It is a Story

Barry Mazur

For Persi Diaconis on the occasion of his 60th birthday

[*Note: These pages were my rough private notes that I made in preparation for the talk I gave. Needless to say, I didn't cover all this material in my talk. I am willing, but hesitant, to put it on the web because of three reasons, the primary one being that I haven't yet put in proper references for any of the results cited. The second reason is that in its next draft I will be separating much of the narrative discussion from the mathematical discussion, and breaking this material into two articles, each of which will be a more expanded, and perhaps more useful, version of what is offered here. The narrative discussion will be related to a talk I hope to give at a conference on "Mathematics and Narrative," in July of this year.*]

I don't know whether or not Archimedes discovered his hydrostatic principle, the *principle of bouyancy,* while taking his bath. Nor do I know whether he immediately jumped from the tub shouting "Eureka," and ran home stark naked, dripping wet. But I do know that we all have our favorite stories that go along with accounts of mathematics; tales that help to explain, to dramatize, to teach, and even to shape in important ways, the mathematical material being recounted.

Persi, I also know the motivating role that stories play in your descriptions of mathematical themes; so, for your birthday, I thought it might be fun to think a bit about *story* per se, and then tell a specific mathematical story that grips me, at present.

Some years ago I read *Death in the Woods*, a short story of Sherwood Anderson*. It is a disarmingly simple account of a woman's death. In the woods. This piece of writing is like a pillow-feather, or a snowflake: it glints well in the sunlight, but I, at least, couldn't quite grasp it, the slightest breeze propelled it from my attention. A willow-the-wisp of a text. On the fourth reading I came across the declaration "It is a story" in the midst of the narrative. My eyes must have raced by those four words three other times, but I was happy to have finally registered them: I now had the author's collusion in being concerned with the question of whether or not it is, in fact, a story. "Yes, it is." he tells me.

What is a story? Is the previous paragraph a story?

Even before giving an answer to this in the context of stories in expository mathematics, we should be clear about whether the stories we will be considering are *ends* or *means.* In fiction, telling the story is the ultimate goal, and everything else is a means toward that goal. I suspect that even Sheherezade, despite her dangerous situation, and the immediate mortal purpose for her storytelling, would agree to this. It is also so in the Sherwood Anderson story, where the story, one discovers, is about the author's relationship to the story.

In mathematical exposition any of its story elements are usually intended to serve the mathematical ideas: story is a means, the ideas are the end. In conformance, however, to the universal truth that *any general statement has at least one counter-example*, I offer Lewis Carroll's *Alice in Wonderland* as a pin to pierce the bubble of the previous sentence: here is a tale that is not often considered to be mathematical exposition– it is most definitely a *story*–and yet its narrative suffuses into mathematical idea and mathematical idea permeates its narrative–all this happens so thoroughly that when we read it we realize that the cliché words *ends, means* can't possibly be relevant to it **.

Let us throw together a provisional taxonomy of "kinds of storytelling" in mathematics. I feel that there are three standard forms, and also a fourth form: the one that I will be most interested in. My names for the

* in preparation for attending a class given by the novelist Robert Boswell.

** Also, my brother has recently published a collection of stories– Joseph Mazur, *Euclid in the Rainforest*, Pi Press New York,– where tale and mathematics interweave, neither element being subservient to the other; see, for example, *Anna's Accusation* in loc. cit.

standard ones are *origin*-stories (stories explaining some original motivation for studying the mathematics being described, this motivation being external to the development of mathematical ideas themselves), *purpose*-stories (again, other than those which describe a purpose within the context of mathematics itself), and *raisins in the pudding.*

Raisins in the pudding are ornamental bits of story meant to provide anecdotal digressions or perhaps a certain amount of relief from the ardors of the main task of the exposition. At the least they are intended to add extra color. But the primary relationship of the stories or story-fragments in this category to the mathematical subject is ornament: they are not required to help in furthering–in any direct way–the reader's comprehension of the material, nor do they fit in as a part of the structure of the argument presented.

Archimedes' "Eureka" is such a raisin–at least when it isn't coupled to the more purpose-filled story of checking the gold content of King Hiero's crown–but other examples are plentiful, and any book in mathematics that has none of these ornaments risks being too aridly single-purposed to make for pleasant reading. The danger, though, is when the weight of ornament deflects the text from its primary purpose, as happens, in my opinion, in two out of the three recent popular books on the Riemann Hypothesis, where the raisins overwhelm the proof that is supposed to be in the pudding**.

The *Purpose*-stories told in mathematical texts are meant to answer the question: to what end–external to the development of mathematics itself– will this particular piece of mathematics be used? A practical end might be envisioned, and, indeed, the great edifice of Applied Mathematics is devoted to mathematics strongly shaped by some particular issue or issues (e.g. in the "real world"). The term Applied Mathematics, it seems to me, covers a spectrum ranging from *commissioned problems* to *applicable methods.* But anywhere along this spectrum, Applied Mathematics has a purpose-story (whether explicit or implicit) at its very foundation.

Some purpose-stories do not connect with ends that could be labeled as practical. It often surprises me how even the tiniest sliver of such a purpose-story manages to focus the mind, and to clarify ideas. To begin to think about this, let us consider the "value-added" comprehension afforded to us, in Archimedes' *Sand-reckoner* (in which Archimedes establishes notation to describe very large numbers) when we read that the author wanted *to denote numbers larger than the number of grains of sand needed to fill up the universe.*

I believe that formulating this purpose– to denote numbers larger than the number of grains of sand in the universe– accomplishes three small, but important, pedagogical missions. First, it provides a very specific touchstone to judge whether the enterprise (finding notation for large numbers) will be deemed a success. A frame– admittedly a very large frame– has been put around the project, and we know what we are aiming for: according to Archimedes the number of grains tote up to 10^{63} (in our notation). Second, it is whispering something to us about the relationship between *matter* and *idea.* Here a language is being developed–Archimedes tells us–in the realm of idea, for quantities that are beyond those that could ever be realized by mere matter. Third, it emphasizes that this notational language can be precisely handled and understood, for the most specific of aims, even in the sublime range of numbers that have no material referent.

The above analysis of the pedagogical importance of the *sand* in Archimedes' treatise may be too simpleminded, but I feel that any bona fide purpose-story in a mathematical exposition will have some subtle effect on the presentation of the pure mathematical content, and what that effect is may be worth understanding.

Origin-stories–how the mathematician came to work on the material, how ideas originated, are critically different from purpose-stories because, at the very least, an origin-story leaves the endgame of the project open, although it claims to pinpoint the beginning-game. An example for this is given by the celebrated treatise of Archimedes (lost to us until 1906). This was the letter written to Eratosthenes that Archimedes called simply *The Method* and in which he proposed to offer what he called the "mechanical" method that preceded and led to many of his mathematical discoveries.

The fourth category of story in mathematical exposition, the one far more difficult to categorize than the others, is the kind of narrative that is traced out by the ideas themselves as they unfold. The fundamental

** John Derbeyshire's book on the Riemann Hypothesis, however, manages to strike a fine balance.

propelling force of any story is made palpable by the energy behind the question “and then?” that the listener asks. The yearning to do this–to ask “What happens next?”–in the realm of ideas, the impulse usually called *intellectual curiosity*, is itself a most curious thing. To begin, it presumes that there is a “next.” Narrative depends critically on our being creatures in time. Any argumentation requires this, as well. Given X we proceed to Y, and thereby weave an ordered succession through an otherwise atemporal intellectual realm. What this entails deserves to be thought through, but I won't try to do that here*. Suffice to say that a pattern of suspense, however muted, must be felt; and the standard engine that drives suspense–namely, some form of conflict or tension–must be present.

All this is preamble to the telling of a current mystifying conundrum in number theory: a tension between heuristics on the one hand, and the data accumulated so far that goes counter to these heuristics on the other, a conflict that has something of a see-saw history. It will be largely *story* and–at least in terms of my personal contribution—no *result* (yet). Nevertheless, some beautiful results will be quoted, and this conflict raises the question of whether we as mathematicians may, at times, face a situation where the substance we study has one shape *asymptotically*, and yet all computational evidence elucidating this substance, even up to very large numbers, seem consistent with the possibility that the data have a different asymptotic shape.

There are some traditional mathematical stories with a similar flavor: where the “early returns” of the data lead one to make incorrect guesses about the general state of things. For example, there is the history of Skewes' Number, a very large number proved to be a lower bound for the first crossing point of $\pi(x)$ and $Li(x)$; this theorem is amusing on many fronts, one of which being that it is one of the few famous mathematical results that, if it were weakened, would become all the more impressive. There is also the lesser known conjecture of Kummer on the distribution of phases of cubic Gauss sums. Kummer conjectured, based on the numerical evidence at his disposal that the phases should occur with frequencies $1:2:3$. Of course, he hardly computed very far, given today's standards, but still it was something of a surprise when Samueal Patterson proved that, despite the data, the actual frequency distribution was $1:1:1$.

But the story I wish to tell has the virtue of being on-going: besides conflict, there is indeed, suspense. I shall try to tell it with a minimum of technical language. It is all about counting the percentage of objects of a certain type, within a larger collection of objects.

Let us begin with the simple question: *what is the percentage of even numbers among all positive integers?* Of course, if we range the positive integers in their natural increasing order and compute the limit of the percentages computed with respect to the increasing collections of positive numbers we encounter, we get 50%. But we also know that if we want to be perverse, we can rearrange the sequence of positive integers to produce percentages that limit to anything we want. Is there, however, an ordering of the positive integers that is natural to consider in terms of evenness and oddness, and with respect to which the even numbers occur more (or less) than 50% of the time. The wild card in the previous sentence is the word *natural,* a word that Aristotle's *Metaphysics* teaches us, if nothing else, to be wary of. I don't happen to have any *natural* candidate for re-ordering that will upset the 50% statistics of even versus odd, and I don't expect to come up with any, but I raise this question only to remind us that − in any such discussion all the data we collect depends upon an initial judgment–the reasonableness of which is difficult in itself to assess–that we have set up a *natural* ordering of the instances we intend to count.

This suggests that we devote a few lines to the structure of “orderings” per se. For the purposes of computing the percentages that we will be interested, we don't really need total orderings. We can make do with something less, that I will call *nestings.* If C is a set, by a **nesting** of C let us mean a nested sequence of *finite* subsets

$$C_1 \subset C_2 \subset \ldots \subset C_i \subset \ldots \subset C$$

* Jean-Francois Burnol recently sent me an e-mail saying that in doing this we are like voyagers exploring a spider web who will “par la force des choses” create a temporal ordering which can be but a pale incarnation of the full structure.

such that $\cup_i C_i = C$. Fix a nesting of C. If $X \subset C$ is a subset, say that X has a **percentage** if the limit of

$$\frac{\#\{X \cap C_i\}}{\#\{C_i\}}$$

exists as i tends to ∞, and say that X **consists of** $A\%$ **of** C if this limit, when expressed as a percentage, is A. This is all relative to the fixed nesting, of course. Any infinite subsequence of a nesting is again a nesting; when we pass from a given nesting to one of its subsequences, let us agree to re-index the subsequence with the indices $1, 2, \ldots$. Say that two nestings $\{C_i\}_i$ and $\{C'_i\}_i$ of C are **equivalent** if there are infinite subsequences of each, which when re-indexed as described above, have the property that the limits of $\frac{\#\{C'_i \cap C_i\}}{\#\{C_i\}}$ and of $\frac{\#\{C'_i \cap C_i\}}{\#\{C'_i\}}$ are 100%. It isn't difficult to see that if $X \subset C$ is a subset that *has a percentage* with respect to each of two nestings $\{C_i\}_i$ and $\{C'_i\}_i$, then if $\{C_i\}_i$ and $\{C'_i\}_i$ are equivalent, the percentage that X is of C, computed with respect to each of these nestings is the same.

So, in terms of the vocabulary of the previous paragraph, if we encounter two nestings of a collection of mathematical objects, and compute the percentage that a given subset of these objects comprises– computed relative to each of the nestings–it is a priori conceivable that we get different statistics, but only if the two nestings are inequivalent.

Now let us count some mathematical objects. For example, quartic fields. What percentage, for example of all quartic fields have the property that the Galois group of their Galois closure is the full symmetric group, S_4? The above discussion surely puts us on warning that all will depend upon what objects, exactly, we propose to count, and how we order, or nest, these objects. Anyone who has taught Galois theory knows how finicky you have to be, when choosing the coefficients of a fourth degree polynomial, if you want a root of that polynomial to generate anything other than a field whose Galois group is other than S_4. Hilbert's irreducibility theorem provides corroboration of this with a proof that if you rank algebraic numbers of degree 4 by the size of the coefficients of their minimal polynomial (monic, over $\mathbf{Q}$) 100% of them have Galois group S_4.

But let us count quartic fields (rather than algebraic numbers that generate them) nested by the size (absolute value) of their discriminant. A few months ago, before I saw Manjul Barghava's astounding preprint, *The density of discriminants of quartic rings and fields.* Manjul first calculates the basic asymptotics of these fields. To give an example of the sort of detailed information he gives, he shows:

$$\frac{\#\text{ of quartic fields, } \{\text{totally real, disc } \in [0, X]\}}{X}$$

tends to

$$\{\frac{1}{48}\}\prod_p (1 + p^{-2} - p^{-3} - p^{-4})$$

as X goes to ∞. Here the product is over all prime numbers p and as indicated, we are considering only totally real fields with positive discriminant. Change, for example, *totally real* to *totally complex* and the $\frac{1}{48}$ changes to $\frac{1}{8}$; put whatever local conditions you want, and the $\frac{1}{48}$ changes accordingly.

This is already a significant extension of a tradition of counting fields of given low degree nested by discriminant. The enterprise was started in earnest for fields of degree 3 by results of Davenport and Heilbrunn who studied asymptotics for the number of $GL_2(\mathbf{Z})$-equivalence classes of binary cubic forms. This was generalized by Shintani via his theory offering analytic continuation of Dirichlet series associated to pre-homogenous vector space (in counting cubic fields, Shintani finds simple poles at $s = 1$ and $s = 5/6$ and computes the residues at each of these poles, thereby getting the dominant term in the asymptotics together with a precise first-error term). All this has been deepened and refined by the work of Wright, Datskowski, and more recently, Yukie.

Bhargava, however, takes a different tack, avoiding Dirichlet series altogether; he thinks of the problem of counting quartic fields as a problem purely in the Geometry of Numbers, and discovers (and actually *proves*) that precisely $90.644\ldots\%$ of these quartic fields have Galois group S_4 and a remaining $9.356\ldots\%$ have Galois group the dihedral subgroup $D_4 \subset S_4$.

Bhargava's result is the first (proved result) of its sort giving at least *one* example of a proper (transitive) subgroup $G \subset S_n$ which is the Galois group of a positive proportion of fields of degree n, when these fields are nested by size of discriminant. [*Note: Add references to general conjectures about this phenomenon that Bob Guralnik mentioned to me.*]

With this as a backdrop, I wish to pass to the counting question I am currently puzzled by, where data– massive though it may be– does not give indubitable confirmation of the heuristics that a large number of mathematicians actually believe; and, nevertheless, the heuristics are indeed believed. The topic is rational points on elliptic curves.

Elliptic curves have been connected to many of the recent advances in the subject, Fermat's Last Theorem being among them. But elliptic curves have played a major role in mathematics, from early on: in mechanics, abelian integrals, Riemann surfaces, the theory of doubly periodic analytic functions, and automorphic forms. These elliptic curves amply repay the obsessive interest that mathematicians have for them: the tightness of their structure gives us unexpected leverage in applying them to solve a host of problems. Elliptic curves seem to be designed to teach us things, and nowhere less than in arithmetic.

Rational points on elliptic curves are the gems of the arithmetic theory: they are, to diophantine geometry, what units in rings of integers are to algebraic number theory, what algebraic cycles are to algebraic geometry. A rational point in just the right context, at one place in the theory, can inhibit, and control–thanks to ideas of Kolyvagin–the existence of rational points and other mathematical structures elsewhere. Despite all that we presently know about these objects, the initial mystery and excitement that drew mathematicians to this arena in the first place remains in full force today. And with all this, rational points on elliptic curves are the most elementary of mathematical objects to describe.

In down-to-earth language, an *elliptic curve* (over $\mathbf{Q}$) is given by a cubic equation in two variables of the form

$$E: \quad Y^2 = X^3 + aX + b$$

where a, b are rational numbers and such that $X^3 + aX + b$ does not have a double or triple root. The *rational points* we will be interested in, on the elliptic curve E, will be pairs of rational numbers (x, y) that solve the equation upon substituting x for X and y for Y. (There is also a single rational point at infinity in the projective plane but I will ignore it, and discuss only the remaining points in the finite plane.) We consider these curves over $\mathbf{Q}$ up to isomorphism, and rank them according to the size of their *conductor**. There are only finitely many elliptic curves over $\mathbf{Q}$ of a given conductor, so that the size of the conductor does indeed give us a nesting.

The cubic curves in the plane given by the equation displayed above have the property that any straight line that intersects the curve in two rational points is either tangent to the curve at one of these points, is vertical, or else intersects the curve at a unique third point, and if it does, that third point is rational. This provides us with a mechanism for getting new rational points on an elliptic curve from old: even starting with one meager rational point, you can consider the tangent line to the curve at that point, and see if the

* The conductor of an elliptic curve E over $\mathbf{Q}$ is a positive integer $N(E)$ that is a measure of how well the elliptic curve *reduces* modulo various primes. A prime $p \geq 5$, for example, divides the conductor $N(E)$ only if there is no way of modifying the defining equation above of E so that when reduced modulo p we obtain an equation over the field $\mathbf{F}_p$ without multiple roots; the maximal power of such a prime p dividing $N(E)$ is 2 and whether it is 1 or 2 is determined by the nature of the "best" reduction of E modulo p, i.e., whether its defining cubic polynomial has a double or a triple root. There is a slightly more involved, but elementary, recipe to give the maximal power of the prime 2 and of the prime 3 dividing the conductor.

line intersects the elliptic curve in another point– if so this new point will be again rational, and you can then draw the line between them and hope for a third intersection point, which– again– will be rational. This manner of propagating points was called in the old literature the *chord and tangent process.* Nowadays one rather prefers to understand this in terms of the algebraic group structure of the elliptic curve, but the older *chord and tangent* language will do perfectly well for the questions we are about to ask.

To distinguish different *rational point behavior* that our elliptic curves may satisfy, let's first introduce a rough-and-ready numerical invariant that I will call **type**. Say that an elliptic curve E is of **type** 0 if its equation, $Y^2 = X^3 + aX + b$, has no rational points at all. Say that it is of **type** 1 if it has a rational point (x, y) such that starting with it, and producing the rational points that can be generated from it by iterated application of the chord and tangent process we encounter all the rational points on E. In general, say that the elliptic curve E is of **type** r (for $r > 0$) if there are r distinct rational points on E such that all rational points of E can be generated by assiduous application of the chord and tangent process to these r points, and r is the smallest number for which this is possible. A theorem of Mordell guarantees that every elliptic curve has a type, as described above; that is, for any E there are finitely many rational points of E such that all other rational points on E can be generated via iterating the the chord and tangent process, starting from these finitely many points.

People familiar with this theory will see that the notion *type* defined above is very close to the rank of the Mordell-Weil group $E(\mathbf{Q})$. This $E(\mathbf{Q})$ is an abelian group whose points consist of the rational points of the equation equation, $Y^2 = X^3 + aX + b$ defining E together with the point at ∞ that I had excluded; moreover, ∞ is taken as the origin of the group, and collinear rational points on E sum to zero, where *sum* is meant in terms of the group law on $E(\mathbf{Q})$. Mordell's theorem, quoted above, tells us that $E(\mathbf{Q})$ is finitely generated, and therefore it is isomorphic to a group of the shape $\mathbf{Z}^r \oplus \Phi$, where the integer r is often called simply the **rank** of E (over $\mathbf{Q}$). The group Φ is a finite abelian group that is either (trivial,) cyclic, or a sum of two cyclic groups. Our *type* then differs from *rank* only if Φ is nonzero.

For any $r = 0, 1, 2, \ldots$ the question we now may ask is: what percentage of elliptic curves (nested according to size of conductor) are of type r, how many of rank r? More correctly, we should ask: do these percentages *exist,* and if so what are they?

A conjecture* (for which there is evidence, both theoretical, and numerical) would tell us that there is no difference if we use *type* or *rank* in computing percentages for the families of elliptic curves that we will be considering. From now on I'll use *rank.* We will be focusing our attention on the friction between available *data* regarding the percentage distribution for the different *ranks* of elliptic curves, and the emerging *heuristics* or *educated guesses* predicting these percentages.

One fairly firm anchor in our theory is a principle that goes under the heading of **parity**. This principle is still only conjectural**, but is amply confirmed numerically in our accumulated data, and we also have theoretical reasons to believe it. The **Parity Principle** is that 50% *of the members of any of the families of elliptic curves we will be considering, nested by conductor, have even rank ($r = 0, 2, 4, \ldots$) and* 50% *have odd rank ($r = 1, 3, 5, \ldots$).*

Now we can introduce a major protagonist; namely, it is a conjecture that might be called **the minimalist principle**. In general terms, the *minimalist principle* proclaims that from the rough viewpoint of percentages, there are as few of these gems, rational points on elliptic curves, as is possible, given the constraint of the parity principle. That is, 50% *of the members of any of the families of elliptic curves we will be considering, nested by conductor, have rank $r = 0$,* 50% *have rank $r = 1$ and the remaining ranks $r \geq 2$ account for* 0% *of the family.*

I don't know whether anyone has actually conjectured this minimalist principle (taken over *all* elliptic curves over $\mathbf{Q}$) in print, and this already should be telling us how much at sea we are. But as one thing or

* that the subset of elliptic curves with nontrivial rational isogenies 0% among all elliptic curves

** *Maybe give a footnote here, for people knowledgeable about these kinds of results, describing the relationship between this assertion, various theorems about analytic parity, the standard parity conjecture...*

another things comes to light in the subject, the *minimalist* position is sometimes favored, and sometimes not. Who knows? Tomorrow, a Bhargava-like surprise might change the landscape.

For certain special families of elliptic curves this minimalist conjecture has long been in print, and has had a wild ride in terms of its being believed, and doubted. For example, restrict to the family of quadratic twists of a single elliptic curve, which in down-to-earth terms means that you should fix an elliptic curve

$$E_1 : \quad Y^2 = X^3 + a_0 X + b_0$$

and consider the specific family of elliptic curves

$$E_d : \quad Y^2 = X^3 + a_0 d^2 X + b_0 d^3$$

where d ranges through square-free integers.

[*Note: Here I will eventually separate my discussion in two different sections of this article, one treating only quadratic twist families, and the other treating the $X^3 + Y^3 = d$ family. Also, I'll go much more specifically into the Random Matrix predictions, with transparencies. So what is done below is very very rough. But to continue ...*]

You might ask for the percentages for ranks $r = 0, 1, 2, \ldots$ computed within this collection of elliptic curves, nested by increasing $|d|$, or, equivalently, nested by increasing conductor. The minimalist conjecture, initially framed by Dorian Goldfeld in 1979, would predict, again, that 50% of them are of rank 0, 50% of them are of rank 1, and the remaining ranks are a flat 0%. It seemed safe when Goldfeld published it, and in fact, was surely–in some vague form–a folk conjecture long before it got to print.

A shadow loomed over Goldfeld's conjecture as people (Zagier-Kramarz in 1987, and later Fermigier) amassed data that seemed not support it. The Zagier-Kramarz data concerned a family of twists of elliptic curves slightly different from the ones discussed above, but the implications drawn from it seemed dangerous for the minimalist's hopes, even for the quadratic twist families. The numerical evidence gathered by Zagier and Kramarz strongly suggested that, at least for a family of cubic twists of the elliptic curve with $j = 0$, i.e., for curves of the form

$$X^3 + Y^3 = d$$

with d cube-free, a positive percentage of members of this family–in fact, a rather high percentage at that–have Mordell-Weil rank equal to 2. Even worse, this percentage seemed (at least visually, if we eyeball the graphs in question) "flat" over a large range.

The story of this diophantine equation,$X^3 + Y^3 = d$, just taken by itself, is a fascinating one. This equation even makes an implicit appearance in the celebrated tale, told by G.H. Hardy, of going to visit the hospitalized Ramanujan, where–to cheer him up, Hardy chatted about cricket scores, and mentioned that the license number of the taxicab that brought Hardy to the hospital was 1729. Ramanujan pointed out that this number, 1729, distinguished itself as the smallest positive integer that can be expressed as the sum of two cubes in two distinct ways (anyone familiar with the theory of elliptic curves will very likely see, right off the bat, one of these ways).

It is fair to say that any conjecture can hold out only for so long in the face of large quantities of numerical evidence accumulating against it. So it is no surprise that Goldfeld's minimalist conjecture suffered a period of benign avoidance by people working in the field in the late 1980s. This period ended, and prospects brightened, with the appearance of some new heuristic ideas. At present, because of these new ideas predicting that Goldfeld's minimalist conjecture might be correct after all, and because these ideas are buttressed by more recent powerful number-crunching, doubts seem to have (definitively?) quieted.

To start with the number-crunching, let us consider the state of the art in "percentage data" for the elliptic curve family cut out by the equation $X^3 + Y^3 = d$. This is given by Mark Watkins' recent preprint

Rank distribution in a family of cubic twists where Watkins computes the numbers of cube-free d's such that the corresponding elliptic curve has even *positive* rank for $d < 10^7$ and gets the following graph.

INSERT WATKIN'S GRAPH, p. 10 OF HIS PREPRINT HERE

The solid line on the graph does start to bend below the straight dotted line. To check this you may need a magnifying glass. Moreover, to actually give comfort to the minimalist conjecture, that solid line must eventually dip below *any* line with positive slope, and at least in the present picture it seems to be doing its dipping awfully slowly, so the comfort is, perhaps, a bit cold.

So, what are the plausibility arguments that suggest that the Goldfeld conjecture for the family of quadratic twists of any elliptic curve might be correct after all? The basic such argument which I'll call the *elementary heuristic* is due I believe to Sarnak, and depends on guessing the odds (given Ramanujan's conjecture) that a coefficient of a modular form of weight 3/2 vanishes. The connection between the coefficients in question and the conjecture of Goldfeld depends on much of the recent arithmetic theory of elliptic curves, and in particular, on the work of Waldspurger, so it is–of course– hardly *elementary.* But a sharper heuristic comes from the work of Katz-Sarnak, and is related to random matrix models.

[Note: I'll give detailed discussion here, with references, of course.]]

Each of these arguments, the "elementary one" and the "random matrix one," offer reasonably precise predictions of the number of rank ≥ 2 elliptic curves of conductor $\leq X$ that are quadratic twists of a given elliptic curve (and these two predictions are consistent: the random matrix prediction being simply more precise). In terms of the rough percentages that we have been discussing, they assure us that we can expect 0% for ranks ≥ 2, despite the reluctance of the accumulated data so far (e.g., for the analogous family of cubic twists) to exhibit this behavior.

Our discussion has been about the relatively restricted question of percentages for ranks $= 0, 1, 2, \ldots$ computed over quadratic or cubic twists of a given elliptic curve. Let us now return to the question of percentages for these ranks computed over all elliptic curves, nested by conductor. As for the "very early returns," Of the first 200 elliptic curves ordered in terms of conductor, 190 of them have rank 0, and the remaining ten have rank 1. Clearly, then for any asymptotic statistics to kick in, we need to compute in much greater bulk.

The main sources of serious data are Brumer-McGuiness (1990), Cremona (1997), Stein-Watkins (2002), and recent work, in preparation, by Bektemirov-Stein-Watkins. The ground-rules for the different data-bases vary. Some of them count only of curves of prime conductor, some of them curves of square-free conductor, and some all conductors. Moreover, since the conductor and rank of an elliptic curve are invariants of the **Q**-isogeny class of the elliptic curve, one has as the option of counting **Q**-isogeny classes, rather than **Q**-isomorphism classes, of elliptic curves, and some of these sources do that; this choice, very likely, will not be skewing the data since taking some known results into account and the conjecture framed in footnote * above, the percentages will be the same if one lists **Q**-isogeny classes or **Q**-isomorphism classes. Most difficult to gauge, though, is whether the data-bases actually contain *all* the elliptic curves that their labels say they contain. This is a serious issue, and in the preprint under preparation by Bektemirov, Stein and Watkins, *Data about average ranks of elliptic curves*, they deal with the question of whether the vast Stein-Watkins data-base of "all" elliptic curves of prime conductor $< 10^{10}$ contain all the curves it is supposed to contain.

To make the point I wish to make in this story, however, all I need do is to borrow two graphs from Bektemirov, Stein and Watkins (*with their permission*). The first graph makes the minimalist philosophy appear to be a hopeless cause. It computes percentages of ranks $0, 1, 2, 3$ for all **Q**-isomorphism classes of elliptic curves of conductor between 10^5 and 10^8:

INSERT FIGURE 2; Bektemirov-Stein-Watkins' graph, p. 6 of their preprint here

Now this is something that should give even the most stalwart minimalist conjecturer pause, for even though we haven't yet reached Archimedes' *Sand-reckoner number*, by any reckoning we have a large collection of data. But let us not give up hope, and just go to larger ranges, and zoom things to a size where we might discern even modest trends. Here, then, is a picture of the average Mordell-Weil rank of elliptic curves in the part of the Stein-Watkins data-base, referred to above, consisting of prime conductors between 10^8 and 10^{10}:

INSERT FIGURE 3; the blue graph in
http://modular.fas.harvard.edu/baur/www/graphs/allcurves-p-counts.html

Here an ever-so-slightly downward sloping graph appears, gladdening the heart of the died-in-the-wool minimalist. The minimalist guess, by the way, is that that this curves approaches 0.50 asymptotically.

In broad outline, then, this is where we are today. We have a network of intuition about the nature of this important mathematical object–rational points on elliptic curves–that is not wholeheartedly endorsed by this data. Only in the range of numbers that challenge our most powerful computers is there a glimmer of a suggestion that the data may yet come around to our expectations.

We also have a closely related, but structurally different, task ahead of us. How can we account for all this bulk Mordell-Weil rank? We have, to be sure, some asymptotic lower bounds for Mordell-Weil ranks ≥ 2 (some of these results are subject to a parity conjecture some not) but at present our proved results–i.e., the asymptotic collections of rational points we can actually produce–are all far below what we expect and our current methods don't give much hope for being sharpened significantly.

But forget all questions of asymptotics. Consider only the range $\leq 10^{10}$ of the Stein-Watkins data-base. Is there an argument other than just computing ranks for each of the elliptic curves in the data-base–is there a *pure thought* heuristic–that explains why we are witnessing so much Mordell-Weil rank? In a sense, these rational points are both analogous, and not analogous, to the physicist's *dark matter.* This large mass of rational points for elliptic curves of conductor $\leq 10^{10}$ is palpably *there.* We aren't in the dark about that. We are merely in the dark about how to give a satisfactory account of their being there.

We are, in a word, just at the very beginning of this story.

[*Endnote: During the conference, one of the participants with access to very powerful computing capabilities offered to try to extend the existent rank-two data to encompass elliptic curves (say of prime conductor) of conductor* $< 10^{11}$. *This would be great.*]