ON THE HARMONIC FREQUENCY OF PRIMES

In Copes, Larry and Frances Rosamond (edd.) *Educational Transformations: Changing our lives through mathematics: A tribute to Stephen Ira Brown*. Bloomington: Authorhouse, 2006, pp. 37-58

Tomás Mario Kalmar
PO Box 402
Hardwick. Vermont
802 472-5978
kalmar.tomas@gmail.com
October 2003

ON THE HARMONIC FREQUENCY OF PRIMES

Dear Steve

For the consideration which your friends have shown to me by admitting me as one of the contributors to this your Festschrift, I believe I can best express my thanks by availing myself at once of the privilege thereby given me to communicate a Steve Brownish inquiry into the harmonic frequencies of prime numbers, a subject which, through the interest shown in it by Gauss, Dirichlet, Riemann, and now the Clay Institute, appears not altogether unworthy of such a communication.

I take as my starting-point for this inquiry the charming discovery made in 2001 by Heather, Morgan, Ezra, Brian and Steve at Goddard College in Vermont that

$$n = \pi \, \mathrm{H}_\pi$$

Here n is any whole number; π is short for $\pi(n)$, the number of primes less than n; and H_π is the π-th Harmonic Number, i.e.

$$\mathrm{H}_\pi = 1 + 1/2 + 1/3 + \ldots + 1/\pi.$$

I The New Kid on the Block Draws a Picture, Tells a Story

To tell the story of how $n = \pi \, \mathrm{H}_\pi$ first came to be written down a couple of years ago, a novelistic narrative would be best, set in California State University Monterey Bay (1995-1999) and Goddard College in Vermont (2000-2001), with dozens of characters and small groups caught in plots and subplots, until—as in a detective story—the formula $n = \pi \, \mathrm{H}_\pi$ pops up at the very end, as a punchline that suddenly resolves all doubts and brings everything together in a pleasing harmony.

We mathematicians do it backwards. First we tell the punchline. Then we tell the joke. *First* you state the theorem. *Then* you give the proof. On this present occasion decorum calls for me to tell you the *shortest* possible version of the joke whose punchline is $n = \pi \, \mathrm{H}_\pi$. I know no version shorter than the Three Minute Version I composed for my maiden speech at 3:40 pm June 9, 2001 as Goddard College's Mathematician-in-

Residence to the Northeastern Section of Mathematical Association of America Spring Meeting in Norwich University, Vermont. About half an hour from where I now live in Hardwick. My presentation was entitled “The New Kid on the Block”

Given the fact that Goddard College had not had a mathematician for fifteen years, it seemed important to establish myself immediately as a fellow professional and not as some court jester…. The first three minutes of my twenty minute talk must prove to everyone in the room that I am a mathematician. It seemed a good idea therefore to try out with my new colleagues my claim that Goddard students had found a new and simple Prime Number Theorem hitherto unknown to mathematicians, had seen simple interweaving harmonic patterns that explained why the primes behave the way they do. A tall claim. I did not intend to devote 20 minutes to the topic, just three or four.

The three minute version went more or less along the following lines.

At Goddard College we don’t teach courses, we facilitate group studies. In the Spring of 2001, I facilitated a group study of *Varieties of Mathematical Experience*. Once a week, for about ten weeks, off and on, I joined Steve, Brian, Ezra, Morgan and Heather for freewheeling conversations lasting up to three hours each. We sat on a couple of beat-up old couches in an old Victorian house and talked about many experiences, religious, mathematical and otherwise. We told jokes. We talked about why it was *obvious* that Steve’s version of a certain joke was better than mine, and along the way we did what the Books all tell us is a mathematical task beyond the capacity of all but the rarest geniuses: we visualized the distribution of the primes.

Draw a Picture, Tell a Story. To draw the picture, help me out with a round of “Guess My Rule.” Give me a number.

You give me 5 I give you 5.

You give me 7 I give you 7.

You give me 20 I give you 5

What’s my rule?

16 comes out 2

24 comes out 3

17 comes out 17

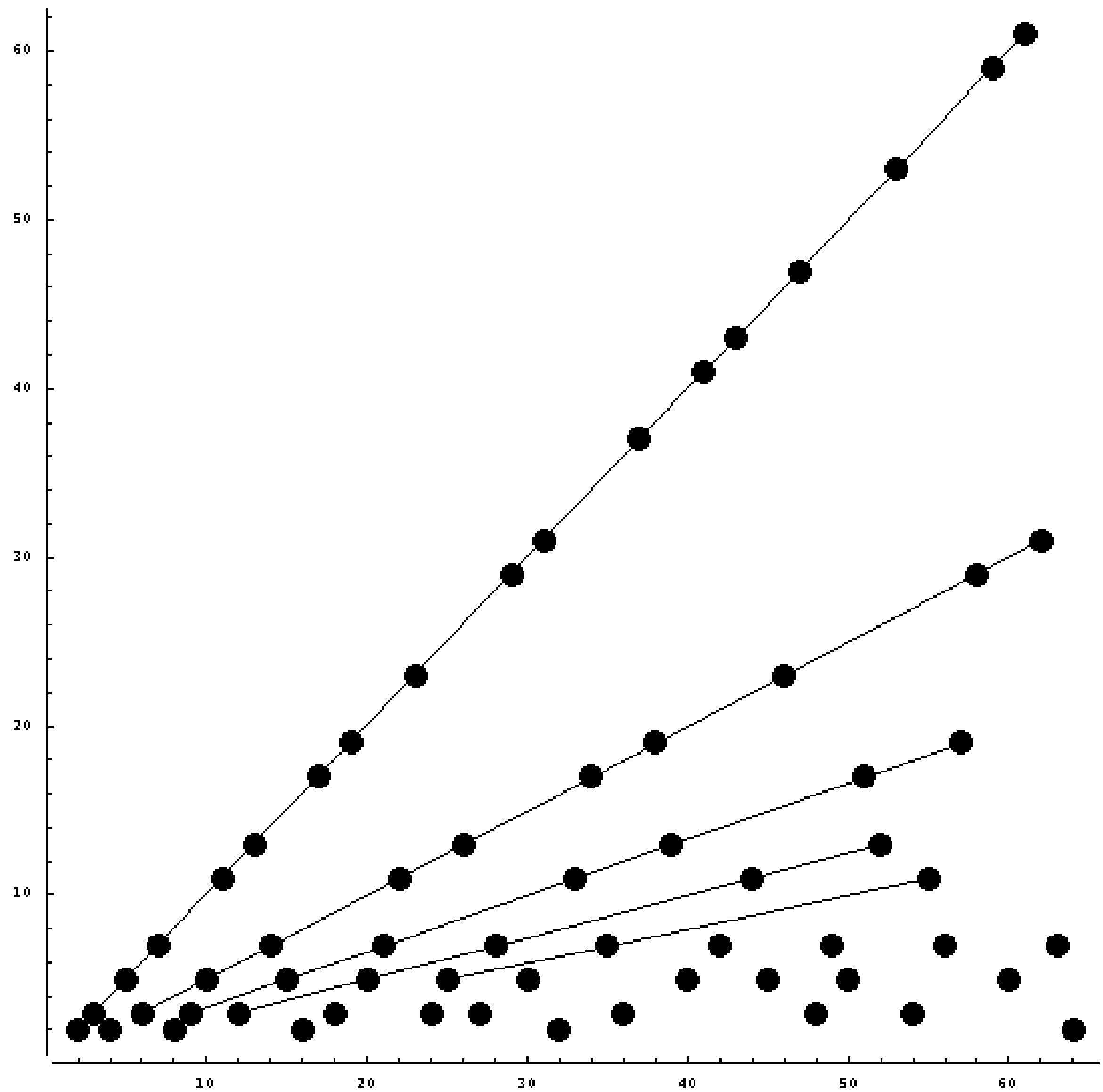
What's my rule?

….

Prime numbers come out prime. So do composite numbers. In plain English, what comes out is the highest prime factor of whatever went in. (*The canonical notation for this is* $P^{+}(n)$ *but I don't use this until the very end of this paper.*)

Your homework is to compile a table of the highest prime factors of the first fifty numbers and plot the function on a graph.

[*This is the moment, dear Reader, to take a break and have some fun doing the homework. Come back when you've had a chance to see the patterns in your own graph. Highlight them. Describe them. And, if you can, explain them. If you read on without doing the homework you may not get the joke — i.e. you may not experience the epiphany.*]



The following week Heather came in with three or four graph sheets pasted together. She said “There seemed to be a second sloping line underneath the first one, so I drew another fifty points, and then there seemed to be a third sloping line under the second one, so I drew another hundred points. Those sloping lines go on forever.”

Ezra, the class genius, had not done the homework. That is to say, he had not actually *drawn* the graph *qua* such.... He had done it in his head. He glanced at Heather’s graph and said “Oh, the slopes are 1, 1/2, 1/3, 1/4 … is there 1/5?”

Morgan had concentrated her attention on the *horizontal* lines. She proved that the highest prime factor of powers of 2 is always 2, and so those points always fall on the first horizontal line $y = 2$. (We called this Morgan's Theorem.) Powers of 2 times powers of 3 fall on the next horizontal lines, $y = 3$. And so on. But only for primes.

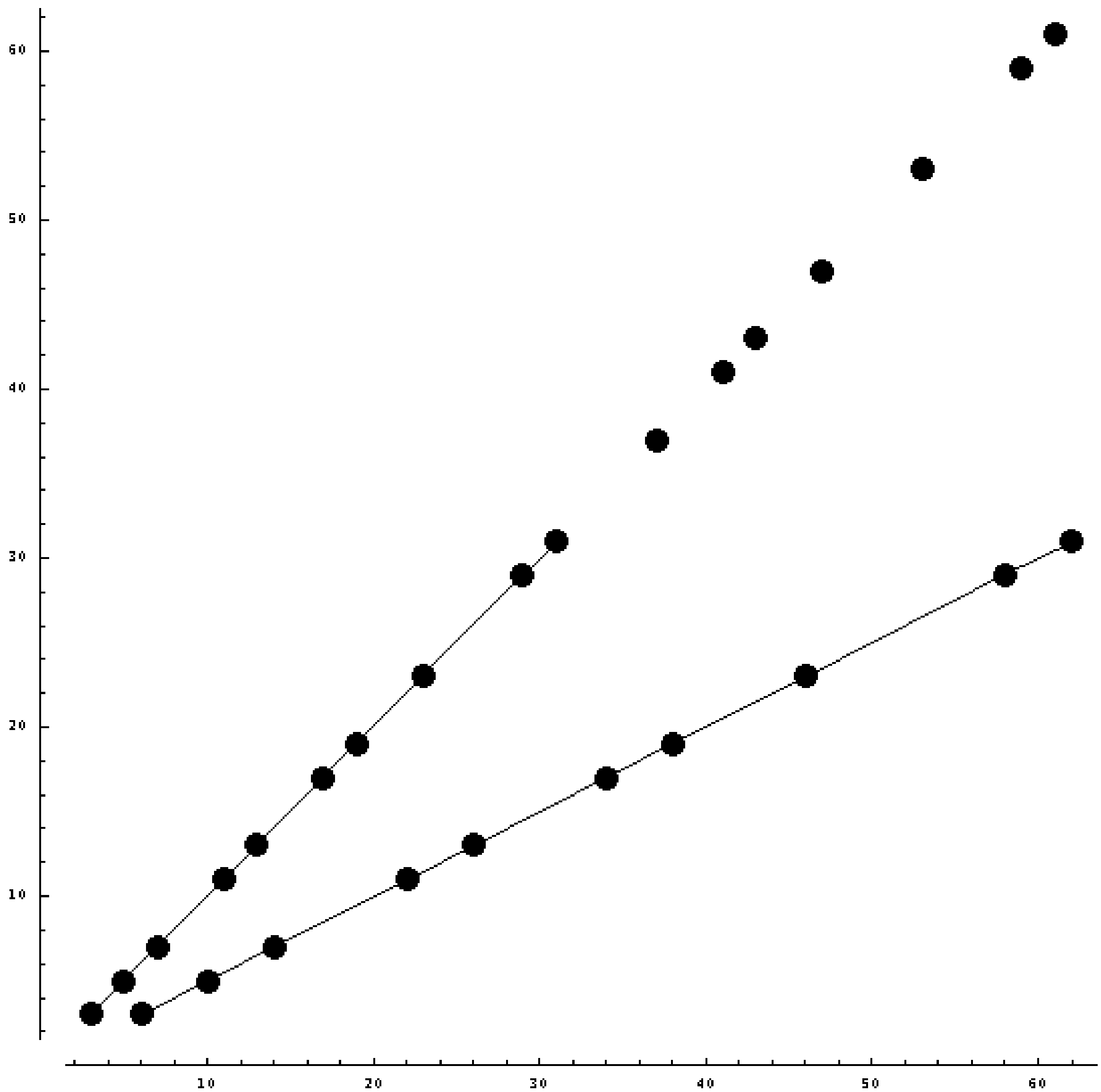
Someone else saw and drew curved patterns linking the points.

The first sloping line, $y = x$, was easy to explain. All primes, and nothing but primes, end up on this line. The highest prime factor of a prime is itself. Call this the prime line.

Points fall on the second sloping line, $y = x/2$, if and only if they are of the form $\{2p, p\}$, where p is prime.

(And so on? Do points fall on the h^{th} sloping line, $y = x/h$ iff they are of the form $\{hp, p\}$, where p is prime?}

The turning point comes when you see that the points on first and second sloping line have *the same pattern*.



The points on the second line are the primes spaced out twice as far apart as on the first line. The same goes for all the other sloping lines. Call them *harmonics* of the prime line. (To keep things simple we called the prime line the *first harmonic*.)

The points on the h^{th} harmonic are the primes spaced out h times as far apart as on the first line.

Now sum it all up. You start out with n points. They fall on a series of "harmonics". The first has $\pi(n)$ points on it, call it π for short. The second has about half as many,

$\pi/2$. The third $\pi/3$. And so on. Add all these points up and you get back to *n*. The number of points you began with is

$$\pi + \pi/2 + \pi/3 + ... = \pi(1 + 1/2 + 1/3 + \ldots)$$

You don't go on forever. Where do you stop? At the last *h*, the last sloping line. What *is* the last harmonic, the last sloping line on this graph? What is its slope $1/h$?

Since there are only *n* points to begin with, you don't need to go beyond $1/n$.

If, as mathematicians customarily do, you write H_n for the *n*th Harmonic Number, you can sum it all up as

$$n = \pi\, H_n$$

This is strikingly like Gauss's famous formula

$$n = \pi \log n.$$

But whereas Gauss had to count millions of primes to see his logarithmic patterns, you can get $n = \pi\, H_n$ by seeing simple patterns in fifty points.

To wrap up this phase of our long conversations, Heather conjectured that in practice there are, on the average, only about as many harmonics actually visible on a given graph as there are points on the prime line, i.e. π. We applauded the Heatherian Hypothesis and the closure it offered us. The higher harmonics do look as if they're lucky to have just one point on them. Going past the π-th harmonic is point-less. If *h* is more than π, $\pi/h < 1$, so, rounding off, it's really 0.

So we ended the conversation and the group study with our improved and more elegant version of the formula:

$$n = \pi + \pi/2 + \pi/3 + ... + \pi/\pi$$

$$= \pi\ (1 + 1/2 + 1/3 + \ldots + 1/\pi)$$

$$\text{i.e. } n = \pi\, H_{\pi}$$

…

As an opening move in my maiden speech, this Three Minute Version worked well. It put Goddard on the mathematical map at the local gathering of the regional MAA. There was a nice long silence as real mathematicians, with Ph D's and all, gazed at the new mystic formula written by hand on the blackboard, and absorbed what it implied. Then I erased the board and for the remaining fifteen minutes I felt free to regale my new colleagues with stories of my adventures as Mathematician in Residence at Goddard College, secure in the knowledge that my liturgical act of writing a new theorem on the blackboard (it has to be true, it has to be new, it has to be by hand, it has to be with chalk) legitimized my membership in the community of mathematicians. I had satisfied Dieudonné's criterion: a mathematician is someone who publishes the proof of a nontrivial theorem. On that day, on that blackboard, on that occasion, for the New Kid on the Block to write

$$n = \pi \, \mathrm{H}_{\pi}$$

was to publish a nontrivial theorem.

II Reflections on an apocryphal theorem.

Credit for this joke, with its memorable punch line, goes to the students at Goddard College and California State University Monterey Bay. It is *their* epiphany that I am celebrating here. As Mathematician in Residence, I served as the scribe whose vocation it is to write down in canonical notation what his clients are settling among themselves in the vernacular.

I confess, dear Steve, that I have filled many pages, nay, notebooks, with reflections on this story. And I must further confess that whenever I write in my math journals I'm always really writing to you, for it was you who taught me that it's OK to write about math the way I do. So I feel it only decent to offer you a glimpse of how I write in the privacy of my notebooks when I'm communing with you. Here for example are some things I wrote in Sevilla, in 1999. (In those days, in my journals I wrote it in the form $\pi(n)=n/\mathrm{H}_{\pi(n)}$ and I called it FULANO'S Prime Number Theorem.)

I regard it as perhaps my prime responsibility as a teacher to wrap up a line of inquiry by "canonizing" what folks have been saying in small groups. To explain, for example "mathematicians have agreed to use this symbol

$$\sigma(n)=\Sigma_1^n \mathrm{d}(i)$$

for what you've been calling such and such …." But before I could start writing about Fulano's Theorem — before I could dare to claim that students on the very margins of academia have seen a vision of the primes distributing themselves harmoniously among the natural numbers — before I could begin writing this sentence, this account — before I could feel authorized to discourse on this august and deeply canonical topic, I had to read and think about the relevant branches of number theory. I was looking to see who else had said what we were saying, seen what we were seeing. Whose vision were we re-experiencing? For if and when I could find where it is (always already) written, in what Book it is already proven, that $\pi(n)=n/H_{\pi(n)}$*, then my task would be complete.*

But in the case of this theorem, the more you look, the less you find it. I've found many interesting discourses and have been able to memorize the canonical traditional master narrative which always pays hagiographic obeisance to the opus of various saints — Euclid, Eratosthenes, Euler, Legendre, Gauss — and always reaches a Wagnerian climax with Riemann in the Nineteenth Century. For nowadays, unless you really understand Riemann's Zeta Function what you say about the Prime Number Theorem must remain at best a romantic discourse, comparable to playing an outdated opening in championship chess.

The canonical opening move in playing the august and most ancient game of number theory is to create a function that counts primes. How many gods are there? Primes are god-like. To see how they interweave with their offspring is to know divinity.

The road not taken, FULANO's uncanonical opening move, is to establish between all primes and all natural numbers a relationship not dependent on counting, not instituted by putting primes and natural numbers in lexical order and establishing *a one to one relationship between them, for this creates an arbitrary relation between n and* p_n*, between, for example, 23 and the 23rd prime number. Linearization is always, even here, at the very threshold of mathematics, just another idealization of reality.*

FULANO's opening move is to see what happens when (as the great John Conway likes to phrase it) you count the features. Allot to each number n, not the n^{th} *prime, but the greatest prime factor of n itself. Do this with care and you get a complete account, a totally precise visualization of how primes behave, how dense they are, how they distribute themselves among their products.*

I studied as deeply as I could all that had been written about the prime number theorem by the most serious mathematicians until eventually I reached my limits. One always does. To reach your limits means discovering, once again, that you merely dabble in the shallow surf of a vast ocean. All mathematicians reach their limits. Even Andrew Wiles reaches his. And then you go out, deeper than you have ever been in your life, and start all over again, beyond your previous limits.

It takes nerve.

One might almost say that one who has not yet reached one's limits has still to become a true mathematician….

My limits were the discourse on Riemann's Hypothesis.

The main thing I found in the canonical discourse on the prime numbers was a deep contemplation of how log n — the natural logarithm of n — is intimately intertwined both with H_n *and via Li(n) with the canonical counting function* $\pi(n)$*. Indeed where Fulano's Theorem reads* $n = \pi H_n$*, the canonical reads* $n = \pi \log n$*.*

But the natural logarithm is simply a fancy version of FULANO's handmade "logarithm" H_n*. And Riemann's zeta function looks like a sort of "super-harmonic" series, just* H_n *writ large. "What Riemann's zeta function looks like when you let it take its Forbidden Argument, 1."*

Yet before all is done someone must explain how it can be that a group of ordinary people on the very margins of academia came to formulate a beautiful prime number theorem unknown to the legitimate mathematical world.

III How accurate is this formula?

Whenever you come up with a new prime number theorem, it's natural to ask how accurate it is. Canonically, if your new one is not as accurate as the best already known, it can't be worth much.

My hope is that you will spend many happy hours on Mathematica measuring the accuracy of FULANO's prime number theorem for yourself. For anyone taking early steps into Mathematica this is a very rewarding point of entry. At least, it was for me…. My breakthrough came when I said to myself, OK, we don't know the slope of the last line, so just call it x

$$n = x\, \mathrm{H}_x$$

Now get Mathematica to solve this for every x.

FULANO's theorem says that as a function of n, this should be about the same as $\pi(n)$.

Anyone can check it out for themselves. Once you figure out how to translate it into Mathematica..

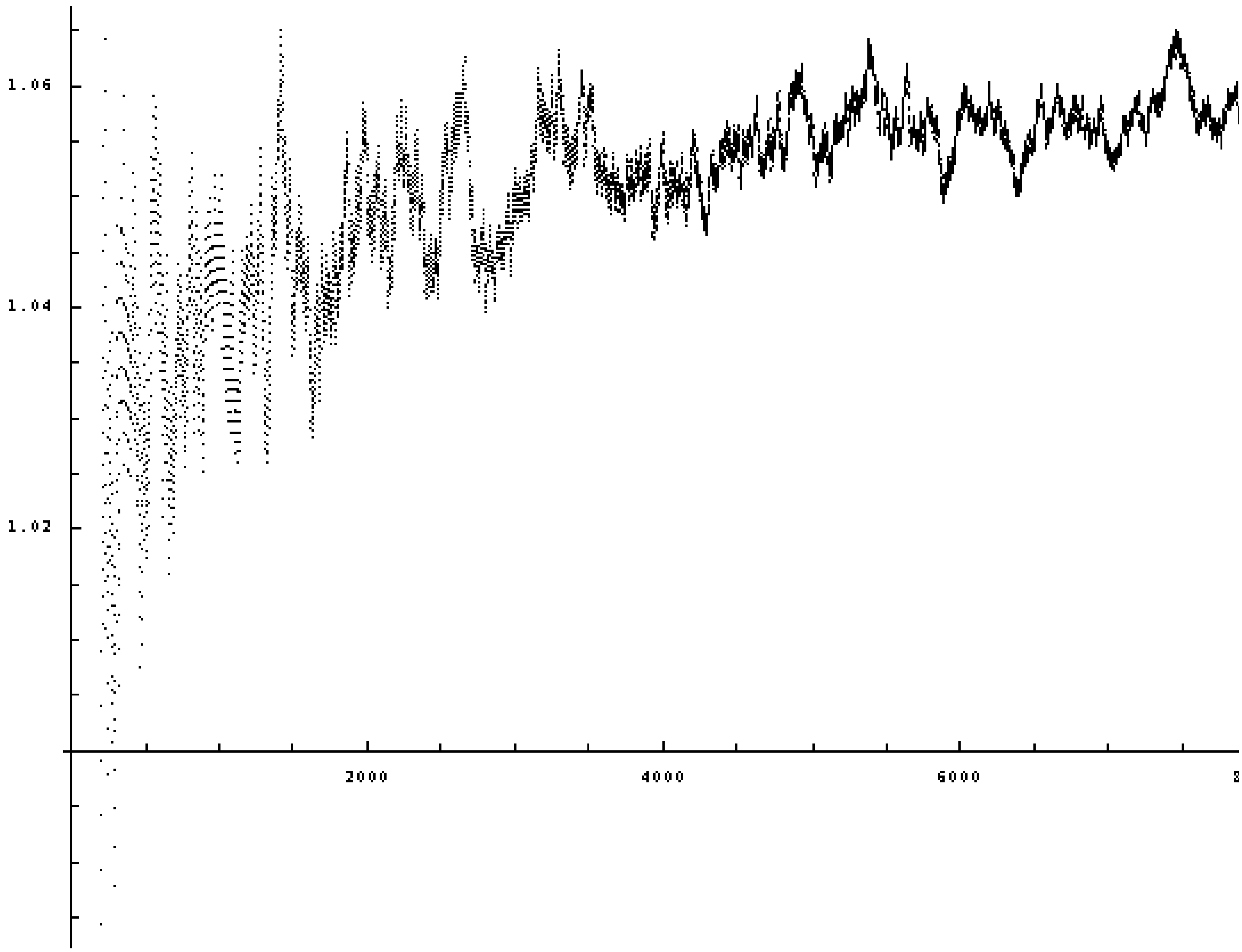
To solve $n = x\, \mathrm{H}_x$ for x write FindRoot[n == x HarmonicNumber[x]]. Define the answer as Fulano's $\pi(n)$:

```
fulanopi[n_] := Floor[x /. FindRoot[n == x h[x], {x, 4.5}]]
```

Compare this with the true value of $\pi(n)$ (which Mathematica can calculate up to some pretty big numbers). To explore the ratio between them define and plot:

```
ratio[n_] := fulanopi[n]/PrimePi[n]
```

Have fun! Here is a sample graph from my early explorations of this ratio.



For the first five hundred billion numbers the ratio between what Fulano predicts for π(*n*) and Mathematica's (true) value keeps up this tango teasingly close to 1. But what is five hundred billion? A mere half-trillion, just a clearing of the divine throat, that's all. Maybe not close enough for God. But close enough for Goddard.

IV Turning the harmonics inside out

In tranquillity now, let's rewind the Three Minute Version to the moment when we applauded the Heatherian Hypothesis: n points are distributed over π harmonics. Now at leisure we can fine-tune our visualization of what's really going on with these harmonics.

How many points can you actually count on each harmonic? What *is* the last harmonic on a finite graph of n points? What is its slope $1/h$?

At California State University Monterey Bay, some groups explored this carefully and discovered that the harmonics don't actually appear in the order 1, 2, 3., ….

h does eventually take on all values, but not in numerical order. Some you expected to find have not yet begun.

One group focussed on this question for quite a while. They listed the harmonics in the order in which they actually "sound":

1, 1, 2, 2, 1, 2, 1, 4, 3, 2, 1, 4, 1, 2, 3, 8, 1, 6, 1, 4, 3, 2, 1, …

18 points and still no 5th harmonic, let alone 7th. The seventh harmonic doesn't "sound" until $n = 49$, whereas the 8th has already sounded at $n = 16$. The 8th, the 16th , the $2^{n\text{-th}}$ harmonics are prominent peaks.

The 4, 3, 2, 1, pattern is striking. If you play this sequence on the piano, with 1 as the tonic, you hear a clear melodic pattern.

At Goddard this story energized Morgan to graph, by hand, the distribution of the harmonics themselves. We called this the dual of the basic graph. For any given n the dual of the original Guess my Rule gives you the harmonic on which n falls. For example:

You give me 5 I give you 1

You give me 7 I give you 1

You give me 20 I give you 4

What's my rule?

16 comes out 8

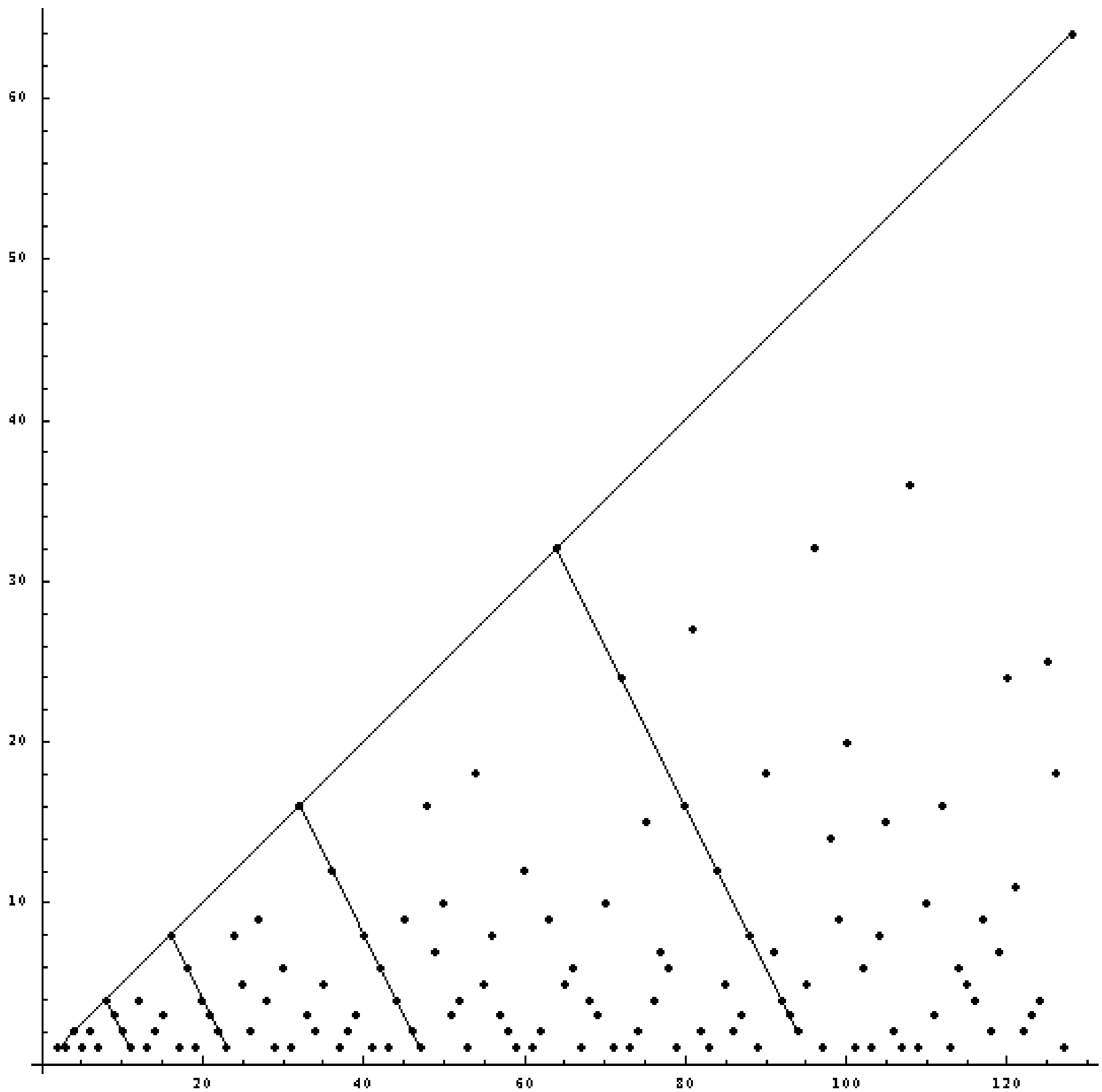
24 comes out 8

17 comes out 1

What's my rule? *h* is all the factors left when you divide *n* by its highest prime factor.

Graphing h as a function of n lets you look at the distribution of primes from a fresh viewpoint. Again, dear Reader go thou and do likewise: pause and draw Morgan's graph by hand before going on. If you don't have the epiphany you may not get the joke!

(To speed up the exposition of what comes next, I've highlighted two linear patterns in the following graph.)



This is the same pattern as before, but inverted in a helpful way. It is as if you have turned the original patterns inside out. The first sloping line is no longer a "harmonic". It has nothing but powers of 2 on it.

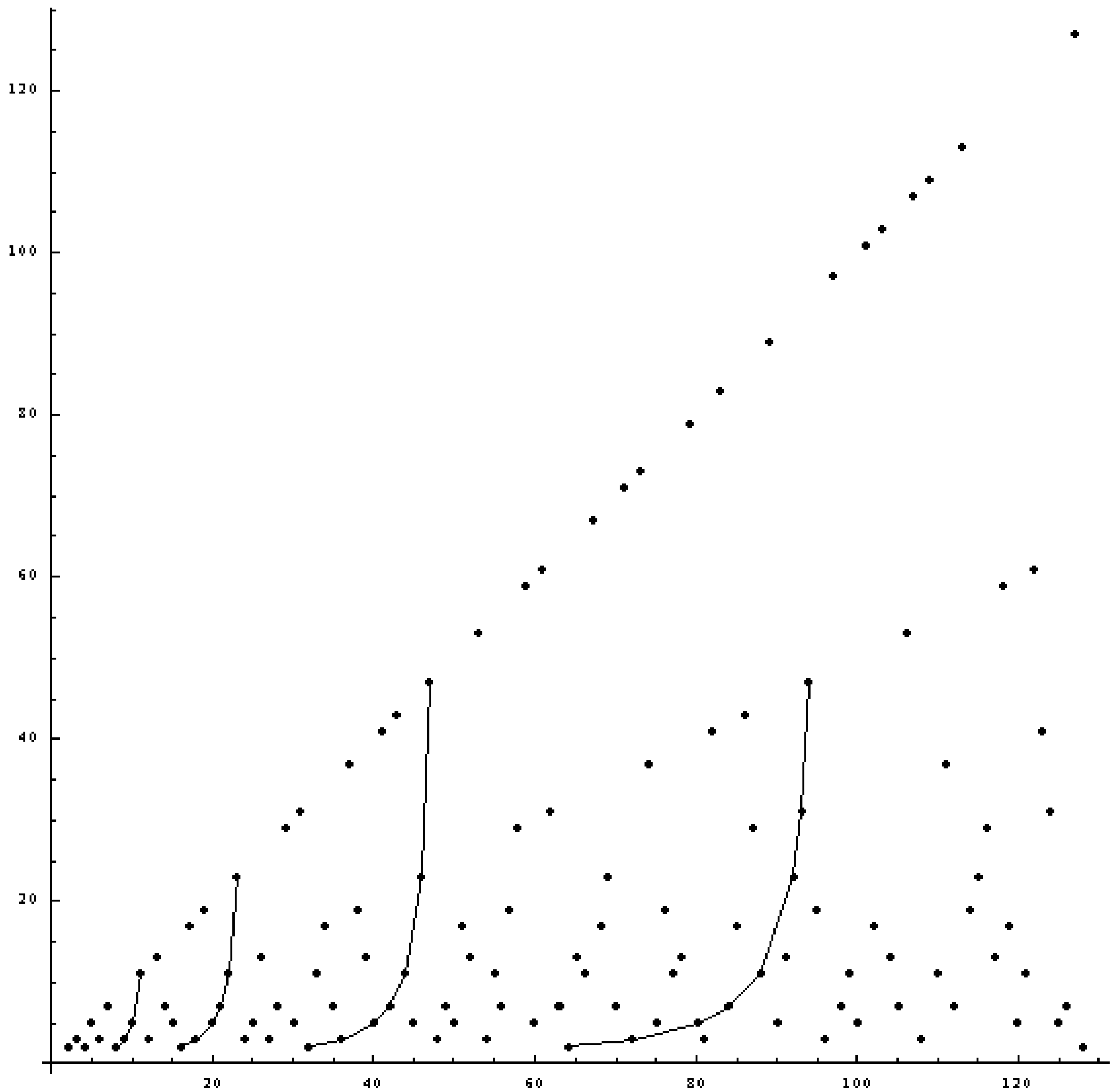
The slopes are no longer 1, 1/2, 1/3, are they? What's the new rule for slopes? (Note that on my graph the aspect ratio is *not* 1.) Does every power of 2 end up on a harmonic that is a power of 2?

Where's the first harmonic, the prime line? (It is now the first *horizontal* line, $h = 1$.) The harmonics are now stacked vertically on top of one another creating unexpected cross-rhythms.

And *vice versa:* the first sloping line on this new graph is the "dual" of the first horizontal line of powers of 2 on the old graph. And the second sloping line is the dual of the second horizontal line on the old graph (powers of 3 times powers of 2). And so on.

Looking at these graphs today, the original graph and its dual, I am reminded of a manipulative talked about by a group at CSUMB. They were convinced it would help 5th and 6th graders experience the harmonic distribution of primes. They proposed making a fan of thin wooden strips, hinged at the origin, laid out (on the original graph) at the correct slopes for the harmonics. On each harmonic, wherever there is actually a point, screw in an eyelet. String a nylon thread through the eyelets on the line $y = p$ and you have the horizontal lines. Today I imagine replacing the horizontal nylon threads by thin dowels (or stiff wires.) Unhinge the harmonics from the origin, lay them on top of one another, and put the dowels through the now-sloping lines of eyelets. Hinge the dowels at the origin. You have this dual, don't you? Or do you?

The 4, 3, 2, 1 pattern we noticed earlier is now clearly visible as a new series of lines sloping *down* from powers of 2 to the prime line. Squint at it from a certain attractive angle of vision (the origin!) and this 2^n "grid" can start to look like its organizing the whole shebang, somehow. These downward lines start at powers of 2 and end on or very near primes. Check it out. They link each power of 2 to a certain prime. With care you can plot the duals (on the original graph) of these straight lines. They look like this:



Epiphany! Those are the curves I alluded to so breezily at the beginning of this paper!

These curves have a hyperbolic sort of shape. A hammock of hyperbola-like curves carrying powers of 2 to certain subsequent primes.

The further out you go the harder it gets to *see* the "hammock" (the hyperbola/line duals). But you can *train* your eye to see a little further each day and to believe that beyond what your eye can discern, this pattern, so clear at its birth, does continue out forever and ever, and that the "hyperbolic" curves are there, though no longer "visible to the naked eye."

If you succeed in training yourself to believe this, find where, in the canonical literature on prime number theory, belief in this "hammock" of connections between powers of 2 and certain primes has already been discussed, explained, and legitimized.

What if not?

V The Joys of Rigor

Go back to the grid of strips and dowels and nylon threads. What we're really talking about here is a lattice. Playing around with it some more can help us fine tune the Goddard guess that there are "about" π/h points on the h^{th} harmonic. Because if you look closely at, say, the 3rd harmonic, you see it's missing it's first point. (Why?) How many points are missing on the 5th harmonic? On the 7th , 11th, 13th harmonics?

Epiphany! If h is itself prime, it's missing the first $\pi(h)$ points!

What about when h is a power of 2? When h is a product of two primes?

"By inspection" you can see that there is 1 point missing at the beginning of lines with slope $1/h$, h=3, 6, 9, 12, 18;

2 points missing on h = 5, 10, 15, 20;

3 on h = 7, 14…

What's the pattern?

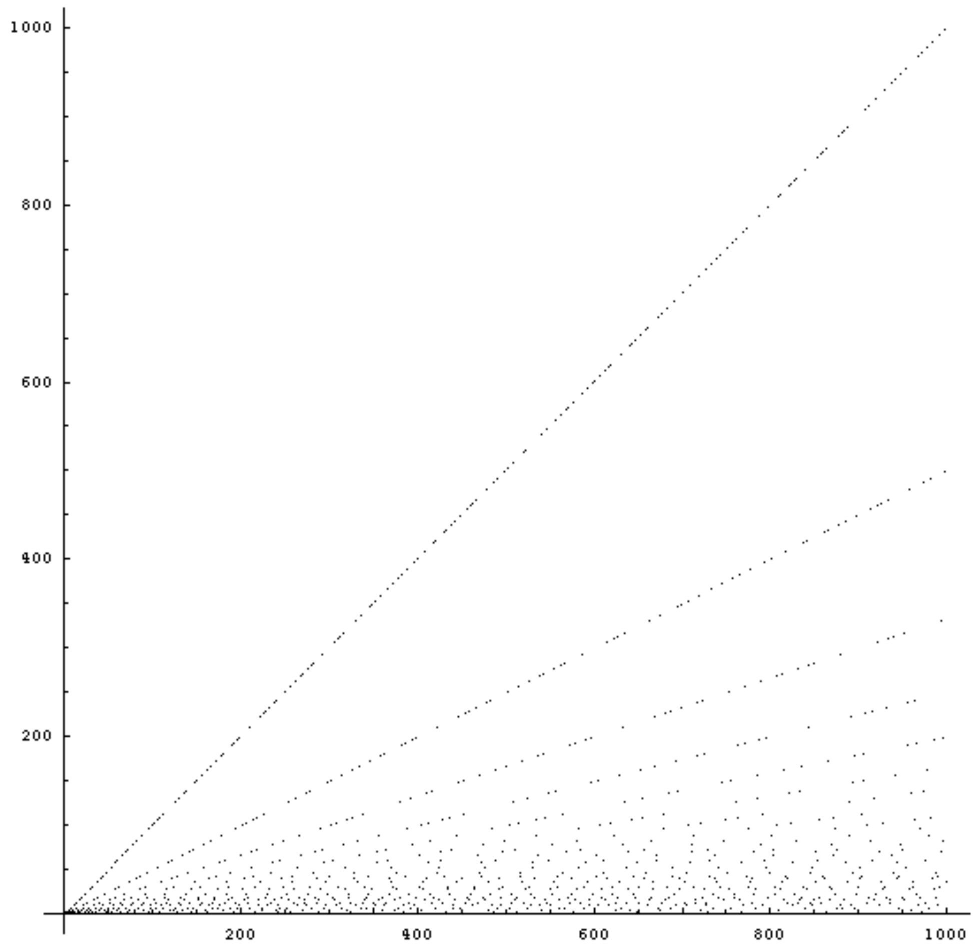
To improve the visualiztion symbolized by $n = \pi H_\pi$ subtract the pattern of these missing points. Their number is determined by the highest prime factor of h. So we can write

$$n/\pi = \sum(\pi/i - \pi\ (P^+(i))) \quad (i \le \pi)$$

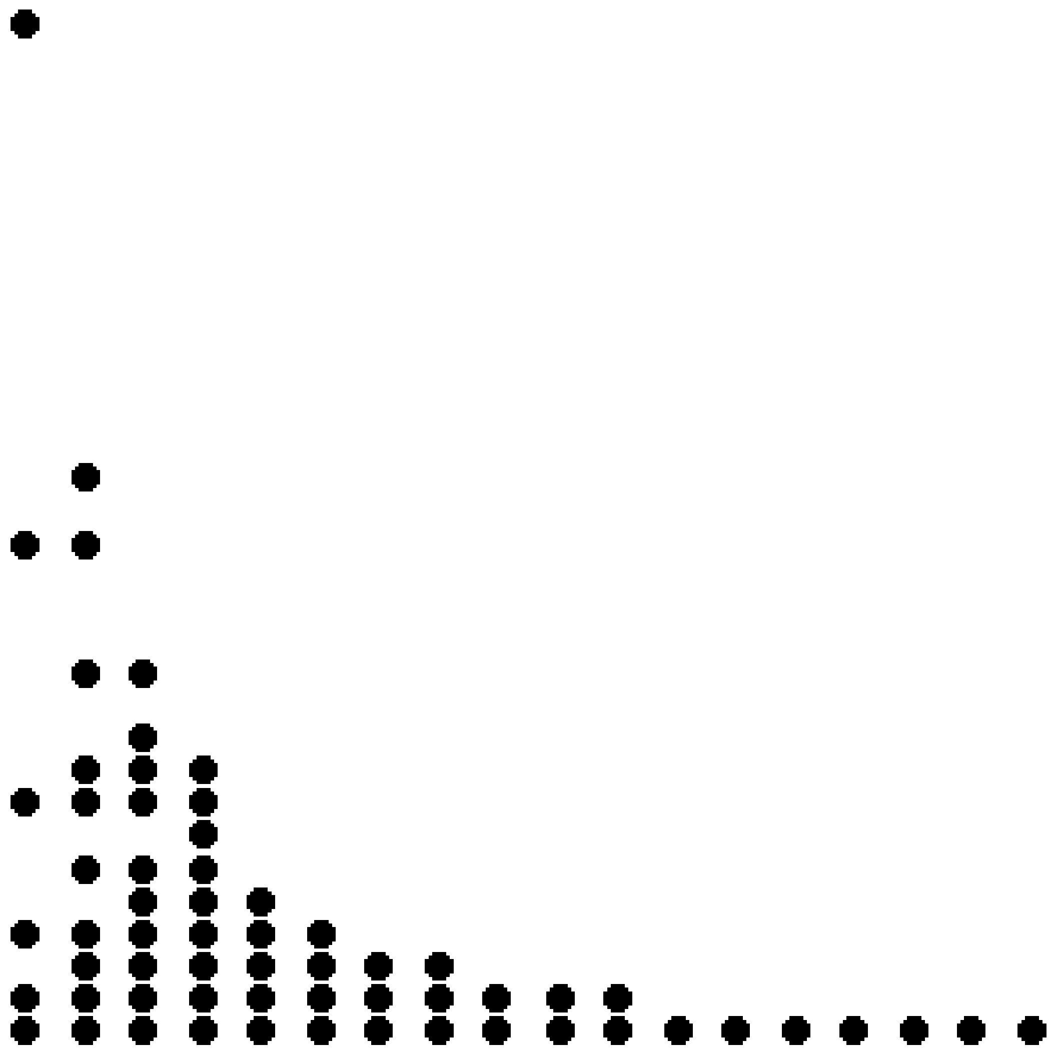
On the other hand if $P^+(h) > \sqrt{n}$, the graph contains no points on the harmonic with slope $1/h$ (it begins "later"). Is this covered by the adjustment? or are we sometimes "overcompensating?

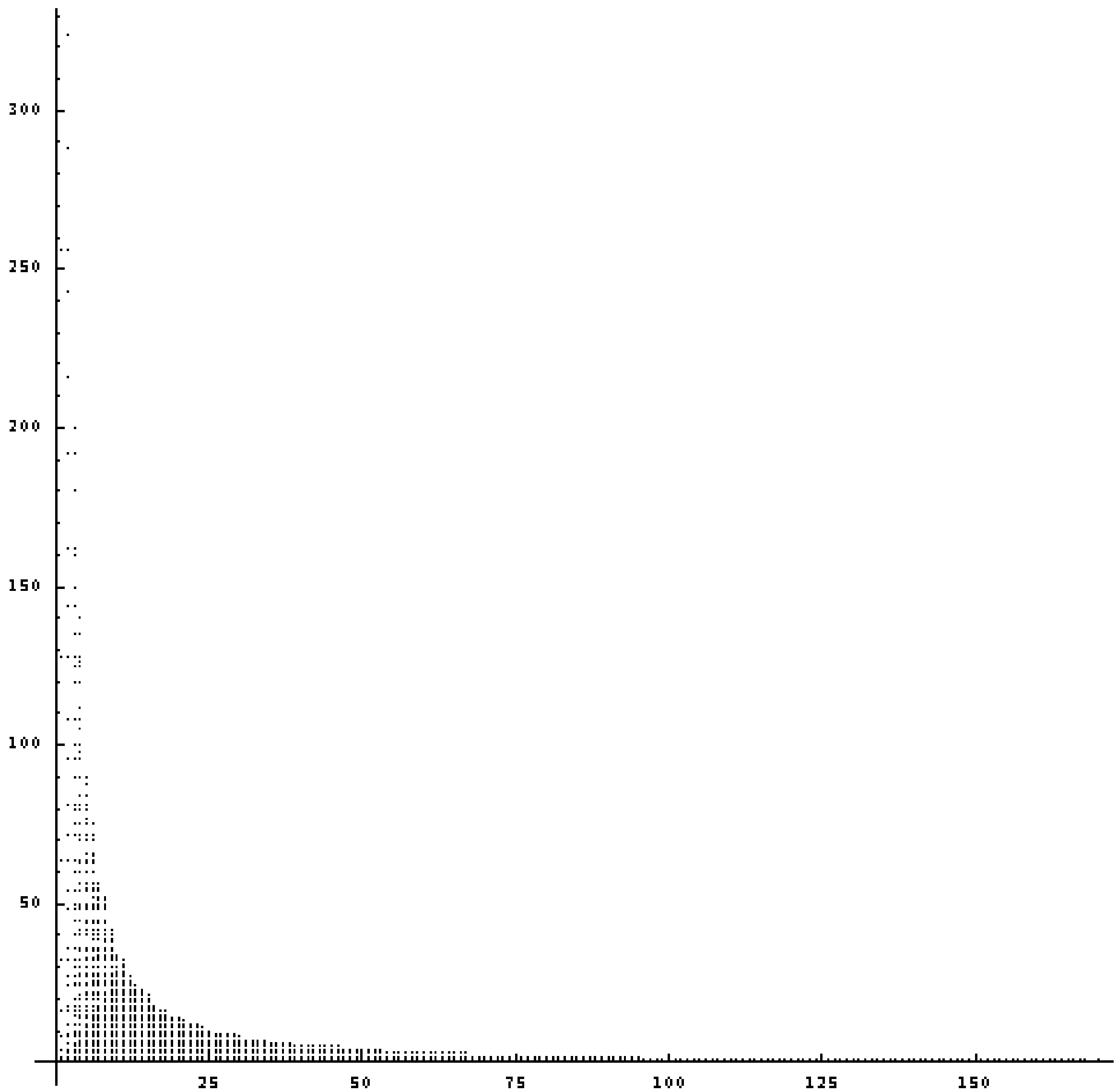
This thickens the plot no little and quite some. πH_π counts more points than there actually are, so we have to subtract an error term based on π of the highest prime factor of h. ("Dude, this is *fractal*! Totally!")

It would be nice to be able to *visualize* this error term. But on the basic graph, the higher the harmonic, the flatter the slope, the harder it is to *see* precisely how many points are missing on each one. Here's a thousand points: how many are missing on, say, the 10^{th} harmonic?



But what if we graph *h* against *p*, against the highest prime factor of *n*? What if we look not at the "dual" space but the "control" space, the "phase" space? Write each *n* as an ordered pair, $\{p,\ h\}$ where p is $\pi(\mathrm{P}^{+}(n))$, and h is the harmonic on which it "sounds", *i.e.* $h = n/\mathrm{P}^{+}(n)$. Here's what you get for the first 64 points,





and for the first thousand points:

The missing points (and missing harmonics) up the left hand side are now easy to see. Are they something to worry about? Are they negligible? Or as *n* gets bigger and bigger do they grow and grow and chew up the pretty picture (and the pretty formula $n = \pi H_\pi$)? Can the appearances be saved?

That juicy question I leave as an exercise for the interested reader.

Once you are hooked, further exercises naturally arise,. For example, since there are no half points on a harmonic, shouldn't we really have writte $n = \pi + [\pi/2] + [\pi/3] + \ldots + [\pi/\pi]$ which is a little less than $\pi + \pi/2 + \pi/3 + \ldots + \pi/\pi$?

This is an interesting function in its own right, canonically connected with Dirichlet's divisor problem. Following Dirichlet, Hardy and Wright[1] reverse our steps. Counting the lattice points under a hyperbola (just like our last graph above, but *without* the missing points), they show that if $d(j)$ counts the divisors of *j,* then

$$d(1) + d(2) + \ldots + d(\pi) = \pi + [\pi/2] + [\pi/3] + \ldots + [\pi/\pi]$$

(To avoid abuse of notation I stick with my π.) This they approximate this by our πH_π. (Which they do *not* connect to the distribution of primes.) This in turn they approximate by $\pi \log \pi$. They thus end with Dirichlet's canonical theorem, invoking Euler's gamma:

$$d(1) + d(2) + \ldots + d(\pi) = \pi \log \pi + (2\gamma - 1)\pi + O(\sqrt{\pi}).$$

Sprinkle n points on the sloping lines the way we did to begin with, add up the points on those lines and you get back to n. That much is true, isn't it? Sooner or later, however, we have to face the fact that the number of primes less than half of n is not exactly half the number of primes less than n.. So we really should have written

$$n = \pi(n) + \pi(n/2) + \pi(n/3) + \ldots + \pi(n/?)$$

Minus the error term that we began looking at above.

And that is the punchline of an entirely different joke.

Or is it?

If we pin down all these issues, won't we have an exact formula for $\pi(n)$*?*

An exact visualization of how harmoniously they are distributed? Such that you and I can sit in an armchair, close our eyes, and see it all, steady and whole?

You, dear Steve, along with the friendly readers over our shoulders, get to decide.

---

[1] *An Introduction to the Theory of Numbers,* Section 18.2, "The average order of *d*(*n*)," pp. 263-266 in the fifth edition, Oxford University Press, 1985.