

[RS12a] Simon Rubinstein-Salzedo. Invariants for $A_4$ fields and the Cohen-Lenstra heuristics. 2012. Preprint at `http://arxiv.org/abs/1210.2773`.

[RS12b] Simon Rubinstein-Salzedo. Period computations for covers of elliptic curves. 2012. Preprint at `http://arxiv.org/abs/1210.4721`.

[RS12c] Simon Rubinstein-Salzedo. Totally ramified branched covers of elliptic curves. 2012. Preprint at `http://arxiv.org/abs/1210.3195`.

[RV12] David P. Roberts and Akshay Venkatesh. Hurwitz number fields. 2012. In preparation.

[TT11] Takashi Taniguchi and Frank Thorne. Secondary terms in counting functions for cubic fields. 2011. `http://arxiv.org/abs/1102.2914`.

[VE10] Akshay Venkatesh and Jordan S. Ellenberg. Statistics of number fields and function fields. In *Proceedings of the International Congress of Mathematicians. Volume II*, pages 383–402, New Delhi, 2010. Hindustan Book Agency.

Department of Mathematics, Dartmouth College, Hanover, NH 03755
*E-mail address*: simon.rubinstein-salzedo@dartmouth.edu