Numerical Investigations of The Sign Parity Conjecture for Elliptic Curves

Atul Pokharel *

January 20, 2002

Mathematics Department
Princeton University
Princeton, NJ 08544

Abstract

It is conjectured that in a family of elliptic curves, in the limit, the sign of the functional equation of a curve in the family is equally likely to be +1 or -1.In these investigations, we explain the main theorems and conjectures leading up to this conjecture, henceforth called the Sign Parity Conjecture. Following this, we assume that the occurrence of sign are independent events and propose that it can thereby be modeled by a symmetric random walk, and give numerical evidence for two "very special" one parameter families of curves of rank 2: $y^2 = x^3 - t^2x + t^4$ and $y^2 = x^3 - t^2x + t^2$ Ultimately, we propose a method of checking whether or not the occurrence of signs are indeed independent events.

*E-mail: pokharel@math.princeton.edu. Many thanks to Prof. Andrew Wiles, Steven J. Miller, Harald A. Helfgott

1 Introduction

In this section, we review some of the basic concepts and terminology concerning Elliptic Curves in a nutshell before moving on to the L-Function. We follow the standard definitions and notation of

Definition 1.1 (Weierstrass Form) *A cubic over a field* $\mathbf{k}$ *in Weierstrass form is given projectively by*

$$y^2w + a_1xyw + a_3yw^2 = x^2 + a_2x^2w + a_4xw^2 + a_6w^3 \tag{1.1}$$

with coefficients in $\mathbf{k}$

Definition 1.2 (Elliptic Curve) *We define an Elliptic curve over a field* $\mathbf{k}$ *to be nonsingular cubic over* $\mathbf{k}$ *that is in Weierstrass Form.*

The behavior of the curve at (0,1,0) at infinity is well known (see Knapp .57), and so we can study much of the behavior of an elliptic curve by studying its affine form:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{1.2}$$

Making the following standard change of variables,

$$\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= 2a_4 + a_1a_3 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2 \end{aligned} \tag{1.3}$$

and

$$c_4 = b_2^2 - 24b_4c_6 = -b_2^3 + 36b_2b_4 - 216b_6 \tag{1.4}$$

If we assume that the character of the field $\mathbf{k}$, char($\mathbf{k}$)$\neq$2,

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6 \tag{1.5}$$

If we assume that char($\mathbf{k}$)$\neq$2, and char($\mathbf{k}$)$\neq$3,

$$y^2 = x^3 - 27c_4x - 54c_6 \tag{1.6}$$

Definition 1.3 (Discriminant) *For any field* $\mathbf{k}$*, the discriminant* $\triangle$ *of the Elliptic Curve is given by the formula*

$$\triangle = -b_2^2 b_8 - 8b_4^3 - 27b_6^2 + 9b_2 b_4 b_6 \tag{1.7}$$

Definition 1.4 (j-invariant) *For an elliptic curve, we define the j-invariant as*

$$j = c_4^3 / \triangle \tag{1.8}$$

The j-invairant is well defined since $\triangle \neq 0$ if the cubic is non singular.[Knapp, Theorem 3.2]. The j-invariant has weight 0, therefore is invariant under all admissible changes of variables. See knapp[p.63] for a definition of admissible change of variables.

Definition 1.5 (Minimal Weierstrass Equation) *An equation of the form1.2 is minimal for a prime* p *if the power of* p *dividing* $\triangle$*cannot be decreased by making an admissible change of variables over* $\mathbf{Q}$ *with the property that the new coefficients are* p*-integral.*

Equation 1.2 is called a global minimal Weierstrass equation if it is minimal for all primes and if the coefficients are integers.

Theorem 1.6 (Due to Neron) *For every elliptic curve* $\boldsymbol{E}$ *over* $\mathbf{Q}$ *there exists an admissible change of variables over* $\mathbf{Q}$ *such that the resulting equation is a global minimal Weierstrass equation.*

Proof: see Knapp[1] p.292

In light of this theorem, henceforth, when we refer to an Elliptic Curve, $\mathbf{E}$ over $\mathbf{Q}$ we will assume it is given by the global minimal Weierstrass equation

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6 \tag{1.9}$$

with discriminant $\triangle$, without any loss of generality.

Next, we mention a rather important theorem by Mordell

Theorem 1.7 (Mordell's Theorem) *If* $\mathbf{E}$ *is an elliptic curve over* $\mathbf{Q}$*, then the abelian group* $\mathbf{E}(\mathbf{Q})$ *is finitely generated.*

Proof: see Knapp, IV

In the context of Mordell's theorem, we will define three more terms, which will be used freely later on.

Definition 1.8 (Naive height) *for P=(x,y) in* $\mathbf{E}(\mathbf{Q})$ *write* $x = \frac{p}{q}$ *with GCD(p,q)=1. We define the Naive height of the point P by*

$$h_0(P) = logmax(|p|, |q|) \geq 0 \tag{1.10}$$

by convention we define $h_0(\infty)=0$

Definition 1.9 (Canonical Height) *There exists a unique function h:*$\mathbf{E}(\mathbf{Q}) \rightarrow \mathbf{R}$ *satisfying (i) h(P)-*h_0*(P) is bounded (ii) h(2P)=4h(P) The function is given by*

$$h(P) = \lim_{n\to\infty} \frac{h_0(2^n P)}{4^n} \tag{1.11}$$

This has h(P)$\geq$ *0 with equallity iff P has finite order. This function is called the canonical height.*

Using Mordells Theorem, we can now define a geometric rank. From Mordell's Theorem,

$$E(Q) \cong \mathbf{Z}^r \oplus F \tag{1.12}$$

The group F is a finite abelian group, uniquely defined as the torsion subgroup. The integer, r is the geometric rank of $\mathbf{E}(\mathbf{Q})$ The analytic rank is the order of vanishing of the L-function (defined below) at s=1.

2 The L-Function

Let $\mathbf{E}$ be an elliptic curve over $\mathbf{Q}$. For each prime p, we consider the reduction $\mathbf{E_p}$ of $\mathbf{E}$ modulo p. $\mathbf{E_p}$ is defined over $\mathbf{Z}/\mathbf{pZ}$ and singular iff $p|\triangle$.

We now define

$$a_p = p + 1 - N_p \tag{2.13}$$

where

$$N_p = \sharp E_p(Z/pZ) \tag{2.14}$$

$\mathbf{N_p}$ is the number of projective solutions of $\mathbf{E_p}$ in $\mathbf{Z_p}$ (i.e. including the solution at infinity)

Definition 2.1 (Local L-Factor) *We define the local L factor for the prime p by*

$$L_p(u) = \begin{cases} \frac{1}{(1-a_p u + p u^2)} & \text{if } p \text{ does not divide } \triangle \\ \frac{1}{1-a_p u} & \text{if } p|\triangle \end{cases} \tag{2.15}$$

The L-function of E is the product of the local L-factors with u replaced in the p^{th} factor by $p^{-s}$and defined to be

$$L(E,s) = \prod_{p|\triangle}(1 - a_p p^{-s})^{-1} \prod_{p\dagger\triangle}(1 - a_p p^{-s} + p^{1-2s})^{-1} \tag{2.16}$$

We will just make a note here that the L-function comes from a more naturally defined zeta function, or a variant of such a function. Refer to Knapp for details.

Theorem 2.2 (Hasse) *Let* $\mathbf{E}$ *be an elliptic curve over* $\mathbf{Q}$ *with integer coefficients. For each prime* $p \dagger \triangle$*, let* $\mathbf{E_p}$ *be the reduction modulo* p*. Then*

$$|a_p| < 2\sqrt{p} \tag{2.17}$$

Corollary 2.3 *The Euler product defining* $\mathbf{L(s, E)}$ *converges for* $Re(s) > \frac{3}{2}$ *and is given there by an absolutely convergent Dirichlet series.*

Theorem 2.4 (Modularity Theorem) *Every elliptic curve over* Q *is modular.*

Conjectured by Taniyama,Shimura,Weil. I am still in search of a sketch of the proof.

Using this, we can prove the following theorem:

Theorem 2.5 $L(E,s)$ *has a meromorphic continuation to all of* C *and satisfies the following functional equation:*

$$N^{\frac{2-s}{2}}(2\pi)^{s-2}\Gamma(2-s)L(E,2-s) = W(E)N^{\frac{3}{2}}(2\pi)^{-s}\Gamma(s)L(E,s) \tag{2.18}$$

Where W(E) is called the root number of E and W(E)=+1 or -1.

Proof: See Knapp[1] p.386
We will rewrite this as

$$\Lambda(s) = \epsilon_E \Lambda(2-s) \tag{2.19}$$

Where $\epsilon_E = \pm 1$ Henceforth, when we refer to the functional equation, we will mean this one. Since W(E) determines the sign, we will proceed by looking further at how it is determined.

It is aa classical result that the root number can be expressed as a product at p finite and p=∞:

$$\mathbf{W}(\mathbf{E}) = \prod_p W_p(\mathbf{E}) \tag{2.20}$$

Where

1. $W_\infty(E) = -1$
2. $W_p(E) = 1$ if the reduction E mod p is good
3. $W_p(E) = -1$ if it is split multiplicative
4. $W_p(E) = 1$ if it is non-split multiplicative
5. $W_p(E) = \frac{-1}{p}$ if it is potentially split multiplicative and p=2
6. $W_p(E) \cong \frac{c_6(E)}{2_2^{\nu}(c_6(E))}$ mod 4 if it is potentially multiplicative and p=2 [Connell]
7. $W_p(E) = \frac{-1}{p}$ for gcd $(\nu_p(\triangle(E)), 12) = 2$ or 6
   $W_p(E) = \frac{-2}{p}$ for gcd $(\nu_p(\triangle(E)), 12) = 3$ if it is potentially good and $p > 3$
   $W_p(E) = \frac{-3}{p}$ for gcd$(\nu_p(\triangle(E)), 12)$ =4 [Rohlrich]
8. $W_p(E)$ given by the tables 1 and 2 of [9] if the reduction is potentially good and p= 2 or 3 [Halberstadt]

1-5 of the above results are classical.

Here, $\triangle(E)$ is the discriminant of the minimal Weierstrass equation of E and $(c_4(E), c_6(E))$ are the corresponding parameters

Looking at the reductions of $\mathbf{E}$ mod p is the same as considering the curve $\mathbf{E_p}$ over $\mathbf{Z_p}$ with the $\mathbf{Z}$ coefficients taken modulo p.

We can classify reductions of $\mathbf{E}$ modulo p as follows:

1. good if $p \dagger \triangle$
   A reduction that is good then the reduction $\mathbf{E}$ mod p gives an Elliptic curve over $\mathbf{Z}/p$.

2. split multiplicative if $p|\triangle, p \dagger c_4$ and $\frac{-c_6}{p} = 1$
   A reduction that is split multiplicative has a node and the slopes at the node lie in $\mathbf{Z}/p$.

3. non-split multiplicative if $p|\triangle$, $p \dagger c_4$ and $\frac{-c_6}{p} = -1$.
   Such a reduction has a node, but the slopes at the node do not lie in $\mathbf{Z}/p$.

4. potentially multiplicative $\nu_p(c_4) = 2, \nu_p(c_6) = 3$ and $\nu_p(lap) > 6$
   This reduction has a cusp, but becomes multiplicative in some extension of $\mathbf{Z}/p$.

5. potentially good otherwise
   This reduction has a cusp, but becomes good in some extension of $\mathbf{Z}/p$.

The function $\nu_p(a)$ returns the number of times the prime p divides a.

Conjecture 2.1 (Sign-Parity Conjecture) *In a family of elliptic curves, in the limit, the sign of the functional equation of a curve in the family is equally likely to be +1 or -1.*

3 Why Parity is Interesting

First we must recall a conjecture by Birch,Swinnerton and Dyer

Conjecture 3.1 (Birch Swinnerton and Dyer, (BSD)) *The BSD conjecture can also be stated as*

$$\textit{Algebraic Rank of E} = \textit{order of vanishing of L(E,s) at s=1} \tag{3.21}$$

The BSD conjecture also suggests the following

Proposition 3.1 $\mathbf{L(1)}$ *= 0 iff* $\mathbf{E(Q)}$ *is infinite*

Proposition 3.2 *(i)W(E)=1 if the order of vanishing of L(E,s) is even (ii)W(E)=-1 if the order of vanishing of L(E,s) is odd*

Therefore, we have that if BSD is true, then

$$\text{root number} = \text{parity of the algebraic rank} \tag{3.22}$$

Let us restate this

Proposition 3.3 *If the root number of an elliptic curve is -1 then the rank is positive.Morevoer, if the root number is +1, and we know that the algebraic rank is not 0, then the rank is "High" i.e. at least 2.*

4 The Theory Underlying the Experiments

In order to get numerical evidence for the sign parity conjecture for the purposes of this paper, we consider a one parameter family of elliptic curves of the form

$$y^2 = x^3 + A(t)x + B(t) \tag{4.23}$$

For every value of t, we get an elliptic curve in the family. The discriminant in this case is $-4A^3 - 27B^2$, and zero if the curve is singular, non zero otherwise.

Conjecture 4.1 (Independence of Sign) *Given an elliptic curve, the sign of its functional equation is independent of the other curves in the family. More precisely, the occurrence of +1 and -1 within a family are independent events.*

Instead of providing a proof, we give numerical data that suggests this may be true. We address the question of independence further in section 7

Definition 4.1 (Slightly Modified Random Walk) *Consider the classic example of a drunk man, staggering about in unit steps. He is restricted to two possible unique motions (usually one step forward or one step backward). Further, his movement is determined by his flipping an unbiased coin. If it turns up heads he goes in one direction, tails, the other.*

We can then ask how likely is it that he will be some desired distance away from where he started.

Let us make estimates of what is to be expected. We define the probability function as usual.
Let us suppose that for every heads,he moves 2 units forward and every tail does nothing. The probability of the coin being heads $P(h) = p$, of tails $P(t) = t$. p=q=$\frac{1}{2}$. Therefore p+q=1.
Suppose we flip the coin N times, the maximum distance that the man can possibly be from the starting position is then M=2N. In a normal random walk, we usually take heads to be +1 and tails to be -1, however, to make calculations simpler, we add the constant 1 to both steps and make heads +2 and talks 0.
Let us suppose we have H heads, and N-H tails. The total distance the man has travelled, d =2H, is independent of the order in which the heads and tails arrive.

Now since the trials are independent,

$$\text{the probability of getting H heads} = \frac{\binom{N}{H}}{2^N} \tag{4.24}$$

Proposition 4.2 *The mean distance, d, after N flips is therefore*

$$\sum_{n=0}^{N} \frac{2n\binom{N}{n}}{2^N} = N \tag{4.25}$$

Proof:

$$(p+q)^N = \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} \tag{4.26}$$

$$\begin{aligned}
\text{Taking the derivative wrt} p, \text{we have,} & \\
(p+q)^{N-1} &= \sum_{n=0}^{N} \binom{N}{n} n p^{n-1} q^{N-n} \\
N &= \sum_{n=0}^{N} \binom{N}{n} n p^{n-1} q^{N-n} \\
\text{Multiplying by} p, & \\
Np &= \sum_{n=0}^{N} \binom{N}{n} n p^{n} q^{N-n} \\
\text{Since p=q=}\frac{1}{2} & \\
\frac{N}{2} &= \sum_{n=0}^{N} \frac{n\binom{N}{n}}{2^N} \\
N &= \sum_{n=0}^{N} \frac{2n\binom{N}{n}}{2^N} \qquad (4.27)
\end{aligned}$$

Since n is equivalent to the number of heads, it is proved.

Proposition 4.3 *For N flips the standard deviation for the total distance, d, from the starting position, is* $\sqrt{N}$

Proof: Taking the second derivative wrt p of 4.26

$$
\begin{aligned}
N(N-1)(p+q)^{N-2} &= \sum_{n=0}^{N} \binom{N}{n} n(n-1)p^{n-2}q^{N-n} \\
\text{Multiplying by } p^2, & \\
N(N-1)p^2 &= \sum_{n=0}^{N} \binom{N}{n} n(n-1)p^{n} \\
&= \sum_{n=0}^{N} \frac{\binom{N}{n} n(n-1)}{2^N} \\
&= \sum_{n=0}^{N} \frac{\binom{N}{n} n^2}{2^N} - \sum_{n=0}^{N} \frac{\binom{N}{n} n}{2^N} \\
\text{from 4.27} & \\
&= \sum_{n=0}^{N} \frac{\binom{N}{n} n^2}{2^N} - \frac{N}{2} \\
\text{We now have} & \\
\frac{N^2}{4} - \frac{N}{4} + \frac{N}{2} &= \sum_{n=0}^{N} \frac{\binom{N}{n} n^2}{2^N} \\
\frac{N^2}{4} + \frac{N}{4} &= \sum_{n=0}^{N} \frac{\binom{N}{n} n^2}{2^N} \qquad (4.28)
\end{aligned}
$$

Again, since n is equivalent to the number of heads, we can further say the Expected value of the distance squared is given by

$$
\begin{aligned}
E(d^2) &= \sum_{n=0}^{N} (2n)^2 \frac{\binom{N}{n}}{2^N} \\
&= 4\sum_{n=0}^{N} n^2 \frac{\binom{N}{n}}{2^N} \\
\text{from 4.28} & \\
&= N^2 + N. \qquad (4.29)
\end{aligned}
$$

Hence the variance of the total distance satisfies

$$\sigma^2 = E(d^2) - E(d)^2 = N^2 + N - N^2 = N \qquad (4.30)$$

$$(4.31)$$

Therefore, the standard deviation in distance after N flips is

$$\sigma = \sqrt{N} \tag{4.32}$$

For every t, we get a value of the sign of the functional equation. Let S be this ordered set containing the signs of the functional equations of the elliptic curves in the one parameter family 4.23 with $0 < t \leq N$.

Simply plotting the sign changes gives us no more information than that the changes occur. Therefore, we began by grouping the data into blocks of size m.

Definition 4.4 (Block) *Given a block size m, we define the i^{th} block as the ordered set B_i of the $m(i-1)^{th}$ through mi^{th} element of S.*

Definition 4.5 (Block Value) *The Block Value, n_i for the i^{th} block, B_i, of block size m, is defined as*

$$n_i = \sum_{k=1}^{m} \epsilon_k \tag{4.33}$$

Where ϵ_k is the k^{th} element of B_i.

If the signs are independently chosen from +1 and -1,then from 4.3, we expect each $n_i \approx \sqrt{m}$

Given the maximum and minumum block numbers, we place the blocks into bins of a histogram centered at 0. We therefore look at the number of blocks that have a sum within a given range. Since we have assumecd that each occurrence of sign is independent of the others, each block can therefore be assumed to be an independent random walk of m coin flips. Therefore, the histogram can be interpreted as showing the distribution of distances that occurred after N/m random walks. We expect this probability distribution to be essentially gaussian.

Definition 4.6 (Bin) *A bin is defined as an interval (a,b] on the real line.*

Definition 4.7 (Bin Size) *Suppose we are given a set of Block Values, V, and a desired number of bins n, we can find r and s and create a histogram with n bins of size $B = \frac{(r-s)}{n}$ each, with the middle interval centered at 0.*

Here, r and s are such that such that maximum block value in V$\leq r$ and minimum block value in V$\geq s$ Therefore, starting at 0, and moving left, we reach another bin after $\frac{B}{2}$. We call B the Bin Size.

Let us suppose that now we consider the "unit" of the histogram baseline to be $\frac{B}{2}$ where B is the bin size. We can now rewrite the intervals of each bin in terms of this new unit. We will have the central bin be the interval $(-1_B, +1_B]$ (since it is centered at 0) and the n^{th} bin to the right be $(n_B$ to $n_B + 2_B]$, where the subscript B denotes that the numbers have been rescaled.

From this definition, we can determine which bin a given value falls into by rescaling the value as follows

$$a_B = \frac{a}{(\frac{\mathrm{B}}{2})} \tag{4.34}$$

and determining which rescaled interval it falls into. That will be its bin. Here B is again the Bin Size.

Proposition 4.8 *If the standard deviation of m flips is σ_N and the bin size is 2b, the distribution of n block values is a renormalized gaussian distribution*

$$P(x) = 2bn\frac{1}{2\pi\sigma^2}e^{\frac{-x^2}{2\sigma^2}} \tag{4.35}$$

with mean 0 and standard deviation $\sigma = \frac{\sigma_m}{b}$

Proof: If we imagine that all the blocks, of size m, are in a single bin, we see that the total area under our histogram is $2bn$ where $b = \frac{B}{2}$, B is the bin size and n is the total number of blocks we have.

If our histogram represents a random walk, the area under this should be equal to a renormalized Gaussian distribution. Therefore, since the area under a normal Gaussian distribution is 1, we must multiply this by $2bn$ to reacale it to fit our histogram. Also, since the standard deviation must also be rescaled accrding to 4.34, we have that the standard deviation of our histogram should be $\sigma = \frac{\sigma_m}{b}$ where σ_m is the standard deviation for m coin flips.

5 Experimental Results

The two curves that we looked at were

$$y^2 = x^3 - t^2x + t^4 \tag{5.36}$$

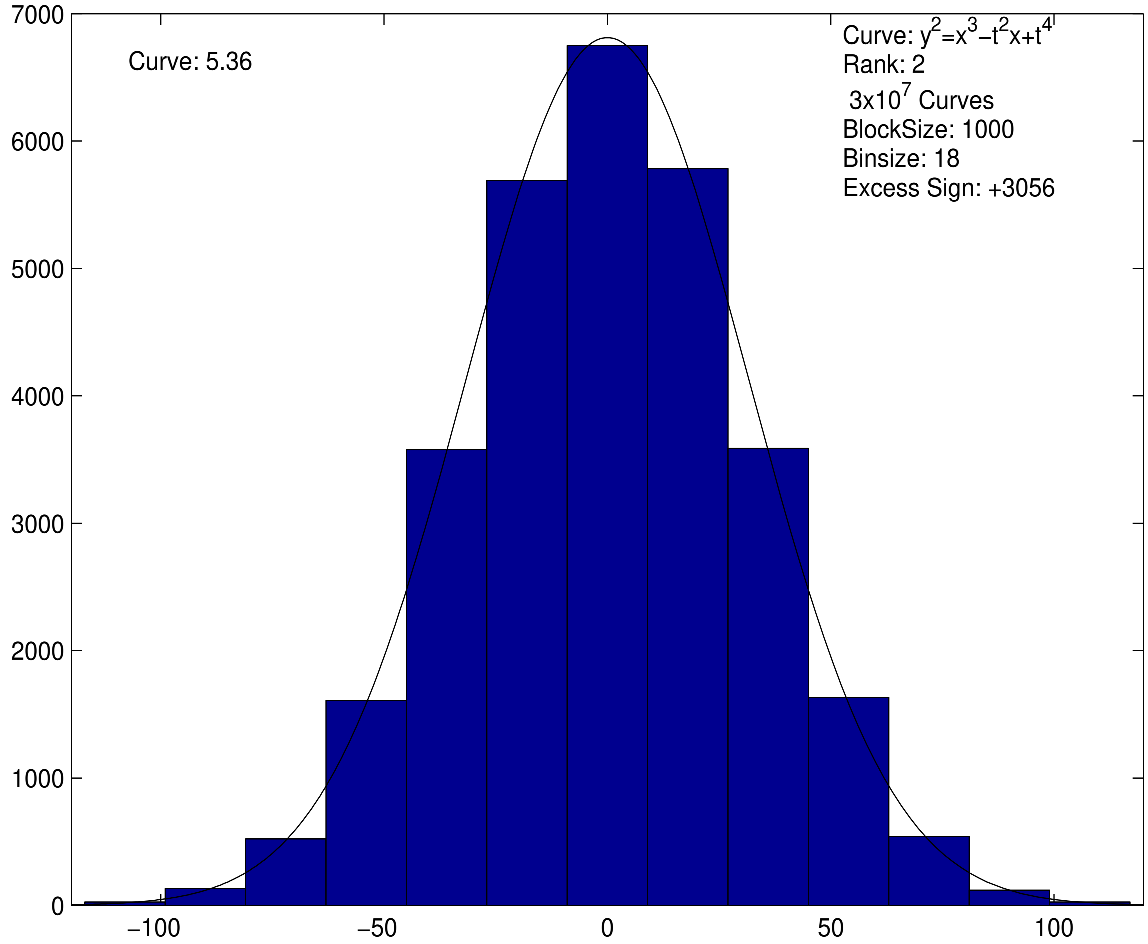
and

$$y^2 = x^3 - t^2x + t^2 \tag{5.37}$$

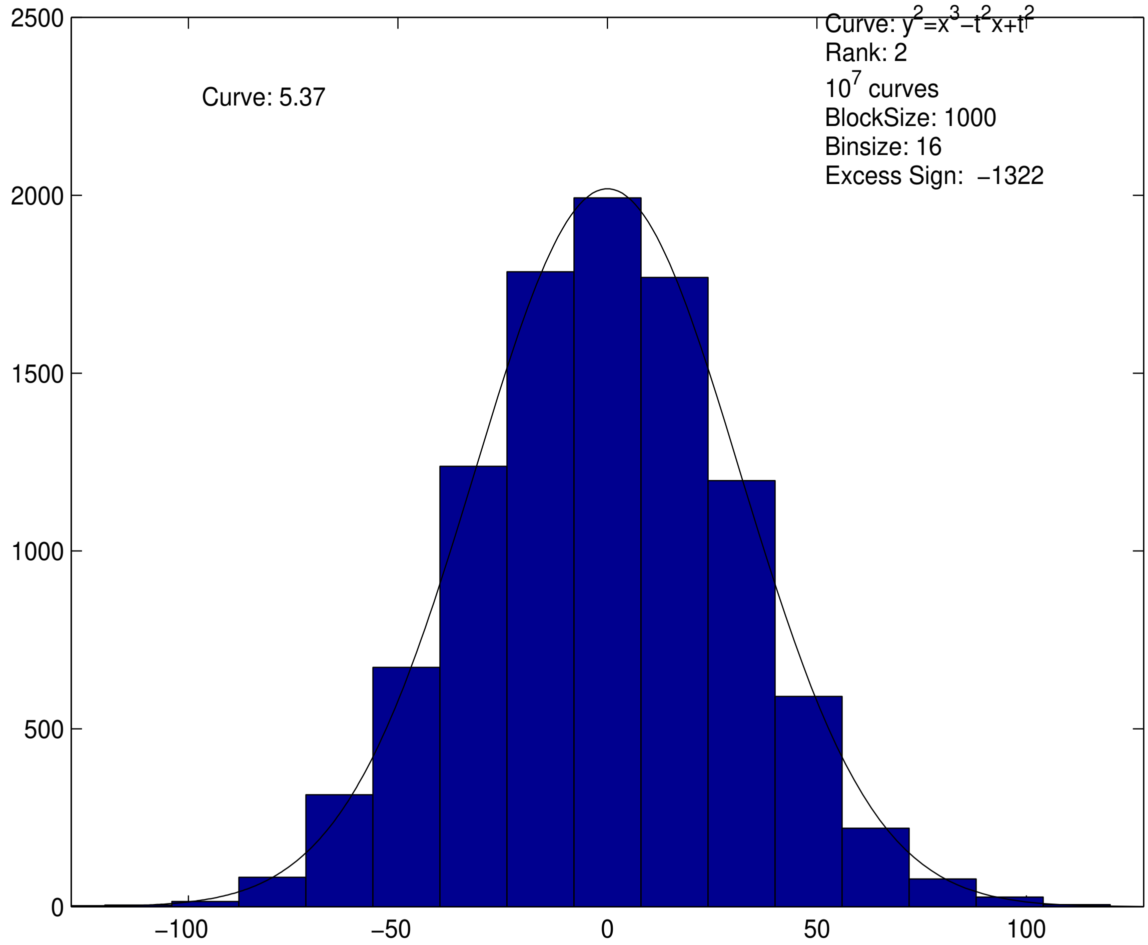
for curve 5.36 we let t go from 0 to $3x10^7$, hence N=$3x10^7$,
and for curve 5.37 we let t go from 0 to $1x10^7$, hence N=$1x10^7$
In order to see something meaningful in the data, we needed then to choose reasonable values for the blocksize, m, and the number of bins, n. Let N be the total number of curves. The only two constraints on these were as follows
-the blocksize,m must be of the order of $\sqrt{N}$, the standard deviation.
-the number of bins,n, must be such that the binsize is much less than $\sqrt{m}$ which is the standard deviation for m trials.

We decided to let m=1000, which is of the order of the standard deviation for both experiments, and the number of bins, n to be 16, which is $<< \sqrt{1000}$. These two parameters were kept constant throughout the experiments.

If the graphs do not appear below, please refer to the following pages for the data.



We also see in the first graph, 5.36 there is an excess positive sign of 3056. The expected deviation is $\sqrt{3x10^7} = 5477$. Therefore it is within expectation.



And in the second graph, there is excess negative sign of -1322 .The expected Deviation is 3162, and therefore, it is again within expectation.

As far as eyballing goes, the graphs suggest that the data fits the expected values very well. Although finer discrepancy meaures are certainly in order. The graphs seem to support our assumptions well.

6 Conclusions

We saw that comparing the occurrence of sign to a symmetric random walk, the expected results fit nicely with what was observed- that the distribution is essentially gaussian. In light of the central limit theorem we may propose that it is not a random walk but a different distribution of finite variance,

although this would also have to be checked by other means.
The main difficulties in programming came from tweaking the code to run fast enough to do a reasonable number of curves in a reasonable amount of time. Most of the difficulty came in understanding the theory enough to set up the data so that we could observe something meaningful instead of many sign changes. Although I would have liked to have done much finer measures of the discrepancy, my background in this was insufficient to do it in the available amount of time and therefore has been left out of this paper.

7 Yet To be Done

An interesting question that was raised while doing these experiments was whether or not the occurrence of sign are actually independent events. Although we assumed in this paper that they were, the next step is to check this claim. Steven and myself will investigate this in the coming weeks. A rough sketch of the method we propose is very simple and as follows:

From now on, let us denote -1 by 0 (this simplifies the explanation and computation considerably)

Once this change is made, the data we have from the curves is a long string of bits. We will look for occurrences of all possible patterns of n bits in this data. From this, we get an observed probability of occurrence for each pattern. We can theoretically predict the probability of occurrence of the pattern, assuming random occurrence. Then we will define a tolerable discrepancy, and look for those patterns that occur more than an expected number of times, within error, if any. We propose that this is a reasonable measure of independence.

For example, suppose we had 50 curves, and we decided on looking for patterns 5 bits wide. More precisely, we want to look at the relative frequencies of 00000, 00001,00010, $\ldots$,11111 within the 50 bits. If the occurrences are independent,then all 32 patterns are equally likely. There are two ways to look for patterns: either we can allow overlapping patterns, or not. For instance, if the pattern we were looking for was 11111, then there is a possibility that if the next bit is 1, we get another occurrence. so 6 bits give us two occurrences. However if we are looking for a pattern 10011 then there is no way that this pattern can recur in less than 11 bits. So we will also have to consider these edge effects. Therefore, depending on which way we search for patterns we can get different expected frequencies.

The difficulties should mainly be computational ones: for instance, I feel the main question will be defining a pattern width that allows us to see recurrences, yet keeps the computation tractable.

8 Code, Optimizations, Suggestions

In this section, I will explain the optimizations made to my code so that I could reach the number of curves that I reached, on the limited resources of the math dept. computers.

The first was the code itself. There are a few things that helped

1. In PARI, the results are stored on an internal stack. Moreover, you have control over the stack pointer. Therefore, I linearized my code so that i didnt need to call the PARI garbage collection routines, but could simply reset the stack pointer to point to before the garbage, once we finished with calculations. This helped once the number of curves became large. Here is example code:

```
long lTop;
//Save the bottom of stack before computations
lTop=avma;
/* this is similar to what we have before, marking
the start of what we'll call the garbage.  */

gaffsg(0,(GEN) F[1]); /* reassigns values of F */
gaffsg(0,(GEN) F[2]);
gaffsg(0,(GEN) F[3]);
gaffsg(-27*(48*t-25),(GEN) F[4]);
gaffsg(-54*(360*t-125),(GEN) F[5]);

E = initell(F,MEDDEFAULTPREC); /* again, initializing E */

/* THE ABOVE IS WHERE WE CHOOSE WHAT CURVE WE'LL BE WORKING
WITH. y^2 = x^3 + A(t)x + B(t) */

sign = ellrootno(E,y);
signSum+=sign;
tsignSum+=sign;

/*restore the stack pointer to what it was before garbage.
This works becuse signSum,tSignsum and sign were allocated
before
we started the calculations.  So they dont lose their values.
```

```
NOTE that E is no longer initialized after this point=>can't
use
it again without calling initell again.
*/
avma=lTop;

/* note here he just sets avma = lTop, and doesn't
use gerepile */
```

There is a good exposition of this in the PARI users guide.
http://modular.fas.harvard.edu/docs/

2. We are given resource quotas (both CPU usage and memory) on math.princeton.edu. However, the amount of memory available to us allows us to make a potential optimzation by writing things to memory first, then dumping it to disk. I tried to cut down on the number of cpu intensive processes as far as possible. This ultimately meant that I stored my results to memory before dumping them in multiple very large chunks to disk. First declare a large(of the order of MB) array and use malloc to get memory for it. Then write all results to this. Then at the end of the calculations, use the write() system call to specify how much of this array to write at once. This is taken from the man page for write

   NAME
   write - write to a file descriptor

   SYNOPSIS
   #include <unistd.h>

   ssize_t write(int fd, const void *buf, size_t count);

   DESCRIPTION
   write writes up to count bytes to the file referenced by
   the file descriptor fd from the buffer starting at buf.
   POSIX requires that a read() which can be proved to occur
   after a write() has returned returns the new data. Note
   that not all file systems are POSIX conforming.

Here, fd would be the file, buf would be your array and count the number of bytes to write. You can invert this advice and use it for read()-ing as well.

3. A very simple optimization is to use int in place of long as far as possible, especially in loops. I found it easy to get mixed up and use long for counters, because I had declared them to be compatible with PARI data types. This just takes up space unnecessarily, and takes time when performing operations.

There was a major hardware optimization that I was lucky to find. Instead of using math.princeton.edu, I discovered two other machines testamd1.princeton.edu and testamd2.princeton.edu that were powerful(1GB RAM, two parallelized Athlon processors at >600Mhz), furthermore, they were relatively unused at the time of writing this. Using these machines is different from math.princeton.edu only in that you must ssh in first to math.princeton.edu, and from there ssh in to testamd1 or testamd2. For instance
from home computer somehow login to math.princeton.edu (use vnc, ssh or something)
then from math.princeton.edu:
ssh testamd1
and login as usual.

After that it is all the same. This really helped the speed of calculations considerably. Unfortunately, these may or may not exist in this state when you try to run your calculations as these are supposedly "experimental" machines.

9 Acknowledgements

I just want to acknowledge those who helped with the paper. Mainly, Professor Wiles, Steve Miller, and Harald Helfgott as the stuff in this paper is mostly their work.

All the code, data and this paper can be found at
`http://www.math.princeton.edu/ pokharel/js`

This paper can be found at
`http://math.princeton.edu/ pokharel/js/paper`
in a plethora of formats.

10 Bibliography

References

[1] Anthony. W Knapp. *Elliptic Curves.* Mathematical Notes, Princeton University Press.

[2] Joseph. H. Silverman John, Tate *Rational Points on Elliptic Curves.* Springer-Verlag Undergraduate Texts in Mathematics.

[3] Yakov Sinai. *Probability Theory An Introductory Course.* Springer-Verlag Translated from the Russian by D. Haughton

[4] Jean Pierre Serre *A course in Arithmetic.* Springer-Verlag,New York 1973

[5] Harald A. Helfgott *Average root numbers in families of elliptic curves and the average of the Moebius function on integers represented by a polynomial.* Work in Progress