Introduction

The present work grew out of an entirely unsuccessful attempt to answer some basic questions about elliptic curves over $\mathbb{Q}$. Start with an elliptic curve E over $\mathbb{Q}$, say given by a Weierstrass equation

$$E: \qquad y^2 = 4x^3 - ax - b,$$

with a, b integers and $a^3 - 27b^2 \neq 0$. By Mordell's theorem [Mor], the group $E(\mathbb{Q})$ of $\mathbb{Q}$–rational points is a finitely generated abelian group. The dimension of the $\mathbb{Q}$–vector space $E(\mathbb{Q})\otimes_{\mathbb{Z}}\mathbb{Q}$ is called the Mordell–Weil rank, or simply the rank, of E. Thus we get a function

$$\{(a,b) \text{ in } \mathbb{Z}^2 \text{ with } a^3 - 27b^2 \neq 0\} \to \{\text{nonnegative integers}\}$$

defined by

$$(a,b) \mapsto \text{the rank of the curve } y^2 = 4x^3 - ax - b.$$

It is remarkable how little we know about this function. For example, we do not know if this function is bounded, or if there exist elliptic curves over $\mathbb{Q}$ of arbitrarily high rank. For a long time, it seems to have been widely believed that this function was bounded. But over the past fifty years, cleverer and cleverer constructions, by Néron [Ner–10], Mestre [Mes–11, Mes–12, Mes–15], Nagao [Nag–20], Nagao–Kouya [Nag–Ko–21], Fermigier [Fer–22], and Martin–McMillen [Mar–McM–23 and Mar–McM–24], have given curves over $\mathbb{Q}$ with higher and higher rank. At this writing in September of 2000 the highest known rank is 24, and the present consensus is that there may well exist elliptic curves over $\mathbb{Q}$ of arbitrarily high rank.

We might then ask if at least we can say anything about the average rank of elliptic curves. What does this question mean? One naive but accessible formulation is this. Since $a^3 - 27b^2 \neq 0$, we might fix a nonzero integer Δ, and look first at the set Ell_Δ defined as

$$\mathrm{Ell}_\Delta := \{(a, b) \text{ in } \mathbb{Z}^2 \text{ with } a^3 - 27b^2 = \Delta\}.$$

Now for each nonzero Δ in $\mathbb{Z}$, the equation

$$X^3 - 27Y^2 = \Delta$$

itself is an elliptic curve over $\mathbb{Q}$. So it has only finitely many solutions (a, b) in integers, by a celebrated result of Siegel giving the finiteness of the number of integral points on an elliptic curve over $\mathbb{Q}$. So the set Ell_Δ is finite. For each integer $N > 0$ we take the union of the sets E_Δ for $0 < |\Delta| \leq N$, and obtain the finite set

$$\mathrm{Ell}_{\leq N} := \{(a,b) \text{ in } \mathbb{Z}^2 \text{ with } 0 < |a^3 - 27b^2| \leq N\}$$

We now form the average

$$\mathrm{avrk}_{\leq N} := (1/\#\mathrm{Ell}_{\leq N})\Sigma_{(a,b) \text{ in } \mathrm{Ell}_{\leq N}} (\text{rank of } y^2 = 4x^3 - ax - b),$$

which is a non–negative real (in fact rational) number.

So now we have a sequence

$$N \to \mathrm{avrk}_{\leq N}$$

of nonnegative real numbers. We do not know if it has a limit. If it does, it would be reasonable to call its limit the average rank of elliptic curves over $\mathbb{Q}$. It is not even known (unconditionally, see [Bru] for conditional results on questions of this type) that the limsup of this sequence is finite.

For a long time, it was widely believed that the large N limit of $\mathrm{avrk}_{\leq N}$ does exist, and that its value is 1/2. Moreover, it was believed that each of the three auxiliary sequences of ratios

fraction of points in $\mathrm{Ell}_{\leq N}$ with rank 0,

fraction of points in $\mathrm{Ell}_{\leq N}$ with rank 1,

and

fraction of points in $\mathrm{Ell}_{\leq N}$ with rank ≥ 2,

has a limit, and that these limits are 1/2, 1/2, and 0 respectively.

Today it is still believed that each of these four sequences has a limit, but there is no longer agreement on what their limits should be. Some numerical experiments ([Brum–McG], [Fer–EE], [Kra–Zag], [Wa–Ta]) support the view that a positive percentage of elliptic curves have rank two or more, i.e., that the fourth limit is nonzero. On the other hand, the philosophy of Katz–Sarnak ([Ka–Sar, RMFEM, Introduction] and [Ka–Sar, Zeroes]) suggests that the limits are as formerly expected, and (hence) that the contradictory evidence is an artifact of too restricted a range of computation.

At this point, we must say something about the L–function L(s, E) of an elliptic curve over $\mathbb{Q}$, and about the Birch and Swinnerton–Dyer conjecture. The curve $E/\mathbb{Q}$ has "conductor" an integer $N = N_E \geq 1$ (whose exact definition need not concern us here) with the property that $E/\mathbb{Q}$ has "good reduction" at precisely the primes p not dividing N. For each such p we define an integer $a_p(E)$ by writing the number of $\mathbb{F}_p$–points on the reduction as $p + 1 - a_p(E)$. The L–function $L(s, E)$ of $E/\mathbb{Q}$ is defined as an Euler product $\prod_p L_p(s,f)$, whose Euler factor $L_p(s, E)$ at each p not dividing N is

$$(1 - a_p(E)p^{-s} + p^{1-2s})^{-1}$$

(and with a recipe for the factors at the bad primes which need not concern us here). The Euler product converges absolutely for $\mathrm{Re}(s) > 2$, thanks to the Hasse estimate

$$|a_p(E)| \leq 2\mathrm{Sqrt}(p).$$

It is now known, thanks to work of Wiles [Wi], Taylor–Wiles [Tay–Wi], and Breuil–Conrad–Diamond–Taylor [Br–Con–Dia–Tay], that every elliptic curve $E/\mathbb{Q}$ is modular. What this means that is that given $E/\mathbb{Q}$, with conductor $N = N_E$, there exists a unique weight two cusp form $f = f_E$ of weight two on the congruence subgroup $\Gamma_0(N)$ of $SL(2, \mathbb{Z})$ which is an eigenfunction of the Hecke operators T_p for primes p not dividing N, whose eigenvalues are the integers $a_p(E)$,

$$T_p f_E = a_p(E) f_E \text{ for every p not dividing N},$$

whose q–expansion at the standard cusp $i\infty$ is q + higher terms, and which is not a modular form on $\Gamma_0(M)$ for any proper divisor M of N.

Now given **any** integer $N \geq 1$ and **any** weight two normalized newform f on $\Gamma_0(N)$, i.e., a cusp form f on $\Gamma_0(N)$ which is an eigenfunction of the Hecke operators T_p for primes p not dividing N, with eigenvalues denoted $a_p(f)$,

$$T_p f = a_p f,$$

whose q–expansion at i∞ is

$$\Sigma_{n\geq 1}\, a_n q^n,\ a_1 = 1,$$

and which is not a modular form on $\Gamma_0(M)$ for any proper divisor M of N, the L–function L(s, f) of f is defined to be the Mellin transform of f. Thus L(s, f) is the Dirichlet series

$$L(s, f) = \Sigma_{n\geq 1}\, a_n n^{-s}.$$

This Dirichlet series has an Euler product $\prod_p L_p(s,f)$ whose Euler factor $L_p(s, f)$ at each p not dividing N is

$$(1 - a_p p^{-s} + p^{1-2s})^{-1}.$$

The Euler product converges absolutely for Re(s) > 2. The function L(s, f) extends to an entire function, and when it is "completed" by a suitable Γ–factor, it satisfies a functional equation under $s \mapsto 2-s$. The precise result is this. One defines

$$\Lambda(s, f) := N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, f).$$

Then Λ(s, f) is entire, and satisfies a functional equation

$$\Lambda(s, f) = \varepsilon(f)\Lambda(2-s, f),$$

where $\varepsilon(f) = \pm 1$ is called the sign in the functional equation.

It turns out that the Euler factors at the bad primes in L(s, E) are equal to those in $L(s, f_E)$, so we have the identity

$$L(s, E) = L(s, f_E).$$

This in turn shows that

$$\Lambda(s, E) := N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, E)$$

extends to an entire function, and satisfies a functional equation

$$\Lambda(s, E) = \varepsilon(E)\Lambda(2-s, E),$$

with $\varepsilon(E)\ (:= \varepsilon(f_E)) = \pm 1$.

The upshot of all this discussion is that L(s, E) is holomorphic at the point s=1, so it makes sense to speak of the order of vanishing of L(s, E) at the point s=1. The basic Birch and Swinnerton–Dyer conjecture for E/ℚ is the assertion that the rank of E/ℚ is the order of vanishing of L(s, E) at s=1. [We say "basic" because there is a refined version which interprets not only the order of vanishing as the rank, but also specifies the leading coefficient in the power series expansion of L(s, E) at s=1.] It is instructive to note that the conjecture was made thirty years before it was known in general that L(s, E) even made sense at s=1.

One calls the order of vanishing of L(s, E) at s=1 the "analytic rank" of E/ℚ, denoted

$\mathrm{rank}_{\mathrm{an}}(E)$:

$$\mathrm{rank}_{\mathrm{an}}(E) := \text{order of vanishing of } L(s, E) \text{ at } s=1.$$

What we now know about the basic Birch and Swinnertion Dyer conjecture can be stated all too briefly:

1) if L(1, E) is nonzero, then E has rank zero.

2) if L(s, E) has a simple zero at s=1, then E has rank one.

In other words, what we know is that

$$\mathrm{rank}_{\mathrm{an}}(E) \le 1 \Rightarrow \mathrm{rank}(E) = \mathrm{rank}_{\mathrm{an}}(E).$$

To emphasize how little we know, it is perhaps worth pointing out that we know neither the a priori inequality

$$\mathrm{rank}(E) \le \mathrm{rank}_{\mathrm{an}}(E),$$

nor the opposite a priori inequality

$$\mathrm{rank}_{\mathrm{an}}(E) \le \mathrm{rank}(E)..$$

[In the "function field case", the analogue of the first a priori inequality holds trivially, cf. [Tate–BSD], [Shio].]

In all the numerical experiments concerning rank of which we are aware, it is the analytic rank rather than the rank which is calculated. Thus the relevance of these experiments to the rank of elliptic curves is conditional on the truth of the Birch and Swinnerton–Dyer conjecture.

A basic observation, due to Shimura (and related by him to Birch at the 1963 Boulder conference in the context of relating twists of modular forms and elliptic curves, cf. [Bir–St]), is that if the sign ε(E) in the functional equation of L(s, E) is −1 [respectively +1], then L(s, E) has a zero of odd [respectively even] order at s=1. So we have the implication

$$\varepsilon(E) = -1 \Rightarrow \mathrm{rank}_{\mathrm{an}}(E) \text{ is } \ge 1, \text{ and odd.}$$

If the Birch and Swinnerton–Dyer conjecture holds, then

$$\varepsilon(E) = -1 \Rightarrow \mathrm{rank}(E) \text{ is } \ge 1, \text{ and odd.}$$

On the other hand, if ε(E) is +1, then rank(E) is forced to be even, so **if** the rank is nonzero, it is at least two. We should point out here that the parity consequence

$$\mathrm{rank}_{\mathrm{an}}(E) \equiv \mathrm{rank}(E) \bmod 2$$

of the Birch and Swinnerton–Dyer conjecture remains a conjecture, sometimes called the Parity Conjecture [Gov–Maz].

The expectation that the average rank of elliptic curves over $\mathbb{Q}$ be 1/2 is based on three ideas: first, that the Birch and Swinnerton–Dyer conjecture holds for all E/$\mathbb{Q}$, second, that half the elliptic curves have sign ε(E) = +1, and half have sign ε(E) = −1, and third, that for most elliptic curves, the rank is the minimum, namely zero or one, imposed by the sign in the functional equation.

The recent conjecture of Katz–Sarnak [Ka–Sar, RMFEM, page 14] about the distribution of the low–lying zeroes of L(s, E) would, if true, make precise and quantify the third idea above, that for most elliptic curves, the rank is the minimum imposed by the sign of the functional equation. We refer to [Ka–Sar, RMFEM, 6.9 and 7.5.5] for the definitions and basic properties of the "eigenvalue location measures" $\nu(+,j)$ and $\nu(-,j)$, j = 1, 2,...on $\mathbb{R}$. What is important for our immediate purposese is that these are all probability measures supported in $\mathbb{R}_{\geq 0}$ which are absolutely continuous with respect to Lebesgue measure.

In order to formulate the conjecture, we must assume the Riemann Hypothesis for the L–functions L(s, E) of all E/$\mathbb{Q}$, namely that all the nontrivial zeroes of L(s, E) (i.e., all the zeroes of Λ(s, E)) lie on Re(s) = 1. If L(s, E) has an even functional equation, its nontrivial zeroes occur in conjugate pairs $1 \pm i\gamma_{E,j}$ with $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq ....$ If E has an odd functional equation, then s=1 is a zero of L(s, E), and the remaining nontrivial zeroes of L(s, E) occur in conjugate pairs $1 \pm i\gamma_{E,j}$ with $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq ....$

We then normalize the heights $\gamma_{E,j}$ of these zeroes according to the conductor N_E of E as follows. We define the normalized height $\tilde{\gamma}_{E,j}$ to be

$$\tilde{\gamma}_{E,j} := \gamma_{E,j}\log(N_E)/2\pi.$$

Now let us return to the set

$$\mathrm{Ell}_{\leq N} := \{(a,b) \text{ in } \mathbb{Z}^2 \text{ with } 0 < |a^3 - 27b^2| \leq N\}.$$

We then break up $\mathrm{Ell}_{\leq N}$ into two subsets

$$\mathrm{Ell}_{\leq N,\pm}$$

according to the sign in the functional equation of the L–function of the E/$\mathbb{Q}$ given by the corresponding Weierstrass equation. It is known to the experts, but nowhere in the literature, that both ratios

$$\#\mathrm{Ell}_{\leq N,\pm}/\#\mathrm{Ell}_{\leq N}$$

tend to 1/2 as $N \to \infty$.

Conjecture (compare [Ka–Sar, RMFEM, page 14]) The normalized heights of low–lying zeroes of L–functions of elliptic curves over $\mathbb{Q}$ are distributed according to the measures $\nu(\pm, j)$, in the following sense. For any integer $j \geq 1$, and for any compactly supported continuous $\mathbb{C}$–valued function h on $\mathbb{R}$, we can calculate the integrals$\int_{\mathbb{R}} hd\nu(\pm,j)$ as follows:

$$\int_{\mathbb{R}} hd\nu(-, j) = \lim_{N \to \infty} (1/\#\mathrm{Ell}_{\leq N,-}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,-}} h(\tilde{\gamma}_{E,j}),$$

and

$$\int_{\mathbb{R}} hd\nu(+, j) = \lim_{N \to \infty} (1/\#\mathrm{Ell}_{\leq N,+}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,+}} h(\tilde{\gamma}_{E,j}).$$

What is the relevance of this conjecture to rank? Take, for each real t > 0, a continuous function $h_t(x)$ on $\mathbb{R}$ which has values in the closed interval [0, 1], is supported in [–t, t], and takes the value 1 at the point x=0, for instance

___/___.

By the absolute continuity of $\nu(\pm, j)$ with respect to Lebesgue measure, we have

$$|\int_{\mathbb{R}} h_t d\nu(\pm, j)| \to 0 \text{ as } t \to 0.$$

Choose N large enough that $\mathrm{Ell}_{\leq N,\varepsilon}$ is nonempty for both choices of sign ε. Denote by $\delta_0(x)$ the characteristic function of {0} in $\mathbb{R}$. Notice that we have the trivial inequality $h_t(x) \geq \delta_0(x)$ for all real x. For the choice +, we have

$$(1/\#\mathrm{Ell}_{\leq N,+}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,+}} h(\tilde{\gamma}_{E,j})$$
$$\geq (1/\#\mathrm{Ell}_{\leq N,+}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,+}} \delta_0(\tilde{\gamma}_{E,j})$$
$$:= \text{fraction of E in } \mathrm{Ell}_{\leq N,+} \text{ with } \mathrm{rank}_{an}(E) \geq j.$$

For the choice –, the L function automatically vanishes once at s=1, but that zero is not on our list $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq \ldots$, so we have

$$(1/\#\mathrm{Ell}_{\leq N,-}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,-}} h(\tilde{\gamma}_{E,j})$$
$$\geq (1/\#\mathrm{Ell}_{\leq N,-}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,-}} \delta_0(\tilde{\gamma}_{E,j})$$
$$:= \text{fraction of E in } \mathrm{Ell}_{\leq N,-} \text{ with } \mathrm{rank}_{an}(E) \geq j+1.$$

Taking the limit as N →∞, and setting j = 1, we find

$$0 = \lim_{N \to \infty} \text{fraction of E in } \mathrm{Ell}_{\leq N,+} \text{ with } \mathrm{rank}_{an}(E) \geq 1,$$

and

$$0 = \lim_{N \to \infty} \text{fraction of E in } \mathrm{Ell}_{\leq N,-} \text{ with } \mathrm{rank}_{an}(E) \geq 2.$$

Therefore, if we assume in addition the Birch and Swinnerton–Dyer conjecture for all E/$\mathbb{Q}$, we find a precise sense in which a vanishingly small fraction of elliptic curves over $\mathbb{Q}$ have rank greater than that imposed by the sign in the functional equation.

As measures on $\mathbb{R}_{\geq 0}$, the $\nu(\pm, j)$ all have densities, and these densitites are the restrictions to $\mathbb{R}_{\geq 0}$ of entire functions, cf. [Ka–Sar, RMFEM, 7.3.6, 7.5.5]. A signifigant difference between the two measures $\nu(-,1)$ and $\nu(+,1)$ is that the density of $\nu(-,1)$ vanishes to second order at the origin x=0, while that of $\nu(+,1)$ is $2 + O(x^2)$ near x=0, cf. [Ka–Sar, RMFEM, AG.0.3 and AG.0.5].

Thus the imposed zero of L(s, E) at s=1 for E of odd functional equation "quadratically repels" the next higher zero $1 + i\gamma_{E,1}$, while for E of even functional equation the point s=1 does not repel the next higher zero $1 + i\gamma_{E,1}$. This is presumably the phenomenon underlying the fact that in the numerical experiments cited above which call into question the "average rank = 1/2" hypothesis, what is found numerically is that about half the curves tested have odd sign, and essentially all of these have analytic rank one, while among the other half of the curves tested, among those with even sign, between twenty and forty percent have analytic rank two or more. What may be happening is that, because $\nu(-,1)$ quadratically repels the origin, while $\nu(+,1)$ does not repel the origin, in any given range of numerical computation, the data on ranks of curves of odd sign will look "better" than the data on ranks of curves of even sign ["better" in supporting the idea that elliptic curves over $\mathbb{Q}$ "try" to have as low a rank as their signs will allow].

An attractive and apparently "easier" question to study is this. Fix one elliptic curve E/$\mathbb{Q}$, with Weierstrass equation

$$E: \quad y^2 = 4x^3 - ax - b$$

and conductor N_E. For each squarefree integer D, one defines the quadratic twist E_D of E by D to be the elliptic curve over $\mathbb{Q}$ of equation

$$E_D: \quad Dy^2 = 4x^3 - ax - b,$$

or equivalently, (multiply the equation by D^3 and change variables to Dx, D^2y)

$$E_D: \quad y^2 = 4x^3 - aD^2x - bD^3.$$

Denote by χ_D the primitive quadratic Dirichlet character attached to the quadratic extension $\mathbb{Q}$(Sqrt(D))/$\mathbb{Q}$. Thus for odd primes p not dividing D, we have

$$\chi_D(p) = 1 \text{ if D is a square in } \mathbb{F}_p, -1 \text{ if not.}$$

For all primes p which are prime to 2×D×N_E, the a_p for E and for E_D are related by

$$a_p(E_D) = \chi_D(p)a_p(E).$$

The conductor of E_D divides (a power of 2)×D^2×N_E. If we take D ≡ 1 mod 4 and relatively prime to N, then the conductor of E_D is D^2N_E. For any D relatively prime to N, E_D has the sign in its functional equation related to that of E by the rule

$$\varepsilon(E_D) = \chi_D(-N_E)\varepsilon(E).$$

Denote by f := f_E the weight two normalized newform attached to E. The normalized newform attached to E_D is $f\otimes\chi_D$, the unique weight two normalized newform of any level dividing a power of $2DN_E$ whose Hecke eigenvalues at primes not dividing $2DN_E$ are given by the rule $a_p(E_D) = \chi_D(p)a_p(E)$ above.

So having fixed E/$\mathbb{Q}$, we can now ask the same questions as above for the family of curves E_D. Thus for real X > 0, we look at the set

$$\mathrm{Sqfr}_{\le X} := \{\text{squarefree integers D with } |D| \le X\}.$$

On this set we have the function

$$D \mapsto \text{rank of } E_D.$$

We can ask whether as X → ∞, the quantities

average of rank(E_D) over $\mathrm{Sqfr}_{\le X}$,
fraction of D in $\mathrm{Sqfr}_{\le X}$, with rank(E_D) = 0,
fraction of D in $\mathrm{Sqfr}_{\le X}$, with rank(E_D) = 1,
fraction of D in $\mathrm{Sqfr}_{\le X}$, with rank(E_D) ≥ 2,

have limits, and, if so, what they are. Or if not, what the limsup's might be. And a more refined version is to break $\mathrm{Sqfr}_{\leq X}$ up according to the sign in the functional equation of L(s, E_D) into two sets $\mathrm{Sqfr}_{\leq X,\pm}$, and repeat the above questions over these sets. There are almost no unconditional results.

If we admit the truth of the Birch and Swinnerton–Dyer conjectures for all the twists E_D, then these are questions about the behavior at s=1 of the L–functions L(s, $f\otimes\chi_D$) as D varies. Let us further assume the Riemann hypothesis for the L–functions L(s, f) attached to all weight two normalized newforms f on all $\Gamma_0(N)$. Then we can formulate the following conjecture.

Conjecture [Ka–Sar, Zeroes, II (b) and pg 21] Fix a weight two normalized newform f on any $\Gamma_0(N)$. Break up the set $\mathrm{Sqfr}_{\leq X}$ according to the sign in the functional equation of L(s, $f\otimes\chi_D$) into two subsets $\mathrm{Sqfr}_{\leq X,\pm}$. [It is known that both the ratios

$$\#\mathrm{Sqfr}_{\leq X,\pm}/\#\mathrm{Sqfr}_{\leq X}$$

tend to 1/2 at $X \to \infty$.]. Then the normalized heights $\tilde{\gamma}_{D,j}$ of the low–lying zeroes of the L–functions L(s, $f\otimes\chi_D$) are distributed according to the measures $\nu(\pm, j)$, in the following sense. For any integer $j \geq 1$, and for any compactly supported continuous $\mathbb{C}$–valued function h on $\mathbb{R}$, we can calculate the integrals $\int_{\mathbb{R}} h d\nu(\pm,j)$ as follows.

$$\int_{\mathbb{R}} h d\nu(-, j) = \lim_{X \to \infty} (1/\#\mathrm{Sqfr}_{\leq X,-}) \Sigma_{D \text{ in } \mathrm{Sqfr}_{\leq X,-}} h(\tilde{\gamma}_{D,j}),$$

and

$$\int_{\mathbb{R}} h d\nu(+, j) = \lim_{X \to \infty} (1/\#\mathrm{Sqfr}_{\leq X,+}) \Sigma_{D \text{ in } \mathrm{Sqfr}_{\leq X,+}} h(\tilde{\gamma}_{D,j}).$$

Exactly as above, the truth of this conjecture for f_E gives us

$$0 = \lim_{X \to \infty} \text{fraction of D in } \mathrm{Sqfr}_{\leq X,+} \text{ with } \mathrm{rank}_{an}(E_D) \geq 1,$$

and

$$0 = \lim_{X \to \infty} \text{fraction of D in } \mathrm{Sqfr}_{\leq X,-} \text{ with } \mathrm{rank}_{an}(E_D) \geq 2.$$

So if we assume in addition the Birch and Swinnerton–Dyer conjecture for all the $E_D/\mathbb{Q}$, we find that as $X \to \infty$, 100 percent of the even twists have rank zero, that 100 percent of the odd twists have rank one, and that the average rank of all the twists is 1/2. That this should be so was first conjectured by Goldfeld [Go].

The numerical experiments so far seem to support this conclusion moderately well for odd twists, but poorly for even twists. Again, the fact that $\nu(-,1)$ quadratically repels the origin, while $\nu(+,1)$ does not repel the origin, may be "why" the numerical data so far is "better" for odd twists then for even twists.

We now turn to the the situation for elliptic curves over function fields over finite fields. Thus let k be a finite field, C/k a proper smooth geometrically connected curve, K := k(C) its function field, and E/K an elliptic curve with non–constant j invariant. Then E/K "spreads out" to an elliptic curve over some dense open set U of C, say $\pi : \mathcal{E} \rightarrow$ U. By the theory of the Neron model, if such a spreading out exists over a given open U, it is unique. Moreover, there is a largest such U, called the open set of good reduction for E/K. [Because E/K has non–constant j invariant, it does not have good reduction everywhere on C.] The finite set of closed points of C at which E/K has bad reduction will be denoted Sing(E/K). By the Neron Ogg Shafarevic criterion, the open set of good reduction can be described as follows. Pick a prime number ℓ invertible in K, pick some spreading out

$$\pi : \mathcal{E} \rightarrow \mathrm{U}$$

of E/K, and form the lisse rank two sheaf $\mathrm{R}^1\pi_*\overline{\mathbb{Q}}_\ell$ on U, which by Hasse [Ha] is pure of weight one. Denoting by j : U $\rightarrow$ C the inclusion, form the "middle extension" (:= direct image) sheaf $\mathcal{F} := \mathrm{j}_*\mathrm{R}^1\pi_*\overline{\mathbb{Q}}_\ell$ on C. This sheaf $\mathcal{F}$ on C is independent of the auxiliary choice of spreading out used to define it, and the open set of good reduction for E/K is precisely the largest open set on which $\mathcal{F}$ is lisse. Thus Sing(E/K) as defined above is equal to Sing($\mathcal{F}$), the set of points of C at which $\mathcal{F}$ is not lisse.

The L–function L(T, E/K) is defined to be the L–function of C with coefficients in $\mathcal{F}$, itself defined as the Euler product

$$\mathrm{L}(\mathrm{T}, \mathcal{F}) := \prod_{\mathrm{x}} (\det(1 - \mathrm{T}^{\deg(\mathrm{x})}\mathrm{Frob}_{\mathrm{x}} \mid \mathcal{F}_{\mathrm{x}})^{-1}$$

over the closed points x of C. At each point x of good reduction, the reduction of E/K at x is an elliptic curve $\mathbb{E}_{\mathrm{x}}$ over the residue field $\mathbb{F}_{\mathrm{x}}$, and

$$\det(1 - \mathrm{TFrob}_{\mathrm{x}} \mid \mathcal{F}_{\mathrm{x}}) = 1 - \mathrm{a}_{\mathrm{x}}\mathrm{T} + (\#\mathbb{F}_{\mathrm{x}})\mathrm{T}^2 \text{ in } \mathbb{Z}[\mathrm{T}],$$

where a_{x} is the integer defined by the equation

$$\mathrm{a}_{\mathrm{x}} := 1 + \#\mathbb{F}_{\mathrm{x}} - \#\mathbb{E}_{\mathrm{x}}(\mathbb{F}_{\mathrm{x}}).$$

Thus the local factors at the points of good reduction are visibly $\mathbb{Z}$–polynomials, independent of the auxiliary choice of ℓ. This is true also of the factors at the points of bad reduction [De–Constants, 9.8].

The cohomological expression for this L–function

$$\mathrm{L}(\mathrm{T}, \mathcal{F}) = \prod_{\mathrm{i}=0,1,2}(\det(1 - \mathrm{TFrob}_{\mathrm{k}} \mid \mathrm{H}^{\mathrm{i}}(\mathrm{C}\otimes_{\mathrm{k}}\overline{\mathrm{k}}, \mathcal{F})))^{(-1)^{\mathrm{i}+1}}$$

simplifies. Because E/K has non–constant j invariant, the middle extension sheaf $\mathcal{F}$ is geometrically irreducible when restricted to any dense open set of $\mathrm{C}\otimes_{\mathrm{k}}\overline{\mathrm{k}}$ on which it is lisse [De–Weil II, 3.5.5]. This in turn implies that the groups H^{i} vanish for i≠1. Thus we end up with the identity

$$\mathrm{L}(\mathrm{T}, \mathrm{E/K}) = \mathrm{L}(\mathrm{T}, \mathcal{F}) = \det(1 - \mathrm{TFrob}_{\mathrm{k}} \mid \mathrm{H}^1(\mathrm{C}\otimes_{\mathrm{k}}\overline{\mathrm{k}}, \mathcal{F})).$$

By Deligne [De–WeII, 3.2.3], $\mathrm{H}^1(\mathrm{C}\otimes_{\mathrm{k}}\overline{\mathrm{k}}, \mathcal{F})$ is pure of weight two. Thus L(T, E/K) = L(T, $\mathcal{F}$) lies in $1 + \mathrm{T}\mathbb{Z}[\mathrm{T}]$ and has all its complex zeros on the circle |T| = 1/q (i.e., $\mathrm{L}(\mathrm{q}^{-\mathrm{s}}, \mathrm{E/K})$ has all its zeros on the line Re(s) = 1).

By the Mordell–Weil theorem, the group E(K) is finitely generated. The (basic) Birch and Swinnerton–Dyer conjecture for E/K asserts that the rank of E(K), denoted rank(E/K), is the order of vanishing of L(T, E/K) at the point T = 1/q, q := #k, or equivalently that rank(E/K) is the multiplicity of 1 as generalized eigenvalue of Frob_k on the Tate–twisted group $H^1(C\otimes_k\bar{k}, \mathcal{F})(1)$. We call this multiplicity the analytic rank of E/K:

$$\mathrm{rank}_{an}(E/K) := \mathrm{ord}_{T=1}\det(1-T\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \mathcal{F})(1)).$$

The group $H^1(C\otimes_k\bar{k}, \mathcal{F})(1)$ has a natural orthogonal autoduality <,> which is preserved by Frob_k, i.e., Frob_k lies in the orthogonal group O := $\mathrm{Aut}(H^1(C\otimes_k\bar{k}, \mathcal{F})(1), <,>)$. Now for any element A of any orthogonal group O, its reversed characteristic polynomial

$$P(T) := \det(1-AT)$$

satisfies the functional equation

$$T^{\deg(P)}P(1/T) = \det(-A)P(T),$$

the sign in which is det(−A).

Applying this to Frob_k, we find the functional equation of the L–function of E/K:

$$T^{\deg(L)}L(1/T, E/K) = \varepsilon(E/K)L(T, E/K),$$

where ε(E/K) is the the sign

$$\varepsilon(E/K) = \det(-\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \mathcal{F})(1)).$$

So just as in the number field case, we have the implications

$$\varepsilon(E/K) = -1 \Rightarrow \mathrm{rank}_{an}(E/K) \text{ is odd, and } \geq 1,$$

$$\varepsilon(E/K) = +1 \Rightarrow \mathrm{rank}_{an}(E/K) \text{ is even.}$$

In the function field case, we also have an a priori inequality

$$\mathrm{rank}(E/K) \leq \mathrm{rank}_{an}(E/K).$$

[But the "parity conjecture", the assertion that we have an a priori congruence

$$\mathrm{rank}(E/K) \equiv \mathrm{rank}_{an}(E/K) \bmod 2,$$

is not known in either the number field or the function field case.]

What about quadratic twists of a given E/K? To define these, we suppose that the field K has odd characteristic. Then E/K is defined by an equation

$$y^2 = x^3 + ax^2 + bx + c$$

where $x^3 + ax^2 + bx + c$ in K[x] is a cubic polynomial with three distinct roots in $\bar{K}$. For any element f in $K^\times$, the quadratic twist E_f/K is defined by the equation

$$fy^2 = x^3 + ax^2 + bx + c.$$

Pick any dense open set U in C over which E/K has good reduction, and over which the function f has neither zero nor pole. Then E_f/K also has good reduction over U, say $\pi_f : \mathcal{E}_f \to U$, and the lisse sheaf $R^1(\pi_f)_*\bar{\mathbb{Q}}_\ell$ on U is obtained from $R^1\pi_*\bar{\mathbb{Q}}_\ell$ by twisting by the lisse rank one Kummer sheaf $\mathcal{L}_{\chi_2(f)}$ on U:

$$R^1(\pi_f)_*\bar{\mathbb{Q}}_\ell = \mathcal{L}_{\chi_2(f)}\otimes R^1\pi_*\bar{\mathbb{Q}}_\ell$$

[Recall that χ_2 is the unique character of order two of $k^\times$, and $\mathcal{L}_{\chi_2(f)}$ is the character of $\pi_1(U)$ whose value on the geometric Frobenius Frob_x attached to a closed point x of U with residue field $\mathbb{F}_x$ is $\chi_2(N_{\mathbb{F}_x/k}(f(x)))$. This twisting formula is the sheaf–theoretic incarnation of the relation

$$a_x(E_f/K) = \chi_2(N_{\mathbb{F}_x/k}(f(x)))a_x(E/K),$$

itself the function field analogue of the number field formula

$$a_p(E_D) = \chi_D(p)a_p(E).]$$

So if we denote by $j : U \to C$, the sheaf $\mathcal{F}_f := j_*R^1(\pi_f)_*\overline{\mathbb{Q}}_\ell$ on C attached to E_f/K is related to the sheaf $\mathcal{F} := j_*R^1\pi_*\overline{\mathbb{Q}}_\ell$ on C attached to E/K by the rule

$$\mathcal{F}_f = j_*(\mathcal{L}_{\chi_2(f)}\otimes j^*\mathcal{F}).$$

And the L–function of E_f/K is thus

$$L(T, E_f/K) = L(T, \mathcal{F}_f) = \det(1 - T\mathrm{Frob}_k \mid H^1(C\otimes_k\overline{k}, \mathcal{F}_f)).$$

Thus when we start with a single elliptic curve E/K, and pick a prime number ℓ invertible in K, we get a geometrically irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C. To the extent that we wish to study the **L–functions** of twists E_f/K (rather than the twists themselves, or their actual ranks) the only input data we need to retain is the sheaf $\mathcal{F}$. Indeed, once we have $\mathcal{F}$, the sheaf $\mathcal{F}_f$ attached to a twist E_f/K is constructed out of $\mathcal{F}$ by the rule

$$\mathcal{F}_f = j_*(\mathcal{L}_{\chi_2(f)}\otimes j^*\mathcal{F}),$$

for $j : U \to C$ the inclusion of any dense open set on which f in invertible and on which $\mathcal{F}$ is lisse.

In the case of twists of an E/$\mathbb{Q}$, we twisted by squarefree integers D, and for growing real $X > 0$ we successively averaged over the finitely many such D with $|D| \le X$. What is the function field analogue?

When the function field K is a rational function field $k(\lambda)$ in one variable λ, every element $f(\lambda)$ of $K^\times$ can be written as $f = g(\lambda)^2h(\lambda)$, with $h(\lambda)$ a polynomial in λ of degree $d \ge 0$ which has all distinct roots in $\overline{k}$ (i.e., h is a square free polynomial). This expression is unique up to $(g, h) \mapsto (\alpha g, \alpha^{-2}h)$ for some α in $k^\times$.

So in this case, we might initially try to look at twists of a given E by **all** squarefree polynomials in λ of higher and higher degree d. We might hope that for a given degree d of twist polynomial h, the L–functions $L(T, E_h/K)$ form some sort of reasonable family of polynomials in T. But the degree of $L(T, E_h/K)$ depends on more than just the degree of the square free h. It is also sensitive to the zeros and poles of h at points of Sing(E/K), the set where E/K has bad reduction. For this reason, it is better to abandon the crutch of polynomials and their degrees, and rather impose in advance the behavior of the twisting function f in $K^\times$ at all the points of Sing(E/K).

Since we are doing quadratic twisting, the local geometric behavior at a point x in C of the twist E_f/K sees $\mathrm{ord}_x(f)$ only through its parity. Let us fix an effective divisor D on C and look only at functions f on C whose divisor of poles is exactly D, and which have d := deg(D) distinct zeros (over $\bar{k}$), none of which lies in $\mathrm{Sing}(E/K)\cap(C-D)$. We denote by

$$\mathrm{Fct}(C, D, d, \mathrm{Sing}(E/K)\cap(C-D)) \subset L(D)$$

this set of functions. Then the interaction between f and Sing(E/K) can be read entirely from the divisor D, in fact, from the parity of $\mathrm{ord}_x(D)$ at each point x in Sing(E/K). In particular, if we want to force local twisting at a given point x in C, in particular at a point in Sing(E/K), we have only to be take an effective D which contains the point x with odd multiplicity. This formulation has the advantage of working equally well over a base curve C of any genus, whereas the polynomial formulation was tied to having $\mathbb{P}^1$ as the base.

The upshot is that if we fix an effective divisor D on C, then as f varies in the space

$$\mathrm{Fct}(C, D, d, \mathrm{Sing}(E/K)\cap(C-D)),$$

all the L–functions $L(T, E_f/K)$ have a common degree. It turns out there is a sheaf–theoretic explanation for this uniformity. For any effective D whose degree d satisfies $d \geq 2g+1$, the space

$$\mathrm{Fct}(C, D, d, \mathrm{Sing}(E/K)\cap(C-D))$$

is, in a natural way, the set of k–points of a smooth, geometrically connected k–scheme

$$X := \mathcal{F}ct(C, D, d, \mathrm{Sing}(E/K)\cap(C-D))$$

of dimension $d + 1 - g$. And there is a lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

on the space X, whose stalk $\mathcal{G}_f$ at a k–valued point

$$f \text{ in } X(k) = \mathrm{Fct}(C, D, d, \mathrm{Sing}(E/K)\cap(C-D))$$

is the cohomology group $H^1(C\otimes_k\bar{k}, \mathcal{F}_f)$, and whose local characteristic polynomial $\det(1 - T\mathrm{Frob}_{k,f} \mid \mathcal{G}_f)$ is given by

$$\det(1 - T\mathrm{Frob}_{k,f} \mid \mathcal{G}_f) = \det(1 - T\mathrm{Frob}_{k,f} \mid H^1(C\otimes_k\bar{k}, \mathcal{F}_f)) = L(T, E_f/K).$$

Moreover, the Tate–twisted sheaf $\mathcal{G}(1)$ is pure of weight zero, and has an orthogonal autoduality, which induces on each individual cohomology group $H^1(C\otimes_k\bar{k}, \mathcal{F}_f)(1)$ the orthogonal autoduality responsible for the functional equation of $L(T, E_f/K)$. And for each finite extension k_n/k of given degree n, the stalks of $\mathcal{G}$ at the k_n–valued points $X(k_n)$ encode the L functions of twists defined over k_n.

In this way, questions about the (distribution of the zeroes of the) L–functions $L(T, E_f/K)$, as f varies in the space

$$X(k) = \mathrm{Fct}(C, D, d, \mathrm{Sing}(E/K)\cap(C-D)),$$

become questions about the sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

on X. Thanks to Deligne's equidistribution theorem [Ka–Sar, RMFEM, 9.2.6], we can answer many of these questions in terms the geometric monodromy group G_{geom} attached to the sheaf $\mathcal{G}$.

For example, **if** the group G_{geom} is the full orthogonal group, we automatically get the following results on average analytic rank.

1) The average analytic rank over k_n of twists defined by f's in $X(k_n)$ tends to 1/2 as $n \to \infty$. [And hence the average rank has a limsup $\le 1/2$ as $n \to \infty$.]

2) for each choice of $\varepsilon = \pm 1$, the fraction $\#X(k_n)_{sign\ \varepsilon}/\#X(k_n)$ of twists with sign ε in the functional equation tends to 1/2 as $n \to \infty$.

3) In the set $\#X(k_n)_{sign\ +}$, the fraction of twists with $rank_{an} = 0$ tends to 1 as $n \to \infty$. [And hence in the set $\#X(k_n)_{sign\ +}$, the fraction of twists with rank = 0 tends to 1 as $n \to \infty$.]

4) In the set $\#X(k_n)_{sign\ -}$, the fraction of twists with $rank_{an} = 1$ tends to 1 as $n \to \infty$. [And hence in the set $\#X(k_n)_{sign\ -}$, the fraction of twists with rank ≤ 1 tends to 1 as $n \to \infty$]

Suppose we take a sequence of effective divisors D_ν on C whose degrees d_ν are strictly increasing. Then we get a sequence of smooth k−schemes

$$X_\nu := \mathit{Fct}(C, D_\nu, d, \mathrm{Sing}(E/K) \cap (C - D_\nu))$$

and, on each X_ν, a lisse sheaf $\mathcal{G}_\nu$, say of rank N_ν. The ranks N_ν tend to ∞ with ν. Suppose that for every large enough ν, the group G_{geom} for the sheaf $\mathcal{G}_\nu$ on X_ν is the full orthogonal group $O(N_\nu)$. Then for each choice of sign $\varepsilon = \pm 1$, and each choice of integer $j \ge 1$, we can obtain the eigenvalue location measure $\nu(\varepsilon, j)$ as the following (weak *)double limit: the large ν limit of the large n limit of the distribution of the j'th normalized zero of the L−functions attached to variable points in $X_\nu(k_n)_{sign\ \varepsilon}$.

It was with these applications in mind that we set out to prove that, at least in characteristic $p \ge 5$, as soon as the effective divisor D on C has degree d sufficiently large, then G_{geom} for $\mathcal{G}$ is the full orthogonal group. Unfortunately, this assertion is not always true. What is true is that G_{geom} is either the full orthogonal group O or the special orthogonal group SO, provided only that E/K has nonconstant j invariant, and that

$$d \ge 4g+4, \text{ and}$$
$$2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(E/K)(\overline{k}), 144).$$

[If p=3, this result remains valid provided that the sheaf $\mathcal{F}$ attached to E/K is everywhere tamely ramified, a condition which is automatic in higher characteristic]

We prove that G_{geom} is O if E/K has multiplicative reduction (i.e., unipotent local monodromy) at some point of Sing(E/K) which is not contained in D.

But there are cases where G_{geom} is SO rather than O. If E/K does **not** have unipotent local monodromy at **any** point of Sing(E/K), and if every point of Sing(E/K) which occurs in D does so with even multiplicity, then $\mathcal{G}$ has even rank, say N, and an analysis of local constants, using [De−Constants, 9.5] shows that G_{geom} lies in SO(N) (and hence is equal to SO(N), for d large). cf. Theorem 8.5.7.

An example of an E/K with nonconstant j but with no places of multiplicative reduction, is the twisted (by $\lambda(\lambda-1)$) Legendre curve

$$y^2 = \lambda(\lambda-1)x(x-1)(x-\lambda)$$

over $k(\lambda)$, $k := \mathbb{F}_p$, p any odd prime, which has bad reduction precisely at 0, 1, ∞, but at each of these points the monodromy is

$$(\text{quadratic character})\otimes(\text{unipotent}).]$$

In this example, it turns out (cf. Corollary 8.6.7) that if the characteristic p is 1 mod 4, then all the L–functions over all k_n have **even** functional equation. But, if p is 3 mod 4, then the L–functions over even [respectively odd] degree extensions k_n have even [respectively odd] functional equations!

The Legendre curve itself,

$$y^2 = x(x-1)(x-\lambda)$$

over $k(\lambda)$, has unipotent local monodromy at both 0 and 1. And so if we twist by polynomials $f(\lambda)$ in $k[\lambda]$ of any fixed degree $d \geq 146$, which have all distinct roots in $\overline{k}$ and are invertible at both 0 and 1, the resulting sheaf $\mathcal{G}_d$ on $X_d := \mathit{Fct}(\mathbb{P}^1, d\infty, d, \{0,1\})$ has $G_{geom} = O(N_d)$, with N_d equal to 2d if d is even, and to 2d–1 if d is odd.

Now the Legendre curve makes sense over $\mathbb{Z}[1/2][\lambda, 1/\lambda(\lambda-1)]$, and the space X_d makes sense over $\mathbb{Z}[1/2]$. For each fixed $d \geq 146$, it makes sense to vary the characteristic p, and ask average rank questions about twists of the Legendre curve over $\mathbb{F}_p(\lambda)$ by points in $X_d(\mathbb{F}_p)$ as $p \to \infty$. We get the same answers as we got by fixing p and looking at twists by points in $X_d(\mathbb{F}_pn)$ as $n \to \infty$. If we vary d as well, we can recover the eigenvalue location measures $\nu(\varepsilon,j)$ as well. For each choice of sign ε and integer $j \geq 1$, we can obtain the eigenvalue location measure $\nu(\varepsilon, j)$ as the following (weak *) double limit: the large d limit of the large p limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_d(\mathbb{F}_p)_{\text{sign } \varepsilon}$.

But there are some basic things we don't know, "even" about this Legendre example, and "even" in equal characteristic p. For example, it is easy to see that for any fixed p, $\#X_d(\mathbb{F}_p) \to \infty$ as $d \to \infty$. [Indeed, an element of $X_d(\mathbb{F}_p)$ is a degree d polynomial $f(\lambda)$ in $\mathbb{F}_p[\lambda]$ with all distinct roots in $\overline{\mathbb{F}}_p$, which is nonzero at the points 0 and 1. For $d \geq 3$, any **irreducible** polynomial of degree d in $\mathbb{F}_p[\lambda]$ will lie in $X_d(\mathbb{F}_p)$. And the number of degree d irreducibles in $\mathbb{F}_p[\lambda]$ is at least

$$(p-1)(1/d)(p^d - (d/2)p^{d/2}).]$$

It is also easy to see that for each choice of sign ε, the ratio

$$\#X_d(\mathbb{F}_p)_{\text{sign } \varepsilon}/\#X_d(\mathbb{F}_p)$$

tend to 1/2 as $d \to \infty$. [For d even, use [De–Const, 9.5] as in 8.5.7. For d odd, use the fact that for α in $\mathbb{F}_p^\times$ a nonsquare, and any f in $X_d(\mathbb{F}_p)$, the twists of the Legendre curve by f and by αf have opposite signs in their functional equations, cf. 5.5.2, case 3).] But for p fixed, we do **not know** any of the following 1) through 4).

1)The average rank of twists defined by f's in $X_d(\mathbb{F}_p)$ tends to 1/2 as $d \to \infty$.
2) In the set $X_d(\mathbb{F}_p)_{\mathrm{sign}\,-}$ the fraction of twists with $\mathrm{rank}_{\mathrm{an}} = 1$ tends to 1 as $d \to \infty$.
3) In the set $X_d(\mathbb{F}_p)_{\mathrm{sign}\,+}$ the fraction of twists with $\mathrm{rank}_{\mathrm{an}} = 0$ tends to 1 as $d \to \infty$.
4) For each choice of sign ε and integer $j \geq 1$, the eigenvalue location measure $\nu(\varepsilon, j)$ is the following (weak *) **single** limit: the large d limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_d(\mathbb{F}_p)_{\mathrm{sign}\,\varepsilon}$.

Let us now stand back and see what ingredients were required in the above discussion of quadratic twists of E/K, an elliptic curve over a function field with a nonconstant j–invariant. The function field K is the function field of a projective, smooth, geometrically connected curve C/k, k a finite field. Over some dense open set U in C, E/K spreads out to an elliptic curve $\pi : \mathcal{E} \to U$. We fix a prime number ℓ invertible in k, and form the lisse sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ on U. It is lisse of rank two, pure of weight one, and symplectically self dual toward $\overline{\mathbb{Q}}_\ell(-1)$. The assumption that the j invariant is nonconstant is used only to insure that $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is geometrically irreducible on U. If k has characteristic $p \geq 5$, then $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is everywhere tamely ramified: this is the only way the hypothesis $p \geq 5$ is used. Denoting by $j : U \to C$ the inclusion, we form the sheaf

$$\mathcal{F} := j_*R^1\pi_*\overline{\mathbb{Q}}_\ell$$

on C. We then fix an effective divisor D on C of large degree. We form the quadratic twists E_f/K of E/K by variable f in L(D) which have deg(D) distinct zeroes (over $\overline{k}$), none of which lies in D or in $\mathrm{Sing}(\mathcal{F})\cap(C-D)$. The L–functions of these quadratic twists are the local L–functions of a lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

at the k–points of a smooth, geometrically connected k–scheme

$$X := \mathit{Fct}(C, D, d, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

of dimension $d + 1 - g$.

The original ellliptic curve E/K occcurs **only** through the geometrically irreducible middle extension sheaf $\mathcal{F}$ on C. Once we have $\mathcal{F}$, we can forget where it came from! Our fundamental result in the elliptic case is the determination of the geometric and arithmetic monodromy groups attached to the lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

on the smooth, geometrically connected k–scheme

$$X := \mathit{Fct}(C, D, d, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

of dimension $\deg(D) + 1 - g$.

In fact, we can study the L–functions of twists, by nontrivial tame characters χ of **any** order, of an **arbitrary** geometrically irreducible middle extension sheaf $\mathcal{F}$ on C. Again in this general set up, the L−functions of such twists are the local L−functions of a lisse $\overline{\mathbb{Q}}_\ell$−sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$$

at the k−points of the same smooth, geometrically connected k−scheme

$$X := \mathit{Fct}(C, D, d, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

of dimension deg(D) + 1 − g that occurred above for quadratic twists of elliptic curves. Again the question is to determine the arithmetic and geometric monodromy groups attached to $\mathcal{G}$.

The rank N of $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ grows with deg(D), indeed we have an a priori inequality

$$N := \mathrm{rank}\mathcal{G} \geq (2g - 2 + \deg(D))\mathrm{rank}(\mathcal{F}).$$

One case of our main technical result (Theorems 5.5.1 and 5.6.1) is this. Suppose that $\mathcal{F}$ is everywhere tamely ramified. Then for any effective divisor D of large degree, the geometric monodromy group G_{geom} for $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ is one of the following subgroups of GL(N):

O(N)
SO(N): possible only if N is even
Sp(N): possible only if N is even
a group containing SL(N).

We can be more precise about which cases arise for which input data $(\mathcal{F}, \chi)$. Unless χ has order two and $\mathcal{F}$ is self−dual on $C\otimes\overline{k}$, G_{geom} contains SL(N). If $\mathcal{F}$ is orthogonally self dual on $C\otimes\overline{k}$, and χ has order two, then $\mathcal{G}$ is symplectically self dual on $X\otimes\overline{k}$, and G_{geom} for $\mathcal{G}$ is Sp(N). If $\mathcal{F}$ is symplectically self dual on $C\otimes\overline{k}$, and χ has order two, then $\mathcal{G}$ us orthogonally self dual on $X\otimes\overline{k}$, and G_{geom} for $\mathcal{G}$ is either SO(N), possible only if N is even, or it is O(N).

We can drop the hypothesis that $\mathcal{F}$ be everywhere tame if we are in large characteristic (the exact condition is $p \geq \mathrm{rank}(\mathcal{F}) + 2$), and if we require in addition that the effective divisor D of large degree contain no point where $\mathcal{F}$ is wildly ramified. [This second condition is automatic for D's which are disjoint from the ramification of $\mathcal{F}$.]

Fix, then, input data $(\mathcal{F}, \chi, D)$ as above. As deg(D) grows, the sheaves $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ have larger and larger classical groups as their geometric monodromy groups. The general large N limit results of Katz−Sarnak [Ka−Sar, RMFEM] then give information about the statistical behaviour of the zeroes of the L−functions of the corresponding twists. This information always concerns a double limit $\lim_{\deg(D)\to\infty}\lim_{\deg(E/k)\to\infty}$. For each D we must consider, for larger and larger finite extensions E of k, the L−functions of all twists $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ as f runs over the E−valued points X(E) of the parameter space

$$X = \mathit{Fct}(C, D, d, \mathrm{Sing}(\mathcal{F})\cap(C-D)).$$

We also work out some refinements of these results, where we change the inner limit. The first refinement is twist only by "primes" in X(E), i.e., by functions f in X(E) whose divisor of zeroes $\mathrm{div}_0(f)$ is a single closed point of $C\otimes_k E$. The terminology "prime" arises as follows. In the case when C is $\mathbb{P}^1$ and D is d∞, an element f in X(E) is a polynomial f(t) in E[t] of degree d which has d distinct roots in $\overline{E}$ and which is invertible at the finite singularities of $\mathcal{F}$. Such an element f is "prime" if and only if f(t) is an irreducible polynomial in E[t]. More generally, we might twist only by f's in X(E) whose divisor of zeroes has any pre–imposed factorization pattern. For instance, we might twist only by f's in X(E) which "split completely" over E, i.e., by f's in X(E) which have d distinct zeroes in C(E).

A second refinement is to start not over a finite field, but over a ring of finite type over $\mathbb{Z}$, for instance over $\mathbb{Z}[1/N\ell]$. Then just as in the case of the Legendre family discussed above, we can look at twists by points in $X(\mathbb{F}_p)$ as p → ∞. We get the same answers as we got by fixing p and looking at twists by points in $X(\mathbb{F}_{p^n})$ as n → ∞. We can combine the two refinements. We can twist only by primes in $X(\mathbb{F}_p)$ as p → ∞, or we can twist only by elements of $X(\mathbb{F}_p)$ which "split completely" over $\mathbb{F}_p$. Under mild hypotheses, the limit results remain the same.

Still working over $\mathbb{Z}[1/N]$, take a sequence of divisors D_ν whose degrees d_ν are strictly increasing. We get thus a sequence of parameter spaces

$$X_\nu := \mathit{Fct}(C, D, d, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

over $\mathbb{Z}[1/N]$. We can recover the eigenvalue location measure (whichever of $\nu(\varepsilon,j)$ or $\nu(j)$ is appropriate to the situation being considered) as the following (weak *) double limit: the large ν limit of the large p limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_\nu(\mathbb{F}_p)$.

If we fix the prime p, and let $\nu \to \infty$, then just as in the Legendre case discussed above, it is natural to ask if we can recover the eigenvalue location measure, whichever of $\nu(\varepsilon,j)$ or $\nu(j)$ is appropriate, as the following (weak *) single limit: the large ν limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_\nu(\mathbb{F}_p)$.

Let us now backtrack, and describe the logical organization of this book. It falls naturally into four parts:

Part I (Chapters 1,2,3,4): background material, used in Part II.
Part II (Chapter 5) twisting, done over an algebraically closed field
Part III (Chapters 6,7,8): twisting, done over a finite field
Part IV (Chapters 9, 10): twisting, done over schemes of finite type over $\mathbb{Z}$.

The first chapter is devoted to results from representation theory. Its main result is Theorem 1.5.1, which depends essentially upon a beautiful result of Zarhin about recognizing when an irreducible Lie subalgebra of End(V), V a finite–dimensional $\mathbb{C}$–vector space, is either Lie(SL(V)) or Lie(SO(V)) or, if dim(V) is even, Lie(Sp(V)). It also requires a remarkable recent result [Wales] of Wales concerning finite primitive irreducible subgroups G of GL(V) containing elements γ of type

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1),$$

with ζ a primitive n'th root of unity, $n \geq 3$, which occurs with multiplicity r, $1 \leq r < \dim(V)$. Wales result is that $\dim(V) \leq 4r$.

Wales's inequality was conjectured in an earlier version of this manuscript, written at a time when less was known. It was known, by Blichfeld's 60° theorem [Blich–FCG, paragraph 70, Theorem 8, page 96], that the case $n \geq 6$ could not arise: no finite irreducible primitive subgroup of GL(V) contains such an element. [Blichfeldt's 60^{o} theorem is that in a finite irreducible primitive subgroup G of GL(N, $\mathbb{C}$), if an element g in G has an eigenvalue α such that every other eigenvalue of γ is within 60^{o} of α (on either side, including the endpoints), then g is a scalar.] A little–known result of Zalesskii showed that in the case n=5 we have dim(V) = 2r. For n = 3 or n=4, there were only results for $r \leq 2$. For r=1, we have Mitchell's theorem [Mit], that a finite irreducible primitive subgroup of GL(V) containing a pseudoreflection of order $n > 2$ exists only if $\dim(V) \leq 4$. For r=2, Huffman–Wales prove that if n=4 then $\dim(V) \leq 4$, and if n=3 then dim(V) ≤ 8, cf. [Huf–Wa, Theorems 2 and 3 respectively].

In an appendix to Chapter 1, we explained at length the result [AZ.1] of Zalesskii, and made some conjectures about what might be true in general. Wales then proved the most optimistic AZ.6.2 of these conjectures. Because his manuscript itself refers to some of the results in the appendix, we have left the appendix unchanged, except to add a note saying that Wales has now proven AZ.6.2. We have, however, simplified the original statement and proof of Theorem 1.5.1 by making use of Wales inequality.

In the second chapter, we use the general theory of Lefschetz pencils over an algebraically closed field to develop some basic facts about the geometry of curves, which were surely well known in the nineteenth century.

The third chapter is concerned with induction of group representations, and with giving algebro–geometric criteria for induced representations to have various properties (e.g., to be autodual, to be irreducible).

The fourth chapter is a brief review of "middle convolution" and its effect on local monodromy as developed in [Ka–RLS]. This material depends in an essential way on Laumon's work on Fourier Transform.

After all these preliminaries, we turn to our subject proper in Chapter 5, which is the technical core of the book. We work over an algebraically closed field, and compute monodromy groups of twist sheaves, using as essential ingredients results of all the previous chapters. [In the earlier version of this manuscript, written before Wales result, we could not twist by characters χ of order 4 or 6 unless the input sheaf $\mathcal{F}$ had rank at most 2.]

In Chapter 6, we explain how to formulate over a general base scheme the set up we considered in Chapter 5.

In Chapter 7, we work over a finite field, and extract the diophantine consequences of the monodromy results of Chapter 5. The essential ingredient here is the work of Deligne in [De–Weil II], both his purity theorem and his equidistribution theorem.

In Chapter 8, we give applications to average analytic rank of twists of a given elliptic curve. This leads us into a long discussion of whether the monodromy group in question is O or SO, and leads us to some very nice examples.

In Chapter 9, we begin to work systematically over a base which is a scheme of finite type over $\mathbb{Z}$, rather than "just" a finite field. We also introduce the notion of twisting by a "prime". We prove an equidistribution theorem for primes in divisor classes, which was presumably well known in the late 1920's and 1930's to people like Artin, Hasse and Schmidt, but for which we do not know a reference. We then analyze when twisting only by primes changes nothing as far as equidistribution properties. This leads us to a simple but useful case of Goursat's Lemma.

In Chapter 10, we give "horizontal" versions (i.e., over $\mathbb{F}_p$ as $p \to \infty$) of all the results we found earlier over a finite field k (where we worked over larger and larger extension fields of the given k)

It is a pleasure to thank Cheewhye Chin for his help in preparing the index. I respectfully dedicate this book to the memory of my teacher Bernard Dwork, to whom I owe so very much.