ABELIAN CHABAUTY

DAVID ZUREICK-BROWN

Abstract. These are expanded lecture notes for a series of four lectures at the Arizona Winter School on “Nonabelian Chabauty”, held March 7-11, 2020 in Tucson, Arizona.

Last update: March 1, 2020

Contents

1. Introduction 1
   1.1. Course outline, and how to read these notes 2
   1.2. Acknowledgements 4
2. Abelian Chabauty 4
   2.2. Computational aspects: an exercise 5
3. The uniformity conjecture 6
   3.5. Evidence and records 6
   3.6. Chabauty–Coleman bounds 7
   3.8. A new hope 7
4. Bad Reduction 7
   4.2. A few examples 8
5. Rank Favorable bounds 9
   5.3. Stoll’s proof of Theorem 5.1 10
   5.4. The rank of a divisor 12
   5.6. Chip firing and the rank of a divisor on a graph 13
   5.12. Semicontinuity of specialization 14
   5.14. Rank favorable bounds for curves with totally degenerate reduction 15
   5.15. Refined ranks 16
6. Tropical Geometry, Berkovich spaces, and Chabauty 16

References 16

1. Introduction

Let $K$ be a number field and let $X/K$ be a nice\footnote{smooth, projective, and geometrically integral} curve of genus $g > 1$ whose Jacobian has rank $r := \text{rank } \text{Jac}_X(K)$.

The Method of Chabauty–Coleman (alternatively: “Chabauty’s method”, “Abelian Chabauty”, or just plain, vanilla, “Chabauty”) is among the most successful and widely

\textit{Date:} March 1, 2020.
applicable techniques for analyzing (either theoretically or explicitly) the set $X(K)$ of $K$-points of a low rank ($r < g$) curve $X$, and is an essential part of the “explicit approaches to rational points” toolbox. In particular, with some luck, Chabauty’s method allows one to

- explicitly determine, with proof, the set $X(K)$, or
- determine an upper bound on $\# X(K)$.

Mordell conjectured in the 20’s that the set $X(K)$ of $K$-points of $X$ was finite. This was famously proved in the 80’s by Faltings [Fal86] (for which he was awarded a Fields medal), with subsequent independent proofs by Vojta and Bombieri [Voj91, Bom90].

Chabauty [Cha41], building on a idea of Skolem [Sko34], gave the first substantial progress toward Mordell’s conjecture. Chabauty’s method, which used $p$-adic techniques to produce $p$-adic “locally analytic” functions, relies on the hypothesis that $r < g$. This technique sat somewhat dormant until the 80’s, when Coleman’s seminal paper [Col85] resurrected and improved Chabauty’s idea.

**Machine Computation.** Computer aided computational tools took some time to catch up. The first big bottleneck was to improve the techniques for compute ranks of Jacobians of curves. To execute Chabauty’s method for a particular, explicit curve, one needs to know the rank $r$ of its Jacobian (to check the $r < g$ condition), and a basis for the Mordell–Weil group of $\text{Jac}(K)$ (or at least a finite index subgroup). An early highlight is due to Gordon and Grant [GG93]; building on work of Cassels and Flynn [Fly90, Cas89, Cas83] they work out the special case of two-descent on the Jacobian of a genus 2 hyperelliptic curve with rational Weierstrass points, and (with the help of a SUN Sparcstation) provably compute the rank of a couple examples.

This ushered in a golden era of computer assisted approaches to rational points on curves (and higher dimensional varieties). Even today there are substantial conceptual and practical improvements; one recent highlight is [BPS16], which greatly expands our abilities to compute ranks of Jacobians of non-hyperelliptic curves.

**Applications.** It is worth highlighting a few of the numerous applications of Chabauty’s method.

- McCallum made substantial progress towards a proof of Fermat’s Last Theorem, via a careful study of certain quotients of the Fermat curve $x^p + y^p = z^p$; for instance, [McC94] proves the “second case” of Fermats Last Theorem for regular primes;
- arithmetic statistics: Poonen and Stoll use Chabauty’s method to prove that 100% of odd degree hyperelliptic curves have only one rational point [PS14];
- analysis of rational points on modular curves, and Mazur’s “Program B” [RZB15];
- resolution of various generalized Fermat equations [PSS07].

1.1. **Course outline, and how to read these notes.** There is already an excellent survey by McCallum and Poonen [MP12]. This is short (16 pages) and a great entry point. I recommend my survey with Katz and Rabinoff [KRZB16] for the connections between $p$-adic and tropical techniques, and the survey [BJ15] by Baker and Jensen for a more geometric and combinatorial perspective. I also recommend attempting the computational Exercise 2.3 below (even if one is ultimately most interested in theory).
These notes will focus on the ideas from \cite{KZB13} and \cite{KRZB}, and the papers \cite{LT02,Sto06,Sto19,Bak08} which inspired our work. In particular, my discussion of the foundations of Abelian Chabauty, and discussion of tropical techniques, will be abridged, and these notes are somewhat of an advertisement for \cite{MP12} and \cite{KRZB16}.

Additionally, while reading these notes, we also recommend attempting Exercise 2.3 from Subsection 2.2 which will help to quickly come up with speed with how to perform Chabauty’s method in Magma.

**Abelian Chabauty.** We will start with a ‘black box’ discussion of the method of Chabauty and Coleman, addressing various points of view; this section is mostly an abridged version of \cite{MP12}, and it is recommended to read their survey alongside this section and before reading future sections.

Exhibiting Abelian Chabauty as a special case of Nonabelian Chabauty is not completely straightforward. These notes do not address this, and we instead recommend Poonen’s excellent set of notes available at

\url{http://www-math.mit.edu/~poonen/papers/p-adic_approach.pdf}

**Bad reduction.** One avenue to improve on Coleman’s bound is to generalize the framework of Chabauty and Coleman’ arguments to the case of bad reduction. We will discuss the advantage of working at bad primes and the difficulties and tradeoffs that arise, starting with the work of Lorenzini–Tucker \cite{LT02}.

**Rank favorable bounds.** When the rank is strictly smaller than $g - 1$, there are more “inputs” to Chabauty’s method and one expects this extra flexibility to lead to improvements to the method, giving rise to “rank favorable” bounds. We’ll discuss the setup, and the translation to the notion of a “rank” of a divisor (due to Stoll \cite{Sto06}). For a curve with good reduction, this notion of rank will be the classical one, and the improved bounds will follow from Clifford’s Theorem \cite[Theorem IV.5.4]{Har77}. In the case of bad reduction, reducible reduction, ranks are no longer as well behaved; instead, we introduce Baker’s notion of “numerical rank” \cite{Bak08} and explain how to repair Stoll’s argument in the special case of a curve with totally degenerate reduction.

**Tropical Geometry and Berkovich spaces.** For a curve with bad reduction at a prime $p$, it had been well understood that “monodromy” and “analytic continuation” of $p$-adic integrals was an issue. Coleman proved that in the case of good reduction, there is no “monodromy” and the various ways of analytically continuing $p$-adic integrals all coincide. In the case of bad reduction, they generally do not coincide (we will discuss a simple example which illustrates this).

Stoll \cite{Sto19} discovered that, while choices of analytic continuation genuinely do differ, they do so in a fairly controlled manner (linear, even), and was able to exploit this to prove a uniformity result for *hyperelliptic* curves of small rank.

These results all argue in the framework of rigid geometry (in the sense of Tate). Great clarification arose from the systematic reformulation via Berkovich spaces, which fill in the “missing” points of rigid spaces and which, at least in the case of curves, are fairly concrete and manageable topological spaces (they’re even Hausdorff). I’ll discuss Chabauty in the setting of Berkovich and tropical geometry and explain how modern tools (e.g., Berkovich’s...}
contraction theorem and Thuillier’s slope formula, exposited in [BPR13] give a clean explanation of Coleman’s “good reduction” theorem, and will discuss my work with Katz and Rabinoff [KRZB] which give uniform bounds for arbitrary (but still small rank) curves.

1.2. Acknowledgements. The author would like to thank Enis Kaya, Jackson Morrow, Dino Lorenzini, and John Voight for useful comments and/or discussions. The author was supported by National Science Foundation CAREER award DMS-1555048.

2. Abelian Chabauty

We give here a quick “black box” version of Chabauty’s method, broken into 3 parts: setup, local analysis, and global coordination. We refer the reader to the excellent survey [MP12] for a more detailed introduction.

As before, let $K$ be a number field with ring of integers $\mathcal{O}_K$ and let $X/K$ be a nice curve of genus $g > 1$. Let $r := \text{rank} \text{Jac}_X(K)$ be the rank of the Jacobian of $X$. Fix a prime $p$ and a prime $\mathfrak{p}$ of $\mathcal{O}_K$ above $p$.

**Setup.** Under the assumption $r < g$, there exist locally analytic functions $f_\omega$ on $X(K_\mathfrak{p})$ (arising as a $p$-adic integral of a differential $\omega$) which vanish on $X(K)$, but not on $X(K_\mathfrak{p})$. More precisely, there exists a subspace $V \subset H^0(X_{K_\mathfrak{p}}, \Omega^1_X)$ such that $\dim_{K_\mathfrak{p}} V \geq g - r$, and with the property that the $p$-adic integral

$$\int_P^Q \omega$$

vanishes for all $P, Q \in X(K)$ and for all $\omega \in V$. We will frequently refer to $V$ as $V_{\text{chab}}$. This is (more or less) enough to conclude finiteness (and is roughly the original argument of Chabauty [Cha41]). See [MP12, Section 4 and Subsection 5.4] for proofs of these statements (culminating in [MP12, Theorem 4.4]).

**Local analysis.** On (residue) discs, the integrals $f_\omega$ are “locally analytic”: they have ($p$-adic) power series expansions, a discrete set of zeroes, and are amenable to fairly explicit study via tools from $p$-adic analysis (Newton polygons, or in more complicated situations, tropical geometry). In Coleman’s original analysis ([Col85 Lemma 3] or [MP12 Lemma 5.1]), one can bound the number zeros of $f_\omega$ in a residue disc in terms of the zeroes of its “derivative”, which we summarize as a ‘$p$-adic Rolle’s theorem’ (in the sense of freshman calculus). In the simplest case one gets Rolle’s theorem on the nose: for $K = \mathbb{Q}$ and $p > 2$, Coleman proves [MP12 Theorem 5.3(1)] that the number of zeroes of $f_\omega$ in a residue disc $D_P$ is at most $1 + n_P$, where

$$n_P = \# (\text{div} \omega \cap D_P).$$

See [MP12 Section 5] for proofs of these statements (culminating in [MP12 Theorem 5.5]).

**Remark 2.1.** An exciting “modern” version of this argument is [BD19 Section 4], where they compare the divisor of a locally analytic function $F$ to the divisor of $\mathcal{D}(F)$, where $\mathcal{D}$ is a “nice” differential operator $\mathcal{D}$.

**Global coordination.** One needs some way to coordinate the different, a priori independent, local bounds (as in Equation 2.0.1), and typically exploits some type of “global”
theorem from the geometry of curves. In Coleman’s proof, Riemann–Roch \cite{Har77} Theorem IV.1.3 suffices; the local bounds (under the $K = \mathbb{Q}$ and $p > 2$ hypotheses) are $1 + n_P$; by Equation 2.0.1 we have that
\[
\sum_{P \in X(\mathbb{F}_p)} n_P = \sum_{P \in X(\mathbb{F}_p)} \# (\text{div } \omega \cap D_P) \leq \deg \text{div } \omega = 2g - 2,
\]
which suffices to prove Coleman’s theorem:
\[
\# X(\mathbb{Q}) \leq \sum_{P \in X(\mathbb{F}_p)} (1 + n_P) = \sum_{P \in X(\mathbb{F}_p)} 1 + \sum_{P \in X(\mathbb{F}_p)} n_P \leq X(\mathbb{F}_p) + 2g - 2.
\]

In the “improvements” to this theorem that we discuss in these notes, one instead needs some other global theorem, e.g., Clifford’s theorem (or Riemann–Roch and Clifford’s theorem for graphs, or for arithmetic curves, or for other refined rank functions). In \cite{KRZB}, which uses (in a sense) the full power of the tools from tropical geometry, this step relies global information about sections of the “tropical canonical bundle” (see \cite{KRZB, Lemma 4.15}).

Again, please see \cite{MP12} (especially the detailed examples in Section 8) for a survey and a more thorough introduction. It is also very useful to attempt the Magma exercise (Exercise 2.3) described in the “Computational aspects” part of Subsection 1.1.

2.2. Computational aspects: an exercise. While reading these notes, we also recommend attempting Exercise 2.3 below, which will help to quickly come up with how to perform Chabauty’s method in Magma.

Magma has a free, limited use online calculator here
\url{http://magma.maths.usyd.edu.au/calc/}
and a thoroughly documented implementation of Chabauty’s method
\url{http://magma.maths.usyd.edu.au/magma/handbook/text/1533}.

Even better is to obtain a copy for your laptop, or ssh access to a departmental server with a copy of Magma. The Simons Foundation has graciously made Magma freely available to mathematicians working in the US
\url{http://magma.maths.usyd.edu.au/magma/ordering/};
your department’s tech staff should be able to help you obtain a copy of Magma through this agreement.

Exercise 2.3. Take Smart’s list (from \cite{Sma97}) of the 427 genus 2 curves with good reduction away from 2, and provably find all of the rational points on them. A temporary folder containing several references, and containing a subfolder titled “preparatory-Magma-exercise” with instructions for this exercise, is available at
\url{http://www.math.emory.edu/~dzb/AWS2020}.

As an entry point to some of the additional computational techniques one might need (such as étale descent), we recommend Poonen’s surveys \cite{Poo96} and \cite{Poo02}.
3. The uniformity conjecture

The uniformity conjecture is one of the outstanding open conjectures in arithmetic and diophantine geometry. Initially, Mazur asked whether one can bound \#X(K) purely in terms of the rank of the Jacobian of X (see [Maz00 Page 223] [Maz86 Page 234]). This was later promoted to the following stronger conjecture.

**Conjecture 3.1** ([CHM97]). Let K be a number field and let \( g \geq 2 \) be an integer. There exists a constant \( B_g(K) \) such that for every smooth curve \( X \) over \( K \) of genus \( g \), the number \( \#X(K) \) of \( K \)-rational points is at most \( B_g(K) \).

The uniformity conjecture famously follows [CHM97 Theorem 1.1] from the Weak Lang conjecture (a higher dimension analogue of the Mordell conjecture), which is the following.

**Conjecture 3.2** ([Lan74], 1.3; see also [Lan86]). Let \( X \) be a smooth proper variety of general type over a number field \( K \). Then there exists a proper closed subscheme \( Z \) of \( X \) such that \( X(K) = Z(K) \).

Alternatively, there are the following stronger pair of conjectures.

**Conjecture 3.3** (Generic Uniform Boundedness [CHM97]). Let \( g \geq 2 \) be an integer. There exists a constant \( B_g \) such that for number field \( K \), there exist only finitely many isomorphism classes of curves of genus \( g \) and over \( K \) such that \( \#X(K) > B_g \).

This follows from the Strong Lang Conjecture.

**Conjecture 3.4** ([Lan74], 1.3; see also [Lan86]). Let \( X \) be a smooth proper variety of general type over a number field \( K \). Then there exists a proper closed subscheme \( Z \) of \( X \) such that for every finite extension \( K \subset L \), the complement \( X(L) - Z(L) \) is finite.

In [CHM97], one applies the Weak (or Strong) Lang Conjecture to symmetric powers of the universal curve \( C \to \overline{M}_{g,n} \). A major aspect of their proof is to show that large enough symmetric powers of \( C \) are of general type (or at least dominate a variety of general type); this is a special case of their “correlation” theorem [CHM97 Theorem 1.2]. See the papers [Pac97][Pac99][Abr97][Abr95][Cap95][CHM95] for improvements, variants and a lot of additional discussion, and the slides [http://www-math.mit.edu/~poonen/slides/uniformboundedness.pdf](http://www-math.mit.edu/~poonen/slides/uniformboundedness.pdf) for a fairly recent discussion and some additional motivation.

3.5. Evidence and records. The following table (taken from [Cap95 Section 4]) gives the best known lower bounds on the constant \( B_g(\mathbb{Q}) \).

<table>
<thead>
<tr>
<th>( g )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>45</th>
<th>( g )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_g(\mathbb{Q}) )</td>
<td>642</td>
<td>112</td>
<td>126</td>
<td>132</td>
<td>192</td>
<td>781</td>
<td>( 16(g + 1) )</td>
</tr>
</tbody>
</table>

The record so far is due to Michael Stoll, who found (searching systematically through several families of curves constructed by Noam Elkies) the following:

\[
y^2 = 82342800x^6 - 470135160x^5 + 52485681x^4 + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600
\]

It has at least 642 rational points, and rank at most 22. See [http://www.mathe2.uni-bayreuth.de/stoll/recordcurve.html](http://www.mathe2.uni-bayreuth.de/stoll/recordcurve.html)
for a full list of the known points.

The families constructed by Elkies arise in the following way: he studied K3 surfaces of the form

\[ y^2 = S(t, u, v) \]

with lots of rational lines, such that \( S \) restricted to such a line is a perfect square.

### 3.6. Chabauty–Coleman bounds

The proofs of Mordell due to Faltings, Vojta, and Bombieri \([\text{Fal97}, \text{Voj91}, \text{Bom90}]\) give upper bounds on \( \#X(K) \). These bounds tend to be astronomical, and are not explicit in their original proofs; moreover, it is unclear (to me) how they depend on \( X \) and \( K \).

One application of Chabauty’s method is to give uniform bounds on small rank curves. Coleman’s original theorem is the following.

**Theorem 3.7** (Coleman, \([\text{Col85}]\)). Let \( X/\mathbb{Q} \) be a curve of genus \( g \) and let \( r = \text{rank}_\mathbb{Z} \text{Jac}_X(\mathbb{Q}) \). Suppose \( p > 2g \) is a prime of good reduction. Suppose \( r < g \). Then

\[ \#X(\mathbb{Q}) \leq \#X(\mathbb{F}_p) + 2g - 2. \]

See \([\text{Col85}, \text{Lemma 3}]\) or \([\text{MP12, Theorem 5.3}]\) for a proof.

Various authors have worked to weaken Coleman’s hypotheses and to improve the bound; see Theorems 4.1, 5.1, 5.2, 6.1, and 6.2 below and the surrounding discussion.

### 3.8. A new hope

The recent work of Dimitrov, Gao and Habegger \([\text{DGH19}, \text{DGH20}]\) give bounds on \( \#X(K) \) which only depend on \( g(X) \), \( \deg K \), and \( \text{rank} \text{Jac}_X(K) \). Combined with the conjectural boundedness of ranks of Jacobians of curves of fixed genus over a fixed number field (as predicted by \([\text{PPVW19}]\) and \([\text{Poo18, Section 4.2}]\)) this would prove uniformity. Their approach is in the spirit of Vojta’s original proof \([\text{Voj91}]\), and relies on their recent other work on improved height bounds.

### 4. Bad Reduction

One avenue to improve on Coleman’s bound is to generalize the Chabauty framework and Coleman’s arguments to the case of bad reduction. We will discuss the advantage of working at bad primes and the difficulties that arise, starting with the work of Lorenzini–Tucker \([\text{LT02}]\).

Coleman’s original bound (3.7) relies on an initial choice of a prime of good reduction. The first such prime could be arbitrarily large (e.g., consider a hyperelliptic curve \( y^2 = f(x) \), and twist it by \( d \), where \( d \) is the product of the first million primes; or, pick singular curves \( X_p \) over \( \mathbb{F}_p \), for the first million primes \( p \), and use the Chinese Remainder Theorem to construct a curve \( X \) over \( \mathbb{Q} \) that \( X_{\mathbb{F}_p} \cong X_p \) for each such prime \( p \)). The problem with this is that the Hasse bound only provides that \( \#X(\mathbb{F}_p) \leq 2g\sqrt{p} + p + 1 \); so as \( p \) increases, Coleman’s bound becomes increasingly worse, and in particular is not uniform.

Whence the appeal of the following theorem of Lorenzini and Tucker.
Theorem 4.1 (Lorenzini, Tucker, [LT02], Corollary 1.11). Suppose $p > 2g$ and let $\mathcal{X}$ be a proper regular model of $X$ over $\mathbb{Z}_p$. Suppose $r < g$. Then

$$\#X(\mathbb{Q}) \leq \#\mathcal{X}^{\text{sm}}_F(\mathbb{F}_p) + 2g - 2$$

where $\mathcal{X}^{\text{sm}}_F$ is the smooth locus of the special fiber $\mathcal{X}^{\text{sm}}_F$.

Recall that a scheme $X$ is regular if for every point $x \in X$, with corresponding maximal ideal $m$ and residue field $k(x)$,

$$\dim k(x)/m^2 = \dim X.$$ 

If $R$ is a DVR with uniformizer $\pi$ and residue field $k$ and $X \to \text{Spec} R$ is a relative curve, then a point $x \in X_k$ is regular if and only if the local equation at $x$ is $yz = \pi$. (By “local equation” we mean the equation for the completion of the étale local ring $\mathcal{O}_{X,x}$.) See the examples in Subsection 4.2, and see [Sil94, Chapter IV] for a leisurely treatment.

The utility of the proper regular model $\mathcal{X}$ of $X$ is that the reduction map

$$r: \mathcal{X}(\mathbb{Q}) \to \mathcal{X}(\mathbb{F}_p)$$

takes values in the smooth locus $\mathcal{X}^{\text{sm}}_F(\mathbb{F}_p)$. In Chabauty’s method, one thus only needs to consider residue classes $r^{-1}(Q)$ of points $Q \in \mathcal{X}^{\text{sm}}_F(\mathbb{F}_p)$. Such residue classes are ($p$-adically analytically) isomorphic to discs; this makes the setup of Chabauty easier, and makes the “local analysis” much easier.

The “$2g - 2$” term in Coleman’s bound (3.7) is derived from Riemann–Roch on $X_{\mathbb{F}_p}$, and is the rationale for the “good reduction” hypothesis. Lorenzini and Tucker recover the $2g - 2$ term via Riemann–Roch on $X_{\mathbb{Q}_p}$ and a more involved $p$-adic analytic argument. A later, alternative proof [MP12, Theorem A.5] instead recovers the $2g - 2$ term via arithmetic intersection theory on $\mathcal{X}$ and adjunction.

The drawback is that $\mathcal{X}_{\mathbb{F}_p}$ could contain an arbitrarily long chain of $\mathbb{P}^1$’s. (For example, if $X$ is an elliptic curve with semistable reduction at $p$ and $v_p(j(X)) = -n$, then $\mathcal{X}_{\mathbb{F}_p}$ is an $n$-gon of $\mathbb{P}^1$’s.) Once again, the size of $\mathcal{X}^{\text{sm}}(\mathbb{F}_p)$ could be arbitrarily large, giving non-uniform bounds.

Stoll [Sto19] had the bold idea to work with a non-regular, minimal model: one contacts each chain of $\mathbb{P}^1$’s into a single node. Such a model is no longer regular, but the number of components is bounded, and the genus of each component is bounded (exercise: verify this). Since the model is no longer regular, rational points no longer reduce to smooth points (exercise: give an example), and might reduce to a node. The residue class of a node is now an annulus (explain in an example). This creates multiple problems: an annulus admits “monodromy” and integrals no longer admit a unique analytic continuation, and local expansions are now laurent, rather than power, series. See Theorem 6.1 below and the surrounding discussion.

4.2. A few examples.

Example 4.3 (A regular model). The relative curve

$$y^2 = (x(x - 1)(x - 2))^3 - 5 = (x(x - 1)(x - 2))^3 \mod 5.$$
is regular at the point \((0,0)\). The local equation at \((0,0)\) analytically looks like \(xy = 5\). Once can see by elementary number theory that no rational point can reduce to \((0,0)\).

**Example 4.4** (Resolving a semistable, but not regular, model). \[
y^2 = (x(x - 1)(x - 2))^3 - 5^4 = (x(x - 1)(x - 2))^3 \mod 5
\]

Now, the local equation at \((0,0)\) looks like \(xy = 5^4\), and \((0,5^2)\) reduces to \((0,0)\). Blowing up along the ideal \((x,y,5)\) gives

and the local equations now look like \(xy = 5^3\) and \(xy = 5\). One of the 2 points is still not regular. After 2 more blowups we get

and now all of the local equations look like \(xy = 5\), giving a regular model.

5. **Rank Favorable bounds**

Lorenzini and Tucker [LT02] ask if one can refine Coleman’s bound (Theorem 3.7) when the rank \(r\) is small (i.e., \(r \leq g - 2\)). This was subsequently answered by Stoll.
Theorem 5.1 (Stoll, [Sto06], Corollary 6.7). With the hypothesis of Theorem 3.7,
\[ \#X(\mathbb{Q}) \leq \#X(\mathbb{F}_p) + 2r. \]

The space of differentials “suitable” for Chabauty’s method has dimension at least \( g - r \). When \( r < g - 1 \), there are thus more “inputs” to Chabauty’s method and this extra flexibility can be exploited to improve the \( p \)-adic analysis. Indeed, Stoll’s idea is: instead of using a single integral, to tailor the choice of integral to each residue class. The additional “global geometric input” is the (classical) notion of “rank of a divisor”, and after translating his setup into this language, improved bounds follow from Clifford’s Theorem.

The application of Clifford’s theorem is also the source of the “good reduction” hypothesis. Using ideas from the “discrete case” of tropical geometry (in particular “chip-firing”), Eric Katz and I generalized Stoll’s theorem to arbitrary reduction types.

Theorem 5.2 (Katz, Zureick-Brown, [KZB13]). Let \( X/\mathbb{Q} \) be a curve of genus \( g \) and let \( r = \text{rank} \text{Jac}_X(\mathbb{Q}) \). Suppose \( p > 2r + 2 \) is a prime, that \( r < g \), and let \( X \) be a proper regular model of \( X \) over \( \mathbb{Z}_p \). Then
\[ \#X(\mathbb{Q}) \leq \#\mathcal{X}_{\mathbb{F}_p}(\mathbb{F}_p) + 2r. \]

Unlike Lorenzini and Tucker’s generalization of Coleman’s theorem, where they replace Coleman’s use of Riemann–Roch on \( X_{\mathbb{F}_p} \) with Riemann–Roch on \( X_{\mathbb{Q}_p} \), it does not seem possible to replace Stoll’s use of Clifford’s theorem on \( X_{\mathbb{F}_p} \) with Clifford’s theorem on \( X_{\mathbb{Q}_p} \). Matt Baker suggested that it might be possible to generalize Stoll’s theorem to curves with bad, totally degenerate reduction (i.e., \( X_{\mathbb{F}_p} \) is a union of rational curves meeting transversely) using ideas from tropical geometry (see the recent survey [BJ15] on tropical geometry and applications), in particular the notion of “chip firing”, Baker’s combinatorial definition of rank, and Baker–Norine’s [BN07] combinatorial Riemann–Roch and Clifford theorems. Baker was correct, and in fact an enrichment of his theory led to the following common generalization of Stoll’s and Lorenzini and Tucker’s theorems.

Baker’s recent work [Bak08] clarifies the relationship between linear systems on curves and on finite graphs. Highlights include a semicontinuity theorem for ranks of linear systems (as one passes from the curve to its dual graph), and graph theoretic analogues of Riemann–Roch and Clifford’s theorem. Baker’s theory works best with totally degenerate curves (i.e. each component is a \( \mathbb{P}^1 \)). Theorem 5.2 requires an enrichment of Baker’s theory if the irreducible components of the reduction have higher genus.

5.3. Stoll’s proof of Theorem 5.1. Let \( p > 2 \). Denote by \( V_{\text{chab}} \) the vector space of all \( \omega \in H^0(X_{\mathbb{Q}_p}, \Omega^1_{X_{\mathbb{Q}_p}/\mathbb{Q}_p}) \) such that \( \int_{P_1} P_2 \omega = 0 \) for all \( P_1, P_2 \in X(\mathbb{Q}) \). Then \( \dim V_{\text{chab}} \geq g - r \).

For each \( \omega \in V_{\text{chab}} \), scale \( \omega \) by a power of \( p \) so that the reduction \( \tilde{\omega} \) of \( \omega \) is non zero, and denote by \( \tilde{V}_{\text{chab}} \) the set of all such reductions; we note that \( \dim_{\mathbb{F}_p} \tilde{V}_{\text{chab}} = \dim_{\mathbb{Q}_p} V_{\text{chab}} \).

For each \( Q \in X(\mathbb{F}_p) \) and \( \omega \in \tilde{V}_{\text{chab}} \), set
\[ n_Q(\omega) := \deg(\text{div} \omega |_{[Q]}) \quad \text{and} \quad n_Q := \min_{\omega \in \tilde{V}_{\text{chab}}} n_Q(\omega) \]
(where we recall that \( [Q] \) denotes the tube or residue class of \( Q \), that is, the set of all points of \( X(\mathbb{Q}_p) \) which reduce to \( Q \)). Since div is compatible with reduction mod \( p \), \( n_Q(\omega) \) is equal to the valuation \( v_Q(\tilde{\omega}) \) (i.e., the order of vanishing of \( \tilde{\omega} \) at \( Q \)).
By the "\(p\)-adic Rolle’s theorem", the number of zeroes of \( \int \omega \) in \( Q \) is at most \( 1 + n_Q(\omega) \), so
\[
X(\mathbb{Q}) \leq \sum_{Q \in X(\mathbb{F}_p)} (1 + n_Q) = \sum_{Q \in X(\mathbb{F}_p)} 1 + \sum_{Q \in X(\mathbb{F}_p)} n_Q = X(\mathbb{F}_p) + \deg (D_{chab}),
\]
where we define \( D_{chab} \) to be the divisor
\[
D_{chab} := \sum_{Q \in X(\mathbb{F}_p)} n_Q Q \in \text{Div} X_{\mathbb{F}_p}.
\]

By Riemann–Roch, \( \deg D_{chab} \leq 2g - 2 \), recovering the bound
\[
X(\mathbb{Q}) \leq X(\mathbb{F}_p) + 2g - 2.
\]

We claim that, in fact, \( \deg D \leq 2r \), which suffices to prove Theorem 5.1. (When \( r = g - 1 \), \( 2r = 2g - 2 \).) Stoll’s main observation is that
\[
\widetilde{V}_{chab} \subset H^0(X_{\mathbb{F}_p}, \Omega^1_{X_{\mathbb{F}_p}, (-D_{chab})}),
\]
and in particular,
\[
\dim H^0(X_{\mathbb{F}_p}, \Omega^1_{X_{\mathbb{F}_p}, (-D_{chab})}) \geq \dim \widetilde{V}_{chab} \geq g - r. \tag{5.3.2}
\]

To justify Equation \( \text{5.3.1} \), given an effective divisor \( E = \sum n_P P \) and a line bundle \( \mathcal{L} \) on a curve \( X \), recall that \( H^0(\mathcal{X}, \mathcal{L}(-E)) \) is the subspace of sections of \( H^0(\mathcal{X}, \mathcal{L}) \) that have at least a zero of order \( n_P \) at \( P \). A differential \( \tilde{\omega} \in \tilde{V}_{chab} \) thus satisfies \( v_P(\omega) \geq n_P \) by definition of \( n_P \! \).

On the other hand, Clifford’s Theorem \([\text{Har} 77 \text{, Theorem IV.5.4}]\) implies that
\[
\dim H^0(X_{\mathbb{F}_p}, \Omega^1_{X_{\mathbb{F}_p}, (-D_{chab})}) \leq \frac{1}{2} \deg (\Omega^1_{X_{\mathbb{F}_p}, (-D_{chab})}) + 1. \tag{5.3.3}
\]

Combining equations \( \text{5.3.2} \) and \( \text{5.3.3} \) gives
\[
g - r \leq \frac{1}{2} \deg (\Omega^1_{X_{\mathbb{F}_p}, Q_p, (-D_{chab})}) + 1 = g - 1 - \frac{1}{2} \deg D_{chab} + 1
\]
and simplifying gives
\[
\deg D_{chab} \leq 2r.
\]

To justify Equation \( \text{5.3.3} \), we switch to the language of divisors. Recall that a divisor is special if \( \dim H^0(X, K - D) > 0 \), where \( K \) is a canonical divisor. (Equivalently, \( D \) is special if and only if it is a subvector of some canonical divisor.) For context: Riemann–Roch reads:
\[
H^0(X, D) = \dim H^0(X, K - D) + \deg D + 1 - g.
\]

This gives a formula for \( H^0(X, D) \) when the degree of \( D \) is large; indeed when \( \deg D > \deg K = 2g - 2 \), \( \deg K - D < 0 \), therefore \( \dim H^0(X, K - D) = 0 \) and \( H^0(X, D) = \deg D + 1 - g \). At the other extreme: if \( \deg D \leq 2g - 2 \), then it is still possible that \( \dim H^0(X, K - D) = 0 \), in which case, again, \( H^0(X, D) = \deg D + 1 - g \). If \( D \) is special, i.e., \( \dim H^0(X, K - D) > 0 \), then by Riemann–Roch, \( H^0(X, D) < \deg D + 1 - g \); in this case, Clifford’s Theorem gives the much stronger bound
\[
H^0(X, D) \leq \frac{1}{2} (\deg D) + 1.
\]

Let \( K = \text{div} \tilde{\omega} \). Then, by definition of \( n_Q \), \( D_{chab} \) is a subvector of the canonical divisor \( K \) (since \( v_Q(D_{chab}) := v_Q(\tilde{\omega}) = n_Q(\omega) \geq n_Q \)); in particular, \( D_{chab} \) is special.
5.4. The rank of a divisor. In the proof of Stoll’s theorem, we implicitly used the notion of rank of a line bundle (or divisor). One can simply define the rank \( r(\mathcal{L}) \) of a divisor \( \mathcal{L} \in \text{Pic}(X) \) to be

\[
 r(\mathcal{L}) := \dim H^0(X, \mathcal{L}) - 1.
\]

This has the following alternative interpretation over an algebraically closed field: \( r(\mathcal{L}) \) is the largest number of independent and generic “vanishing conditions” one can impose on sections of \( \mathcal{L} \). Recall that for a closed point \( P, H^0(X, \mathcal{L}(-P)) \) is the subspace of sections of \( H^0(X, \mathcal{L}) \) that have at least a simple zero at the point \( P \). More generally, for an effective divisor \( E = \sum n_P P, H^0(X, \mathcal{L}(-E)) \) is the subspace of sections of \( H^0(X, \mathcal{L}) \) that have at least a zero of order \( n_p \) at \( P \). By [Har77, Proof of Theorem IV.1.3],

\[
 \dim H^0(X, \mathcal{L}(-P)) \geq \dim H^0(X, \mathcal{L}) - 1,
\]

and in particular,

\[
 \dim H^0(X, \mathcal{L}(-E)) \geq \dim H^0(X, \mathcal{L}) - \deg E. \tag{5.4.1}
\]

Similarly, one can define the rank \( r(D) \) of a divisor \( D \in \text{Div}(X) \) to be

\[
 r(D) := \dim H^0(X, D) - 1.
\]

This has the following alternative interpretation over an algebraically closed field: \( r(D) \) is the largest number of points of \( X_{\overline{k}} \) (allowing for multiplicity) one can remove from \( D \) before \( D \) is no longer equivalent to some effective divisor. Equivalently, the rank is the largest number of points (allowing for multiplicity) that you can demand occurs as a subdivisor of some effective divisor \( D' \) equivalent to \( D \).

More formally, we make the following definitions.

**Definition 5.5.** Let \( D \in \text{Div} X \) be a divisor. The linear system associated to \( D \) is the collection \(|D|\) of effective divisors linearly equivalent to \( D \). We define the rank \( r(D) \) of \( D \) to be -1 if \(|D|\) is empty (i.e., if \( D \) is not equivalent to an effective divisor). Otherwise, we define

\[
 r(D) := \max\{n \in \mathbb{Z}_{\geq 0} : |D| = \emptyset, \forall E \in \text{Div}_{\geq 0}^n(X_{\overline{k}})\},
\]

where \( \text{Div}_{\geq 0}^n(X_{\overline{k}}) \) is the subset of \( \text{Div}(X_{\overline{k}}) \) of effective divisors of degree \( n \).

The linear system \(|D|\) is naturally isomorphic to the projective space \( \text{Proj} H^0(X, D) \cong \mathbb{P}^r - 1 \), where \( r = \dim H^0(X, D) \).

By Equation 5.4.1

\[
 r(D) \geq \dim H^0(X, D) - 1.
\]

The converse follows from the observation (exercise!) that if \( r(D) = n \), then there exists \( P \in X_{\overline{k}} \) such that \( r(D-P) = n-1 \). Note that one must take \( k \) in the definition of rank. Indeed, consider a non hyperelliptic curve \( X \) with \( X(k) = \emptyset \), but \( X(k') \neq \emptyset \) for some quadratic extension \( k' \) of \( k \). Then for \( P \in X(k') \) with conjugate point \( Q, D := P + Q \in \text{Div}^2 X \). Then \( r(D) = 1 \); but, \( \text{Div}_{\geq 0}^1 X \) is empty, so

\[
 \max\{n \in \mathbb{Z}_{\geq 0} : |D| = \emptyset, \forall E \in \text{Div}_{\geq 0}^n(X)\} = 1
\]
5.6. **Chip firing and the rank of a divisor on a graph.** I recommend taking a quick look at Matt Baker’s short expository article available at

http://people.math.gatech.edu/~mbaker/pdf/g4g9.pdf.

For a short selection of other references: the papers [Bak08, BN07] by Baker and collaborators are my preferred starting point; [BJ15] is also a great survey.

Let $\Gamma$ be a connected graph, with vertex set $V(\Gamma)$ and edge set $E(\Gamma)$. We define $\text{Div} \, \Gamma$ to be the set of maps from the vertices of $V(\Gamma)$ to $\mathbb{Z}$; this is isomorphic to the free group $\mathbb{Z}[V(\Gamma)]$ generated by the set of vertices of $\Gamma$, and will sometimes write $D = \sum_v n_v(v)$ for the function that has value $n_v$ at $v$. The degree of $D$ is $\text{deg} \, D = \sum_v D(v)$. We’ll refer to an element $D \in \text{Div} \, \Gamma$ as a divisor or configuration, and will typically represent them visually as in Figure 5.6 and refer to $D(v)$ as the “number of chips” or “dollars” at the vertex $v$.

The goal of the “dollar game” or “chip firing” is to get a divisor out of debt. We say that a divisor $D$ is effective, and write $D \geq 0$, if $D(v) \geq 0$ for all vertices $v$ of $\Gamma$.

One formalizes lending and borrowing as follows: given $f \in \text{Div} \, \Gamma$, we define the principal divisor associated to $f$ to be

$$\operatorname{div} f = \sum_{v \in \Gamma} D(v) \cdot \left( -(\text{deg} \, v)(v) + \sum_{w \neq v} \#\{\text{edges between } w \text{ and } v\}(w) \right).$$

In particular, $\operatorname{div} \delta_v$ is

$$-(\text{deg} \, v)(v) + \sum_{w \neq v} \#\{\text{edges between } w \text{ and } v\}(w)$$

which has the effect of the vertex $v$ “lending” one chip to each adjacent vertex (see Figure 5.6). We say that two divisors $D$ and $D'$ are equivalent if there is a sequence of lends and borrows which transforms $D$ into $D'$, and we define the Jacobian or Picard group $\text{Pic} \, \Gamma$ to be the abelian group of equivalence classes of divisors on $\Gamma$. more formally, there is an exact sequence

$$0 \rightarrow \mathbb{Z} \rightarrow \text{Div} \, \Gamma \xrightarrow{\text{div}} \text{Div} \, \Gamma \rightarrow \text{Pic} \, \Gamma \rightarrow 0,$$

where the first map sends 1 to the function $\sum_v \delta_v$ (i.e., every vertex lends, which has no effect).

The vector space $H^0(\Gamma, D)$ doesn’t make sense for a graph. Baker’s insight from [Bak08] is that the “practical” definition of rank (Definition 5.5) does generalize nicely to divisors on graphs.
Definition 5.7. Let $D \in \text{Div} \Gamma$ be a divisor. The linear system associated to $D$ is the collection $|D|$ of effective divisors linearly equivalent to $D$. We define the rank $r(D)$ of $D$ to be -1 if $|D|$ is empty (i.e., if $D$ is not equivalent to an effective divisor). Otherwise, we define

$$r(D) := \max\{n \in \mathbb{Z}_{\geq 0} : |D - E| \neq \emptyset, \forall E \in \text{Div}^n_{\geq 0}(\Gamma)\},$$

where $\text{Div}^n_{\geq 0}(X_k)$ is the subset of $\text{Div}(X_k)$ of effective divisors of degree $n$.

Equivalently, the rank is the largest number of points (allowing for multiplicity) that you can demand occurs as a subdivisor of some effective divisor $D'$ equivalent to $D$. In other words, the rank of a divisor $D$ is the “amount of damage” necessary to make the chip firing game unwinnable.

Definition 5.8. Let $\Gamma$ be a graph. We define the canonical divisor on $K$ to be the divisor $K_\Gamma := \sum_{v \in V(\Gamma)} (\deg v - 2)v$.

The genus $g(\Gamma)$ (alternatively first Betti number $h^1(\Gamma)$) of $\Gamma$ is the number of edges minus the number of vertices; note that $\deg K_\Gamma = 2g(\Gamma) - 2$. We say that a divisor $D$ is special if $D$ is effective and $|K - D|$ is non empty.

This definition is motivated by adjunction.

Theorem 5.9 ([BN07], Theorem 1.12 and Corollary 3.5). Let $D \in \text{Div} \Gamma$. Then the following are true.

1. Riemann–Roch: $r(D) - r(K - D) = \deg D + 1 - g$.
2. Clifford’s Theorem: if $D$ is special, then $r(D) \leq \frac{1}{2} \deg D$.

Corollary 5.10 ([BN07], Theorem 1.9). The chip firing game is winnable for the configuration $D$ if $\deg D \geq g$.

Remark 5.11. It is unknown whether one can deduce these from the analogous theorems from the geometry of curves.

5.12. Semicontinuity of specialization. Let $R$ be a complete discrete valuation ring with maximal ideal $\pi$, residue field $k$, and fraction field $K$. Denote by $\eta$ the generic point of Spec $R$, and by $b$ the closed point. (Most of what we say below works just as well if we replace Spec $R$ by an integral scheme $B$.) Let $C \to B$ be a relative curve of genus $g$ (i.e., a smooth proper morphism such that for every $x \in \text{Spec} R$, the fiber $C_x$ is a smooth proper curve of genus $g$ over the residue field $k(x)$).

Let $D = \sum n_P P \in \text{Div} C_\eta$. The dimension of $C$ is 2, and we can extend $D$ to a divisor $\mathcal{D}$ on $C$ by taking the closure of its support; in other words, $\mathcal{D} := \sum n_P \overline{P} \in \text{Div} C_\eta$, where $\overline{P}$ is the closure of $P$. Intersecting $\mathcal{D}$ with the special fiber $C_b$ thus gives a specialization map

$$\text{sp}: \text{Div} C_\eta \to \text{Div} C_b.$$ 

Proposition 5.13. Let $D \in \text{Div} C_\eta$. Then $r(\text{sp}(D)) \geq r(D)$.

The inequality can certainly be strict. Indeed, consider $C$ with hyperelliptic special fiber and non-hyperelliptic generic fiber, and let $D = P + Q$ where $P, Q \in C(K)$ are points whose reductions are hyperelliptic conjugate. Then $r(D) = 1$ but $r(\text{sp}(D)) = 2$. 

14
More generally, if $L$ is a line bundle on $C_\eta$, there is a unique (up to isomorphism) extension of $L$ to a line bundle $\mathcal{L}$ on $C$ (i.e., a line bundle $\mathcal{L}$ on $C$ such that $\mathcal{L}_\eta$ is isomorphic to $L$.

(Indeed, let $s \in L(U)$ be a section over some non empty open set $U \subset C_\eta$; then $\mathcal{L} = \mathcal{O}_C(\mathcal{T}/s)$ extends $L$.)

There is thus an analogous specialization map

$$\text{sp}: \text{Pic} C_\eta \to \text{Pic} C_b,$$

and (since $r(D) = r(\mathcal{O}(D)))$, Proposition 5.13 equivalently implies that $r(\text{sp}(L)) \geq r(L)$.

Proposition 5.13 is a special case of Semicontinuity of Cohomology [Har77, Theorem III.12.8] (taking $i = 0$ and $\mathcal{F} = \mathcal{O}_C(D)$). Unsurprisingly, this is overkill; we sketch a direct proof that will generalize to the “discrete” case.

Proof of Proposition 5.13 First, note that a section $s \in H^0(C_\eta, D)$ can be scaled by a power of the uniformizer $\pi$ to a section of $H^0(C, \mathcal{O}(\mathcal{D}))$. There are a few ways to see this: if we think of $s$ as a function

Conversely, Since $C_\eta \subset C$ is open, and since the vanishing locus of a section of a line bundle is closed, the map

$$H^0(C, \mathcal{O}(\mathcal{D})) \to H^0(C_\eta, D)$$

is injective. In particular,

$$\text{rank}_R H^0(C, \mathcal{O}(\mathcal{D})) = \text{rank}_K H^0(C_\eta, D).$$

Alternatively, it follows directly from flatness that

$$H^0(C, \mathcal{O}(\mathcal{D})) \otimes_R K \cong H^0(C_\eta, D).$$

On the other hand, to be continued

□

5.14. Rank favorable bounds for curves with totally degenerate reduction. One can associate to such a singular curve with transverse crossings its dual graph as in Figure 1. Component curves become nodes, and intersections correspond to edges.

In this subsection, I’ll discuss the special case of a “Mumford curve” (i.e., a curve with “totally degenerate” reduction, in that the reduction is a collection of $\mathbb{P}^1$’s meeting transversely, and in particular represents an isolated point on the moduli space of curves); for such curves, one only needs Baker’s original notion of rank (which we call “numerical rank”); for a discussion of the “abelian rank”, see our paper [KZB13].
5.15. **Refined ranks.** In a certain sense, the numerical rank only “sees” the component group of the Néron model. The enriched notion of abelian rank from [KZB13] and [AB15] sees the abelian part of the Néron model. One can ask if there is a corresponding notion of “toric” or “unipotent” rank. In [KZB13, Subsection 3.3], we define a “toric” rank, and in [KZB13, Example 5.5] demonstrate that these ranks differ; we have yet to find a useful application.

### 6. Tropical Geometry, Berkovich spaces, and Chabauty

A recent breakthrough [Sto19] fully removed, in the special case of hyperelliptic curves, the dependence on a regular model and derived a uniform bound on $\#X(\mathbb{Q})$ for small ($r \leq g - 3$) rank curves.

**Theorem 6.1** (Stoll, [Sto19]). Let $X$ be a hyperelliptic curve of genus $g$ and let $r = \operatorname{rank}_\mathbb{Z} \operatorname{Jac}_X(\mathbb{Q})$. Suppose $r \leq g - 3$. Then

$$\#X(\mathbb{Q}) \leq 8(r + 4)(g - 1) + \max\{1, 4r\} \cdot g.$$ 

A main ingredient in Stoll’s proof is to understand the discrepancy between the different flavors of integration. Eric Katz noticed that this discrepancy “factored through the tropicalization of the torus part of the Berkovich uniformization of $X$”. After a thorough reinterpretation of the method of Chabauty and Coleman via Berkovich spaces, and harnessing the full catalogue of tropical and non-Archimedean analytic tools, we were able to improve Stoll’s result to arbitrary curves of small rank.

**Theorem 6.2** (Katz–Rabinoff–Zureick-Brown, [KRZB]). Let $X$ be any curve of genus $g$ and let $r = \operatorname{rank}_\mathbb{Z} \operatorname{Jac}_X(\mathbb{Q})$. Suppose $r \leq g - 3$. Then

$$\#X(\mathbb{Q}) \leq 84g^2 - 98g + 28.$$ 

For more details, see our survey [KRZB16].

**References**


Dept. of Mathematics, Emory University, Atlanta, GA 30322 USA

E-mail address: dzb@mathcs.emory.edu