PRIMES

Barry Mazur

April 26, 2014

(A discussion of 'Primes: What is Riemann's Hypothesis?,' the book I'm currently writing with William Stein)

William:

https://vimeo.com/90380011

Figure: William

The impact of the Riemann Hypothesis

Figure: Peter Sarnak

"The Riemann hypothesis is the central problem and it implies many, many things. One thing that makes it rather unusual in mathematics today is that there must be over five hundred papers—somebody should go and count—which start 'Assume the Riemann hypothesis,' and the conclusion is fantastic. And those [conclusions] would then become theorems ... With this one solution you would have proven five hundred theorems or more at once."

An expository challenge

The approach you take when you try to explain anything depends upon your intended audience(s). In our case we wanted to reach two quite different kinds of readers (at the same time):

- High School students who are already keen on mathematics,
- A somewhat older crowd of scientists (e.g., engineers) who have a nonprofessional interest in mathematics.

What sort of Hypothesis is the Riemann Hypothesis?

Consider the seemingly innocuous series of questions:

- *How many primes (2, 3, 5, 7, 11, 13, ...) are there less than 100?*
- *How many less than 10,000?*
- *How many less than 1,000,000?*

More generally, how many primes are there less than any given number X?

Riemann's Hypothesis tells us that a strikingly simple-to-describe function is a "very good approximation" to the number of primes less than a given number X. We now see that if we could prove this *Hypothesis of Riemann* we would have the key to a wealth of powerful mathematics. Mathematicians are eager to find that key.

An expository frame—and goal

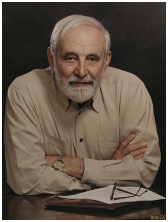


Figure: Raoul Bott (1923–2005)

Raoul Bott, once said—giving advice to some young mathematicians—that whenever one reads a mathematics book or article, or goes to a math lecture, one should aim to come home with something very specific (it can be small, but should be *specific*) that has application to a wider class of mathematical problem than was the focus of the text or lecture.

Setting the frame

If we were to suggest some possible *specific* items to come home with, after reading our book, three key phrases – **prime numbers**, **square-root accurate**, and **spectrum** – would head the list.

PRIMES: order appearing random

Figure: Don Zagier

"**[Primes]**

- *are the most arbitrary and ornery objects studied by mathematicians: they grow like weeds among the natural numbers, seeming to obey no other law than that of chance, and nobody can predict where the next one will sprout.*
- *exhibit stunning regularity . . . they obey their laws with almost military precision."*

How to nudge readers to feel the orneriness of primes

There is something compelling about ‘physically’ hunting for a species of mathematical object, and collecting specimens of it. Our book emphasizes this approach for our readers. Here are some routes that allow you to 'pan' (in different ways) for primes:

Factor trees and **Sieves**

and

Euclid’s Proof of the Infinitude of Primes.

Factor trees

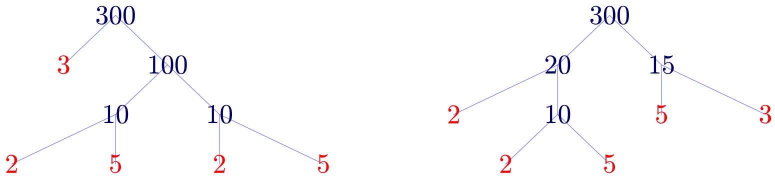


Figure:

Sieves

2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ~~20~~ ~~21~~ ~~22~~ 23 ~~24~~ ~~25~~ ~~26~~

The ubiquity of primes

Figure: Don Quixote and “his” Dulcinea del Toboso

Numbers are obstreperous things. Don Quixote encountered this when he requested that the “bachelor” compose a poem to his lady Dulcinea del Toboso, the first letters of each line spelling out her name.

The stubbornness of primes and knights

The "bachelor" found

> *"a great difficulty in their composition because the number of letters in her name was* 17, *and if he made four Castilian stanzas of four octosyllabic lines each, there would be one letter too many, and if he made the stanzas of five octosyllabic lines each, the ones called* décimas *or* redondillas, *there would be three letters too few..."*

"It must fit in, however, you do it," pleaded Quixote, not willing to grant the imperviousness of the number 17 to division.

The Art of asking questions

Questions anyone might ask

spawning

Questions that shape the field

Gaps: an example of a 'question anyone might ask'

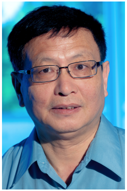


Figure: Yitang Zhang

In celebration of Yitang Zhang's recent result, consider the *gaps* between one prime and the next.

Twin Primes

As of 2014, the largest known twin primes are

$$3756801695685 \cdot 2^{666669} \pm 1$$

These enormous primes have 200700 *digits each.*

Gaps of width k

Define

$$\text{Gap}_k(X) :=$$

number of pairs of *consecutive* primes (p, q) with $q < X$ that have "gap k" (i.e., such that their difference $q - p$ is k).

NOTE: $\text{Gap}_4(10) = 0$.

Gap statistics

Table: Values of $\mathrm{Gap}_k(X)$

X	$\mathrm{Gap}_2(X)$	$\mathrm{Gap}_4(X)$	$\mathrm{Gap}_6(X)$	$\mathrm{Gap}_8(X)$	$\mathrm{Gap}_{100}(X)$	$\mathrm{Gap}_{252}(X)$
10	2	0	0	0	0	0
10^2	8	7	7	1	0	0
10^3	35	40	44	15	0	0
10^4	205	202	299	101	0	0
10^5	1224	1215	1940	773	0	0
10^6	8169	8143	13549	5569	2	0
10^7	58980	58621	99987	42352	36	0
10^8	440312	440257	768752	334180	878	0

How many primes are there?

$$\pi(X) := \# \text{ of primes} \leq X$$

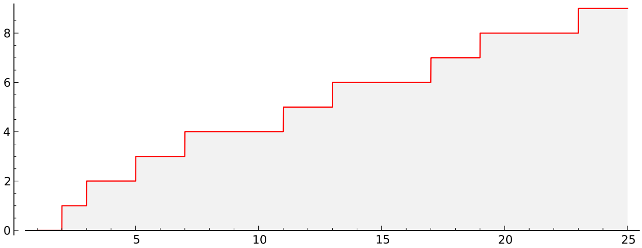


Figure: Staircase of primes up to 25

How many primes are there?

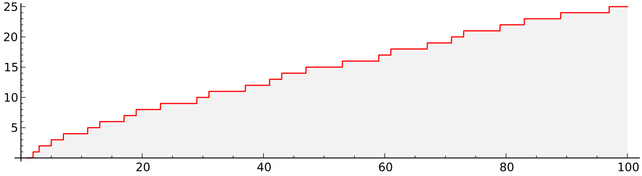


Figure: Staircase of primes up to 100

Prime numbers viewed from a distance

Pictures of data magically become smooth curves as you telescope to greater and greater ranges.

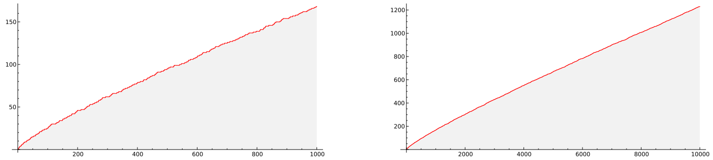


Figure: Staircases of primes up to 1,000 and 10,000

Proportion of Primes

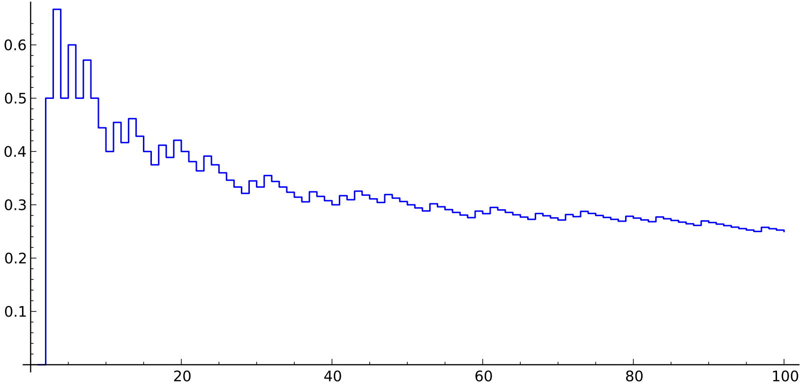
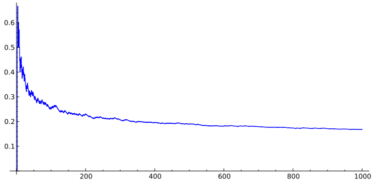


Figure: Graph of the proportion of primes up to X for each integer $X \leq 100$

Proportion of Primes at greater distance



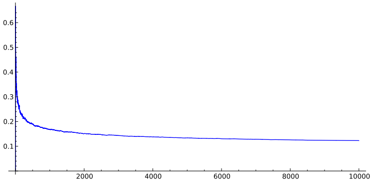


Figure: Proportion of primes for X up to 1,000 (left) and 10,000 (right)

Gauss

Unter	gebtes Primzahlen	Integral $\int \frac{dn}{\log n}$	Differ	Ihre Formel	Abweich.
500 000	41 556	41606,4	+50,4	41596,9	+40,9
1000 000	78 501	78627,5	+126,5	78672,7	+171,7
1500 000	114 112	114263,1	+151,1	114374,0	+264,0
2000 000	148883	149054,8	+171,8	149233,0	+350,0
2500 000	183016	183245,0	+229,0	183495,1	+479,1
3000 000	216745	216970,6	+225,6	217308,5	+563,6

Dass Legendre sich auch mit diesem Gegenstande beschäftigt hat, war mir nicht bekannt; auf Veranlassung Ihres Briefes habe ich in seiner Théorie des Nombres nachgesehen, und in der zweiten Ausgabe einige darauf bezügliche Seiten gefunden, die ich früher übersehen (oder seitdem vergessen) haben muss. Legendre gebraucht die Formel

$$\frac{n}{\log n - A}$$

Figure: A Letter of Gauss

Gauss' guess

The 'probability' that a number N is a prime is proportional to the reciprocal of its number of digits; more precisely the probability is

$$1/\log(N).$$

This would lead us to this guess for the approximate value of $\pi(X)$:

$$\mathrm{Li}(X) := \int_2^X dX/\log(X).$$

Approximating $\pi(X)$

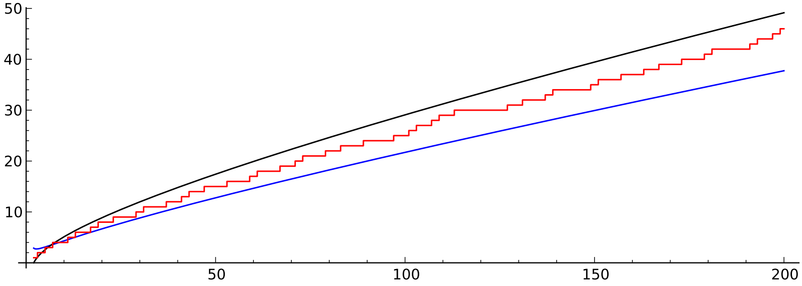


Figure: Plots of $\mathrm{Li}(X)$ (top), $\pi(X)$ (in the middle), and $X/\log(X)$ (bottom).

The Prime Number Theorem

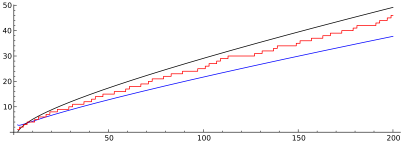


Figure: Plots of $\mathrm{Li}(X)$ (top), $\pi(X)$ (in the middle), and $X/\log(X)$ (bottom).

All three graphs *tend to* ∞ at the same rate.

Ratios

PNT:

The ratios

$$\frac{\pi(X)}{Li(X)} \quad \text{and} \quad \frac{\pi(X)}{X/\log(X)}$$

tend to 1 as X goes to ∞.

Ratios versus Differences

Much subtler question: what about their differences?

$$|\,\mathrm{Li}(X) - \pi(X)|?$$

Riemann's Hypothesis

The Riemann Hypothesis (first formulation)

$\pi(X)$ is approximated by $\mathrm{Li}(X)$, with **essentially square-root** accuracy.

More precisely ...

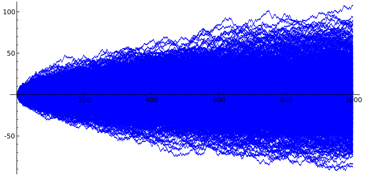
RH is equivalent to:

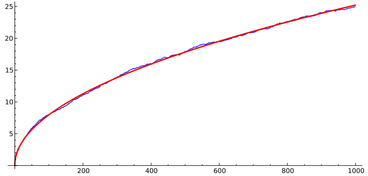
$$|\,\mathrm{Li}(X) - \pi(X)| \leq \sqrt{X}\log(X)$$

for all $X \geq 2.01$.

The gold standard for empirical data accuracy

Discussion of random error, and random walks





The mystery moves to the error term

$$\begin{aligned} \mathit{Mysterious\ quantity}(X) &= \\ &= \mathit{Simple\ expression}(X) + \\ &+ \mathit{Error}(X). \end{aligned}$$

Our mystery moves to our error term

$$\text{Mystery} \quad = \quad \text{Simple} \; + \; \text{Error.}$$

$$\pi(X) \quad = \quad Li(X) - \left(Li(X) - \pi(X)\right)$$

That 'error term'

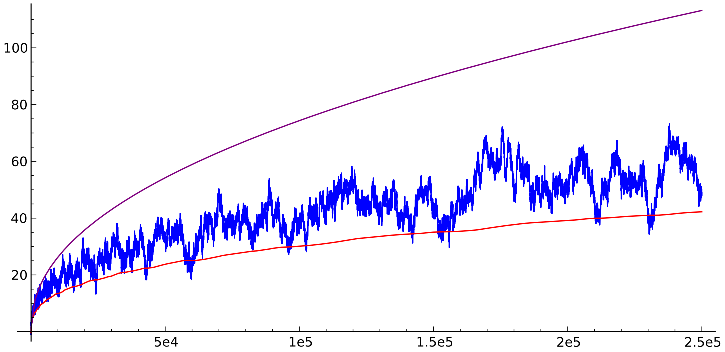


Figure: $\mathrm{Li}(x) - \pi(x)$ (blue middle), its Césaro smoothing (red bottom), and $\sqrt{\frac{2}{\pi}} \cdot \sqrt{x/\log(x)}$ (top), all for $x \leq 250{,}000$

The tension between data and long-range behavior

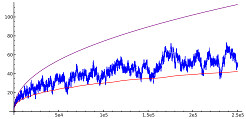
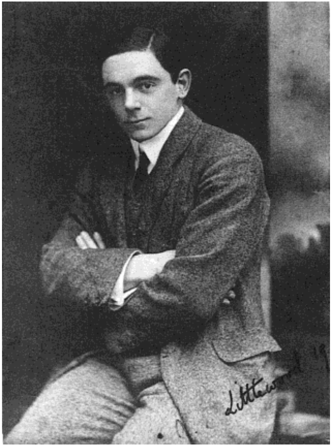


Figure:

The wiggly blue curve which seems to be growing nicely 'like $\sqrt{X}$' will descend below the X-axis, for some value of $X > 10^{14}$.

Skewes Number

The tension between data and long-range behavior



$$10^{14} \quad \leq \quad \text{Skewes Number} \quad < \quad 10^{317}$$

Spectrum

From Latin:
"image," or "appearance."

Spectra and the Fourier transform

(The essential miracle of the theory of the Fourier transform:)

$$G(t) \quad \leftrightarrow \quad F(s)$$

Each behaves as if it were the *'spectral analysis'* of the other.

packaging the information given by prime powers

$$g(t) \quad =$$

$$= \quad -\sum_{p^n} \frac{\log(p)}{p^{n/2}} \cos(t \log(p^n).)$$

$p^n \leq 5$

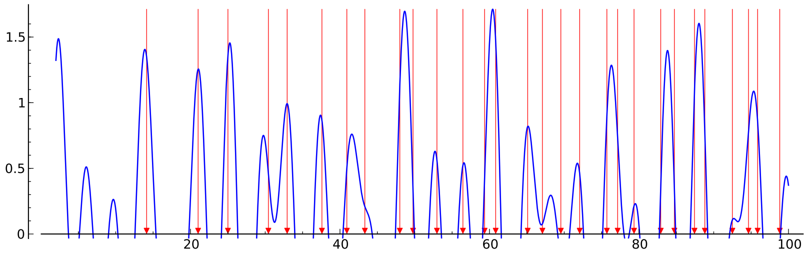


Figure: Plot of $-\sum_{p^n \leq 5} \frac{\log(p)}{p^{n/2}} \cos(t \log(p^n))$ with arrows pointing to the spectrum of the primes

$p^n \leq 20$

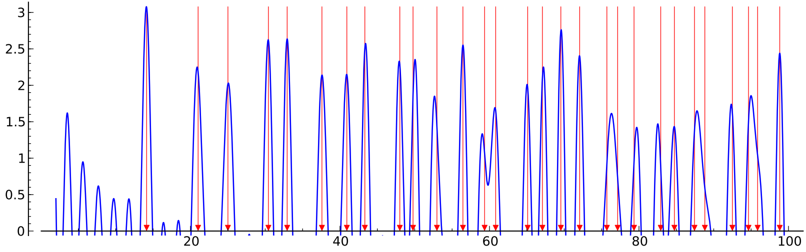


Figure: Plot of $-\sum_{p^n \leq 20} \frac{\log(p)}{p^{n/2}} \cos(t \log(p^n))$ with arrows pointing to the spectrum of the primes

$p^n \leq 50$

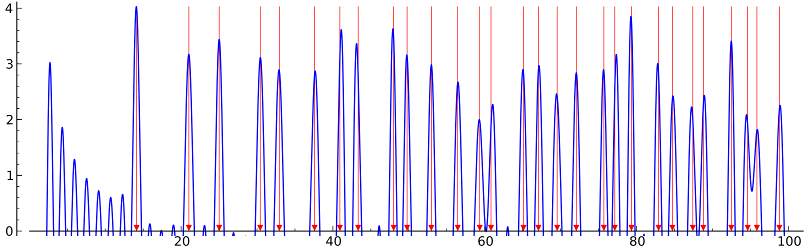


Figure: Plot of $-\sum_{p^n \leq 50} \frac{\log(p)}{p^{n/2}} \cos(t \log(p^n))$ with arrows pointing to the spectrum of the primes

$p^n \leq 500$

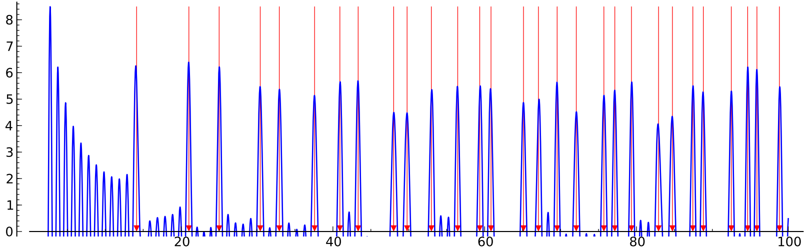


Figure: Plot of $-\sum_{p^n \leq 500} \frac{\log(p)}{p^{n/2}} \cos(t \log(p^n))$ with arrows pointing to the spectrum of the primes

From primes to the Riemann Spectrum

Conditional on RH, $g(t)$ converges to a distribution with singular spikes at the red vertical lines: the Riemann spectrum,

$$\theta_1, \theta_2, \theta_3, \ldots$$

From the Riemann Spectrum to primes

$$
\begin{aligned}
f(s) &= \\
&= 1 + \sum_i \cos(\theta_i \cdot \log(s)).
\end{aligned}
$$

From the Riemann Spectrum to primes

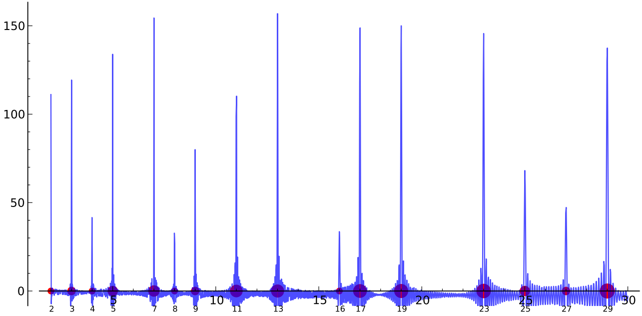


Figure: Illustration of $-\sum_{i=1}^{1000} \cos(\log(s)\theta_i)$, where $\theta_1 \sim 14.13, \ldots$ are the first 1000 contributions to the Riemann spectrum. The spikes are at the prime powers p^n, whose size is proportional to $\log(p)$.

From the Riemann Spectrum to primes

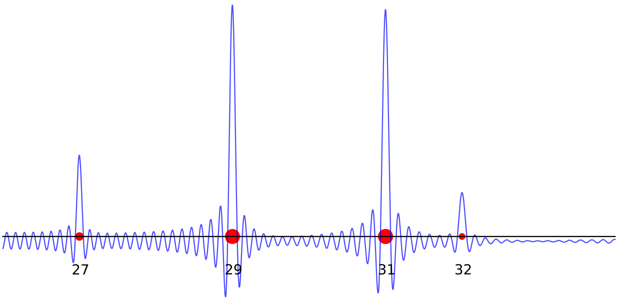


Figure: Illustration of $-\sum_{i=1}^{1000} \cos(\log(s)\theta_i)$ in the neighborhood of a twin prime. Notice how the two primes 29 and 31 are separated out by the Fourier series, and how the prime powers 3^3 and 2^5 also appear.

From the Riemann Spectrum to primes

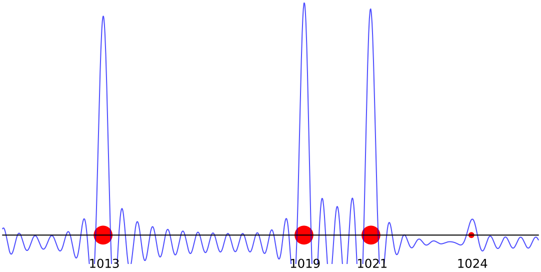


Figure: Fourier series from $1,000$ to $1,030$ using 15,000 of the numbers θ_i. Note the twin primes 1019 and 1021 and that $1024 = 2^{10}$.

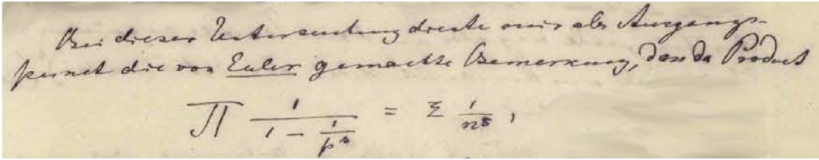
Information and Structure

The Riemann spectrum holds the key to the position of prime numbers on the number line.

What even deeper structure of primes can they reveal to us?

Riemann

Figure: Bernhard Riemann (1826–1866)



Bei dieser Untersuchung diente mir als Ausgangspunkt die von Euler gemachte Bemerkung, dass das Product

$$\prod \frac{1}{1-\frac{1}{p^s}} = \sum \frac{1}{n^s},$$

Figure: From Riemann's 1859 Manuscript

William

`https://vimeo.com/90380011`

Figure: William