Appendix 9

A proof of Riemann's hypothesis via Denjoy's equivalent theorem

Description of the proof

The Riemann Hypothesis is true if and only if the numbers of positive and negative signs of $\mu(n)$ are asymptotically equal, and I prove it by showing that every other disposition of the signs is impossible.

Prolegomena

On 2006 06 27 I announced on the net a theorem very much stronger than Riemann's. Write $R_2(n)$ for $\mathrm{li}\,n - 1/2\,\mathrm{li}\,n^{1/2}$, and let $x^{1/2}$ denote the positive square root of x. Then for all $n>1$

(1) $R_2(n) - 1/2(R_2(n))^{1/2} < \pi(n) < R_2(n) + 1/2(R_2(n))^{1/2}$.

My theorem in (1) is about as much stronger than the RH as the RH is stronger than the PNT. In working up a more-general account of this theorem, with rigorous proofs of its validity, for eventual publication in print, I discovered a neat proof, this time of Riemann's hypothesis only, on entirely different lines.

For this second proof we refer to what is called Denjoy's probabilistic interpretation, notably that the RH is equivalent to the proposition that any square-free number, taken at random, has an equal probability of containing an odd or an even number of distinct prime divisors.

Legendre's formula for $\pi(n)$

Legendre's formula (*Essai* 2nd edition, Paris 1808, pp 412sq) is a recipe for calculating the exact number $\pi(n)$ of primes $\leq n$ without identifying them all. It can be written

(2) $\pi(n) = \pi(n^{1/2}) + (\Sigma\, \mu(d)\ [n/d]) - 1$

where μ is the Möbius function, and the denominators (d) are all the natural numbers that have no large prime p in their decomposition. A prime p is 'large' (in relation to n) if $p > n^{1/2}$.

Analysis. The formula in (2) works correctly because its summation term yields the number of numbers $\leq n$ that are not struck out by the Eratosthenes procedure of striking out those of them that are divisible by a prime q that is 'small' in relation to n, i.e. is such that $2 \leq q \leq n^{1/2}$. The procedure will obviously still work if we redefine one or more of the large primes as 'small'. The unstruck numbers include 1, which is not nowadays classed as a prime. Students of arithmetic born before 1900 were taught that 1 is the least prime, making Goldbach’s conjecture apply to all even numbers including 2. Present-day arithmeticians find it more convenient to exclude 1 from the class of prime numbers, making it the unique natural number whose number ofdistinct prime divisors is zero.

The function $\mu(d)$ can now be defined as equal to 1 if the number of distinct prime divisors of d is even, –1 if it is odd, and 0 if d has a repeated divisor (other than 1). Since 1 is not struck out by the sieve of Erastosthenes and is also included in the count of “primes” calculated by the section $\Sigma\mu(d)\ [n/d]$[*], the count must be reduced by one in either case, and then to get the complete answer the number of small primes (q) used as strikers must be added to the total.

Illustration of Legendre's formula with $n = 20$

d	$f(d) = \mu(d)\ [n/d]$
1*	+20
2*	–10
3*	–6
5	–4
6*	+3
7	–2
10	+2
11	–1
13	–1
14	+1
15	+1
17	–1
19	–1
	Σ +1

$\Sigma\mu(d^*)\ [n/d^*] = 7$

$7 - 1 + \pi(n^{1/2}) = 8 = \pi(20)$

The (d*) are the denominators with no large prime in their decomposition. The small primes 2, 3 must be known explicitly, then the number of large primes 5, 7, 11, 13, 17, 19 can be calculated without any of them being identified.

In the table above we see an illustration of the use of Legendre's formula to calculate $\pi(n)$ for $n = 20$. The starred terms, with no large prime divisors of d, are used to calculate the number of large primes $\leq n$. Notice I have used all the (d) that yield an $f(d)$ other than zero, and the sum of these, for any n, must always be 1, since only one number, 1 itself, remains unstruck if we use all the primes.

The stage is now set for my proof of Denjoy’s equivalent to Riemann’s hypothesis.

First we get rid of the 1, which is the only number left standing after my extension of Legendre’s procedure. To do this we remove it at the beginning, quite legitimately,

---

[*] The fact that, with d unrestricted, the formula $\Sigma\mu(d)\ [n/d] = 1$ is true for all n was first noted by Meissel in *Observationem quaedam in theoria numerorum*, Berlin 1850, but he failed to discover my easy proof of it (use Hardy and Wright T 260) or to find a use for it.

because it is neither prime nor composite, and so does not belong to either of the two complementary classes, composites and primes, to which we reduce the number system in accordance with modern practice.

We thus further rectify the procedure by making the following change:

use $f(d) = \mu(d)\ [(n-1)/d]$ for $d=1$

and use $f(d) = \mu(d)\ [n/d]$ for all other values of d.

Now rework $n=20$ using the new procedure

d	$f(d)$		
1	+19		
2	–10		
3	–6		upper section
5	–4		
6	+3		
7	–2	sum to half way	
10	+2	+2	
11	–1		
13	–1		
14	+1		lower section
15	+1		
17	–1	sum in 2nd half of n	
19	–1	–2	
	0	sum complete.	

I have divided the terms into two sections, in the upper of which each $f(d)$ consists of $\mu(d)$ multiplied by some positive number >1, and in the lower each $f(d)$ is simply $\mu(d)$. (We should note that every value of $\mu(d)$ other than the first must appear in the lower section of one or more natural n.) In the upper section the $f(d)$ sum need not be exactly the sum of the $\mu(d)$ pluses and minuses to this point, but it is obviously positively correlated to it. (For example if all the terms were negative the answer would be negative, and vice versa.) In the lower section the sum of the terms is exactly the sum of the $\mu(d)$ in the section.

Recall that, by Denjoy's equivalent, the Riemann hypothesis is true if and only if the algebraic sums of the pluses and minuses of $\mu(d)$, taken progressively at unit increments of d, vary asymptotically around zero*. Suppose it is untrue. This can only mean that the sums must vary asymptotically around some number other than zero**.

---

* This means we can get it as close as we like to an average of zero difference between the two.

** This means we cannot get the average difference as close as we like to zero after n has reached a certain size, but can get it as close as we like to some other number.

Suppose this number is such that the upper sections of all numbers, taken progressively, vary in aggregate around +2, which we recognize is the sum of the upper-section terms for $n = 20$, so we may take this n as a typical example. Now the sum of the lower-section terms for this n must be exactly –2 to compensate the upper section. So double *n to* $2n = 40$. The aggregate of pluses and minuses in the upper section of $2n = 40$ is exactly what it was in the whole of $n = 20$. But it contains two more minus signs than did the upper section of $n = 20$, so its sum is likely to be reduced towards or beyond zero. Suppose by an unlikely chance it is still +2. The lower section of $2n = 40$ must again be –2 to compensate this, so repeat the procedure by doubling $2n$ to $4n = 80$. Now the upper section of this new number $4n = 80$ must contain two more extra minus signs, making it even more likely to be reduced towards or beyond zero.*

These unlikely chances cannot continue for ever, because every time we doubled the argument we would have to add an average of two more minus signs to the upper-section terms of the new doubled argument, so there must come a time when the sum of the upper-section terms of the new doubled argument is reduced to or beyond zero. Suppose it is reduced to zero. Then the sum of the lower-section terms for this argument will also be zero, and there will be no tendency in either direction when it is doubled again. But suppose the sum in the upper section is reduced beyond zero to a negative value. Now the lower-section sum for this argument must be positive, and the whole process must play itself out again, this time in the opposite direction.

Recall, finally, that the lower sections of every argument consist of increasingly protracted sets of consecutive terms of $\mu(d)$, and that all values of $\mu(d)$ except the first must be presented in the lower sections of one or more natural n.

Because of the negative feedback between the two sections, to suppose the average difference between the plus and minus signs of $\mu(d)$ were to differ from zero by any quantity, however small, is unsustainable, because if this were so, then the absolute difference between the signs would continue to increase without limit in the same direction until there would be a large excess of numbers of a particular prime parity in one half of the numbers up to a given n, and a large deficiency in the other. But since the contents of the two halves are exchanged as n grows larger, this state of affairs is impossible to maintain. Therefore the average difference between the signs can be asymptotic only to zero and the Riemann hypothesis must therefore be true.

---

* We must distinguish, in the upper sections, the aggregate of the plus and minus *signs* from the aggregate of the $f(d)$ *terms*, which in the upper sections are not necessarily the same. In the lower sections they must be the same, and this is what allows the proof. For example, when $n = 20$ the aggregate of the upper-section *terms* is +2, but the aggregate of the upper-section *signs* is –1, i.e. there is a total of one more minus signs than plus signs. In the upper section for $n = 40$ there are three more minus signs than plus signs, i.e. two more than before, as predicted.

Aftermath

The process of negative feedback in successive Möbius values is continuous, and begins long before the big jumps I used to illustrate it. Those of my readers who remember their course in radio telephony will recall that negative feedback leads to oscillation (called 'hunting') about a mean value. The negative feedback in the case I considered is quite pronounced, with an average gain of 1.7857 in the early stages, as the following table will confirm.

n	upper-section deviation	added	swing in direction of addition
5	+2	–	–
10	–1	–2	–3
20	+2	+1	+3
40	–3	–2	–5
80	+4	+3	+7
160	–2	–4	–6
320	–1	+2	+1
		absolute sum 14	absolute sum 25

average gain = 1.785714...

From the table we can see that the successive Möbius values are not at all random, as many commentators have mistakenly supposed, but follow what is called by unsophisticated gamblers a "maturity of chances" hypothesis. This is a supposition that, for example, after a run of successive black numbers at the roulette table, the probability of a red number appearing next is increased to "redress the balance". But the (to some extent) empirically verified hypothesis of probability determines axiomatically that the result of the next spin will be independent of previous results, so the player loses from the inclusion of a zero number that renders the probability of red or black slightly less than 1/2. But if the casino were naïve enough to offer evens against a + or – appearing in any continous set of consecutive values of $\mu(n)$, the player could win a fortune by betting against the trend.

Offhand I can think of no series of naturally-produced numbers, other than $\mu(n)$, for which a maturity-of-chances hypothesis happens to be true.* Can any of my readers?

An unintended confirmation of this is provided by John Derbyshire in his book *Prime obsession* (New York 2003), which is the best and most complete account of the Riemann hypothesis I have seen. It contains, moreover, fewer serious mistakes than any other account I have read, though it is of course impossible to write a book of this size (422

---

* Of course it is also true of primes and composites, but these categories are not independent of $\mu(n)$.

pages) without including some mistakes. He makes the common mistake of suggesting that a series of consecutive values of $\mu(n)$ might be random in respect of their + and – signs (p 322). In a previous page (250) he quotes sets of Mertens’s function (cumulative $\mu(n)$) that he says tell us very little except that their absolute value increases as n does. In fact they tell us a great deal, and if he had noticed it he might, admittedly with a fair amount of further detective work, have discovered my beautiful proof of Riemann’s hypothesis several years before I did.

On page 322 he correctly points out that the average difference between n randomly-produced +1’s and –1’s is $\sqrt{n}$. To convert Mertens’s function into a corresponding function for square-free (n) we must multiply each n by $6/\pi^2$, or about 0.608. From his second set of figures we find arguments

1000 2000 3000 4000 5000 6000 7000 8000 9000 10000

conversions

608 1216 1824 2432 3040 3648 4256 4864 5472 6080

roots of conversions

25 35 43 49 55 60 65 70 74 78

Mertens’s values for original arguments

2 5 –6 –9 2 0 –25 –1 1 –23.

It is evident that the final set of values (considered absolutely as differences) are ridiculously below what they should be if the original sets of +1’s and –1’s were randomly produced. I was going to do a table of his third set of arguments, in millions, whose values are equally impressive, but the set I have tabulated all denote unrandomness so obviously that I will leave the tabulation of his third set to the reader, for the good feeling of being part of the research. What they show is that the successive Möbius +1’s and –1’s are unrandom to an enormous degree, being hugely biased towards a maturity-of-chances hypothesis.

My strong theorem at the beginning of this memoir shows that the primes are similarly unrandom, in the sense of being much more evenly-spread than if they had been randomly placed in the sequence under review.

I had experimented with Legendre’s method of counting primes for many years, convinced that it could lead to a very elementary proof of the PNT, but ironically could not see how to do it until faced with the more heroic prospect of proving the Riemann hypothesis. In this case it seems to be impossible to prove the one without simultaneously proving the other.

In common with other experts in the field, I had begun to suspect that the RH is a problem in elementary arithmetic and not in analysis, as many of us, probably including

Riemann, had previously thought.

Telephone +44 1985 844 855

Abbreviations

PNT = prime number theorem, notably that $\pi(n)/(n/\log n) \to 1$ as $n \to \infty$

RH = Riemann('s) hypothesis that

$\zeta(s) = \Sigma n^{-s} = \Pi(1-p^{-s})^{-1}$

in which n runs through the natural numbers 1, 2, 3, ... and p through the primes 2, 3, 5, ... cannot be equal to zero for nonreal s other than of the form $s = 1/2 + iy$ with $i = \sqrt{-1}$ and y real.

This proof by Professor George Spencer-Brown of Georg Friedrich Bernhard Riemann's hitherto unproven hypothesis, originally proposed in 1859, was first published on the internet in 2008 04 24, and in hard copy in this book in 2008 09 15.

An analysis of the proof

My Denjoy proof of Riemann's hypothesis looks so simple that there seems to be little we can say about it except to reenact it. But one or two things are worth remarking.

Consider the natural counting numbers from 1 up to n and decompose each of them into its prime components. Add up all the exponents of the prime components of each of them. A well-known equivalent to Riemann's hypothesis is the proposition that, for any one of these numbers selected at random, the sum of the exponents of its prime components is equally likely to be even or odd.*

What Professor Denjoy discovered is that we need not consider any of the numbers $\leq n$ with a prime exponent of two or more. In short, if all square-free numbers (d), say, could be shown to be equally likely to have an even or an odd number of prime divisors, the truth of the Riemann hypothesis would follow. Hence all that is required is to show that the average of the differences between the plus and minus terms of $\mu(d)$ varies around and is asymptotic to zero as n increases without limit.

We should next note that we can reclassify one or more of the large primes $\leq n$ as "small" without affecting the result. Suppose for $n=20$ we use, as we must, the small primes 2 and 3, and add to them the extra primes 7 and 13. For the relevant $f(d) \neq 0$ we now have

d	$f(d)$
1	+19
2	−10
3	−6
6	+3
7	−2
13	−1
14	+1
	Σ +4

$\Sigma f(d)=4$. Add to this the redefined "small" primes 2,3,7,13 gives $4 + 4 = 8 = \pi(20)$.

It is further evident that the use of the extra "small" primes as strikers in the sieve of Eratosthenes will make no difference to the result, apart from the fact that they will strike out themselves, so provided we add them back to the total the result will be the same as that of the Legendre/Spencer-Brown procedure, say LSB.

Now notice what I did to prove the Riemann hypothesis. By using all the square-free $(d) \leq n$ I in effect reclassified *all* the primes $\leq n$ as "small". This ensures that whenever I use the rectified Legendre procedure to count the remaining large primes $\leq n$, I shall get the answer zero. I thus made the LSB procedure useless for counting primes, but

---

* Cf Borwein and others, *The Riemann hypothesis*, Burnaby 2008

extremely useful for proving the Riemann hypothesis, since with an $f(d)$ total of zero for every n, I can split the series of $f(d)$ terms anywhere I choose into two sections, and whatever the total $+t$ (say) in one section will be balanced against a total of $-t$ in the other.

The most instructive place to split the series is at the point where $(n/d)<2$. We then get a lower section of $f(d)$ terms all of value ±1. They comprise all the values of $\mu(d)$ for square-free (d) from this point up to n. The upper section of the series will now consist of the early terms of $\mu(d)$ for square-free (d) magnified by a factor ranging from 2 up to $n - 1$.

The fact that the plus and minus values of $\mu(d)$ are magnified in the upper sections ensures that any excesses of positive or negative terms of $\mu(d)$ that appear in the lower sections become overcorrected in the upper sections, to which they are transferred as n increases. Effectively what happens is that both sections correct each other towards an asymptotic average of zero.

The Riemann hypothesis, as we have noted, must be true if the pluses and minuses of $\mu(d)$ are merely equiprobable. But the fact that any excess of one sign over the other that begins to appear in a lower section gets magnified when this part of the lower section gets incorporated in an upper, ensures that the difference between them stays closer to zero than it would if the successive plus or minus signs of $\mu(d)$ were merely randomly distributed like successive falls of a coin.

Consider the last two terms of my rectified $(f(d))$ for $n=20$, notably with $d=17$ and $d=19$. Both of these terms must, with complete certainty, be -1. In a random sequence of terms, the value of each term within the range considered must be completely unaffected by the values of previous terms. But in this case, as we see, the values of these terms are entirely determined, and therefore completely predictable, by the values of the previous terms. This means that whatever patterns the values of $(\mu(d))$ can display, they cannot be random.

In particular it means that both of the arguments 17, 19 for $n=20$ can have only an odd number of prime divisors, and since the cube root of 20 is less than three and both of these numbers are odd, they must both be prime. Thus in general, once we have ascertained the number of prime divisors of $d=1$ and found it to be the even number zero, making $\mu(1) = +1$, all subsequent values of $\mu(d)$ for square-free d can be found by a simple elementary algorithm without having to know any prime divisor of any $d > 1$. *

The astronomer August Ferdinand Möbius was born in Schulpforta in 1790 11 17, and his

---

* The fact that the prime parity of 17 is odd, for example, completely determines the fact that the prime parities of 237 618 987 and 1 009 003 027 are both even. Although this is unavoidable when we think about it, the way we present it to ourselves seems to invest it with an aura of incredibility. Certainly this astonishing fact was neither known nor suspected before this publication.

famous number-theoretic function predates Riemann's paper by at least two decades, and the use of it by others, e.g. Euler, by much longer. That the mathematical profession has to this day failed to notice the most significant property of this function, and has compounded the failure by wrongly supposing its terms to be in some respect "random", defies belief.*

So we can discover the parities of the numbers of primes in all square-free numbers simply by considering the parities of the numbers of primes in lesser numbers hitherto thought to be unrelated to them, something that Hardy and Wright supposed could not be done.

I remember in the 1950's Lord Cherwell, then a senior colleague of mine at Christ Church, and I used to write to the surviving author inclosing various formulas to predict prime numbers larger than the last one determined, suggesting he incorporate them in the next edition of the book. He always refused to do so, we thought wrongly.

None of the suggestions was as ingenious as the one I have just proved, notably that the successive nonzero values of $\mu(d)$ behave like the falls of a magic coin that, from the moment it is struck, remembers exactly how many times it has fallen with one side or the other uppermost, and whenever one side exceeds the other, biases itself towards the other until the excess is eliminated.

This self-correcting property is more typical of living organisms, and thus surprising when we find it in what we thought of as an inanimate number system. It of course makes no difference to the truth of the Riemann hypothesis, which would still be true if the successive values of $\mu(d)$ for square-free (d) merely behaved like the falls of an ordinary unbiased coin, instead of like a being that watched what it was doing and modified its behaviour accordingly. But the fact that the differences between the plus and minus values for square-free (d), because of this self-correcting tendency, stay *closer* to zero than would the differences between the two sides of successive falls of an unbiased coin, suggests that the Riemann hypothesis might be in some way *more than* true.

This could be interpreted as saying that Riemann's 1859 proposition might be too weak, and that my stronger propositions below might more nearly represent the true state of affairs.

By Professor Denjoy's equivalent theorem, the Riemann hypothesis is seen to be true if and only if the successive +1's and –1's in $\mu(d)$ are equiprobable. Since it is meaningless to select a term at random from an infinite set, this can only mean that

---

* This terrible mistake stems from the prevailing myth that the primes are somehow "randomly" distributed, if not wholly so, than at least partially. On the contrary, the primes are an example of the most beautiful and paradoxical form of order imaginable, a perfectly ordered series, i.e. completely unrandom, that never repeats itself. Every point in it signposts the way to every other point, and no two points are confused. And as with this, so with August Möbius's beautiful function that perfectly reflects it.

their differences over the number of LSB terms displayed, which we may call their average differences, tend towards zero as n increases without limit. It does not necessarily mean that the nonzero terms have to be randomly distributed, like the falls of an ordinary unbiased coin, and evidently, as I have proved, they are not.*

This indicates that not only the Riemann hypothesis itself was unclear in the minds of its previous investigators, but that other things about it were unclear too, since if this associated lemma had been clarified, it is hardly likely that such an obvious clue** would not have led to a speedy proof of the hypothesis. But this was not the case, and I cannot recall any mathematical problem I have solved where the muddle in the minds of most of its previous investigators was more extensive or complete.***

For example I have seen no previous account of the Riemann hypothesis and its environs that does not make the mistake of suggesting that the nonzero terms of $\mu(n)$ must be randomly distributed, with probabilities of a half each, for the RH to be true. All that is required is that they be equiprobable. That they be equiprobable *and* random is clearly not required, and is equally clearly not true. By proving them to be equiprobable I proved the RH, but by simultaneously proving them to be non-random in the particular way they are, I proved something that suggests Riemann's original guess was not strong enough, and that the real theorems associated with this branch of arithmetic are in fact much more constraining than he supposed. This turns out to be the case, as we shall presently see.

It also highlights another example of the muddled thinking of previous investigators. Although not explicitly stated, it is nevertheless implied in all the accounts I have seen, that Riemann's guess must be the holy grail of all numeric theorems, and therefore that it must impose the narrowest possible limits on the range of the prime count. In fact these limits are much narrower than Riemann's guess requires them to be.

What mathematicians tend to do when they cannot prove or disprove a proposition is to invent equivalent statements that they think might be easier to decide. Denjoy's probabilistic interpretation of Riemann's guess, that I found easier to prove, is a case in point.

---

* Much of the confusion here springs from the fact that the Möbius +1's and –1's comprise neither proper numbers nor proper signs, but are merely convenient ways of saying whether the prime parity of a number n is even or odd. Boole made a similar mistake, using numbers and signs to represent truth values, which I corrected in the calculus of indications by eliminating both. We could do the same here, substituting 'even' for +1 and 'odd' for –1, but even this is too specific, since in a calculus of only two values all we need to do is to distinguish them without saying which is which. Thus we can generalize $\mu(n)$ to $\sigma(n)$, say, with a single ambiguity that we can resolve at any point, conveniently at $n=1$. Then if $\sigma(1)=+1$, we have $\mu(n)$, and if $\sigma(1)=-1$ we have $-\mu(n)$. In either case the values denote prime parities for which we need not factorize n.

** In fact there was another obvious clue in the surprisingly small sizes of the values of Mertens's function, but it too was overlooked.

*** Of course I do not discount explorers such as von Koch, Denjoy, and Littlewood, without whose preliminary findings my task would have been much harder.

Another such interpretation, by von Koch, concerns constraints on the errors of the prime count from some proven asymptote to it, such as $n/\log n$ or li n. von Koch proved in 1901 that Riemann's conjecture is equivalent to the errors of the prime count $\pi(n)$ being constrained, for all n above some large-enough value, to within the range of li n $\pm K\sqrt{n}\log n$ with K constant.

Improving on von Koch's (and therefore Riemann's) extremely gross limits to the prime count, Littlewood, in 1907, correctly concluded that 'if the P.N.T. were true 'with error about $\sqrt{n}$', the R.H. would follow.' (*Miscellany**, Cambridge 1953.) He means here by P.N.T. the proposition that $\pi(n)/\mathrm{li}\,n \to 1$ as $n \to \infty$, rather than the more-usual $\pi(n)/(n/\log n) \to 1$. Both propositions have been proven true, but the former converges faster than the latter.

Littlewood made no sustained attempt to investigate his or von Koch's not-too-difficult restatement of Riemann's hypothesis, either theoretically by trying to prove it, or empirically by comparing it with known prime-counts. Even more astonishingly, for the next 99 years, nobody else looked for an empirical verification of von Koch's equivalent until I announced on the net, in 2006, the stronger theorem that

(3) $\mathrm{li}n - (\mathrm{li}n)^{1/2} < \pi(n) < \mathrm{li}n + (\mathrm{li}n)^{1/2}$

is true for all known prime counts when $n > 1$ and $x^{1/2}$ is the positive square root of x. (Or we can be more sophisticated and apply the negative square root to the left-hand side of the inequality and the positive square root to the right.)

This is increasingly stronger than Littlewood's lemma, and therefore stronger than Riemann's hypothesis, because it substitutes $(\mathrm{li}n)^{1/2}$ for Littlewoods's $n^{1/2}$, and in the range considered, li$n$ is greater than 1 and lesser than n.

The reason I decided to publish first my Denjoy proof of Riemann's much weaker theorem (formerly known as his 'hypothesis'), is because it requires no principle or axiom that was not available to Euclid, and therefore makes no demands on the reader other than those that have been traditional and available in school text books for the last two thousand years.

My proofs that do make further demands, are of much stronger theorems than Riemann's. Calling

1. $m(n) = \mathrm{li}n - 1/2(\mathrm{li}n^{1/2})$, and
2. $m'(n) = S_d(n)$ with $d=0$ (see p 221),

and calling $r(n) = m(n)$ or $m'(n)$, I can prove that

(4) & (5) $r(n) - 1/2(r(n))^{1/2} < \pi(n) < r(n) + 1/2(r(n))^{1/2}$

assuming, as usual, that $x^{1/2}$ denotes the positive square root of x.

---

* For consistency I exchanged n for x in the text.

These I call my strong theorems because, as is evident from a cursory glance, they are very much stronger than anything Riemann conjectured, and also stronger than my theorem in (3).

I will not prove them here, because there is no need to burden the reader with additional new ideas until a later date, when he, or she, will have become familiar with at least one proof of Riemann's hypothesis and is ready for more.*

My Denjoy proof, in summary, runs as follows.

1. What Professor Denjoy showed is that the RH is equivalent to the proposition that the number of primes in a square-free $d \leq n$ of any size is even or odd with equal probability.

2. I rectify Legendre's method of counting large primes in n and then corrupt it to give the answer zero for all (n).

3. I split the rectified Legendre terms into two sections, upper and lower, so that the lower sections eventually include all values of $\mu(d)$ for square-free $(d) > 1$.

4. I show that the average algebraic sum, i.e. the sum divided by the number of LSB terms displayed, in each section varies around and is asymptotic to zero as n for the $f(d)$ terms increases without limit.

5. Since the lower sections eventually include the values of $\mu(d)$ for all square-free (d) > 1, and their signs are by the previous proposition equiprobable in the limit, the RH, quod erat demonstrandum, must be true.

6. In addition, since the upper sections eventually contain all the values of $\mu(d)$ for square-free (d), but magnified by various factors ranging from 2 up to $n-1$, that are independent of the signs of $\mu(d)$, and the average differences between the plus and minus values of these magnified terms also tend to zero as n increases, this fact constitutes a second proof of Riemann's hypothesis, since if an average of a set of magnified differences tends to zero, then the average of the same set of differences unmagnified must also tend to zero.

It should be noted that in proving the RH in this elementary way, I have also made a simple elementary proof, much simpler than Selberg's, of the prime number theorem, since the one implies the other. I have furthermore decided other propositions, such as a version of Goldbach's conjecture, that Riemann's hypothesis implies. I can also show that my strong theorems imply the truth of all forms of Goldbach's conjecture, and of other previously undecided propositions about prime numbers.

---

* It will also give us something further to discuss when I am invited to talk about my proofs to academic audiences.

Spencer-Brown's cascade

We can redescribe the Legendre/Spencer-Brown procedure LSB to be

(6) $\mathrm{LSB}(n) = n - 1 + \Sigma\mu(d)[n/d]$ with various restrictions on d.

(6.1) Restricting the (d) to primes not greater than $\lfloor n^{1/2} \rfloor$ and their mutual multiples yields the number of large primes not greater than n.

(6.2) Derestricting the (d) to all integers > 1 makes the $\mathrm{LSB}(n)$ equal to zero for all (n).

The fact that (6.2) = 0 for all (n) allows us to demonstrate one of the most astonishing facts of arithmetic, notably that we can know how any natural number will factorize without factorizing it, and by an algorithm so simple that a child of six can do it.

It is done by what I call a cascading algorithm. This is a spinoff from the procedure I adopted in my Denjoy proof.

From our knowledge that $\mu(1) = +1$ because 1 has zero prime divisors and zero is an even number, we proceed to discover $\mu(n)$ for every subsequent number n without having to factorize any such n, as follows.

Take $n = 2$. The LSB terms for this n will be

d	$n-1$ $+f(d)$	running total
	+1	+1
2	−1	0

Since the running total, by (6.2), must reach zero for every n, the last value of $f(d)$ must in this case be −1 to reach this total, so $\mu(2) = -1$ and 2 therefore has an odd number of prime divisors.

Take $n = 3$. Now the LSB series will be

d	$n-1$ $+f(d)$	running total
	+2	+2
2	−1	+1
3	−1	0

so $\mu(3) = -1$ and so 3 has an odd number of prime divisors.

Take $n = 4$

d	$n-1$ $+f(d)$	running total
	+3	+3
2	−2	+1
3	−1	0
4	0	0

since the penultimate total is already zero, $\mu(4) = 0$ and so 4 has a square divisor. Therefore we can ignore 4 in subsequent cascades.

Take $n = 5$

d	$n-1$ $+f(d)$	running total
	+4	+4
2	−2	+2
3	−1	+1
5	−1	0

so $\mu(5) = -1$ and so 5 has an odd number of prime divisors.

Take $n = 6$

d	$n-1$ $+f(d)$	running total
	+5	+5
2	−3	+2
3	−2	0
5	−1	−1
6	+1	0

so $\mu(6) = +1$ and so 6 has an even number of prime divisors.

Take $n = 7$

d	$n-1$ $+f(d)$	running total
	+6	+6
2	−3	+3
3	−2	+1
5	−1	0
6	+1	+1
7	−1	0

so $\mu(7) = -1$ and so 7 has an odd number of prime divisors.

Take $n = 8$

d	$n-1$ $+f(d)$	running total
	+7	+7
2	−4	+3
3	−2	+1
5	−1	0
6	+1	+1
7	−1	0
8	0	0

so $\mu(8) = 0$ and so 8 has a square divisor and can be ignored in subsequent cascades.

Take $n = 9$

d	$n - 1$ $+ f(d)$	running total
	+8	+8
2	–4	+4
3	–3	+1
5	–1	0
6	+1	+1
7	–1	0
9	0	0

so μ(9) = 0 and so 9 has a square divisor and can be ignored in subsequent cascades.

Take $n = 10$

d	$n - 1$ $+ f(d)$	running total
	+9	+9
2	–5	+4
3	–3	+1
5	–2	–1
6	+1	0
7	–1	–1
10	+1	0

so μ(10) =+1 and so 10 has an even number of prime divisors.

Notice we are entirely unconcerned whether the d-arguments are prime or not, or whether or not the divisions n/d are exact. The cascade does not require this information. All it requires is what the previous cascades have told it. We also see that there is no need to list the final term for any n, since the answer must be the penultimate term with the sign reversed.

It is easy to seee how our six-year-old will continue to find the values of μ(n) for every subsequent value of n without ever having to factorize, or even to find one divisor of, any n whatever.

Whenever I make a discovery like this that the best of the rest appear to have overlooked for the past eight thousand years or so, I find it difficult not to suspect that somebody else might have thought of it before. I usually consult an "authority" to make sure. One such "authority" is H M Edwards, *Riemann's Zeta Function*, New York 1974, a detailed account in 336 pages of how not to prove the Riemann hypothesis.

In p 268 he remarks that it is 'plausible to say that successive evaluations of μ(n) are "independent" since knowing the value of μ(n) for one n would not seem to give any information about its value for other values of n.'

This tells us that my law of succession of ($\mu(n)$), how to find its next value from its previous values, was not known before I announced it. Unfortunately nearly all such useful "authorities" are secondary sources, they copy uncritically what other authors have written, including mistakes, and do no experiments of their own that might have led to an independent observation. Consequently they are useful only for verifying historical facts, since whenever they venture an opinion it is never based on arithmetical evidence.

Mr Edwards's opinion above is both wrong and nonsensical, wrong because it is untrue (he suggests there is no law of succession for $\mu(n)$, and I have just demonstrated a very simple one), and nonsensical because it is arithmetically evident that we must always be able to decide what will come next in a calculus whose elements are 1, 2, 3, 4, ..., and in any definite function of these elements, so it must be possible, at least in principle, to find the next value of $\mu(n)$ from its previous values, and what I did was discover a way of doing so that is simple enough for a child of six to operate.

It is nice because it gives us a way to find the value of $\mu(n)$ for any n merely from its previous values, instead of from factorizing n and checking to see which of its factors is prime and then checking again to see if any of them is repeated, and if not counting the prime divisors to see if their cardinal number is even or odd.

It is important because it shows that apparently complicated properties of n, whether it has a repeated divisor or if not, an odd or an even number of prime divisors, previously thought to be determinable only by factorizing n and counting its various kinds of factors, are actually determined by its ordinal place in the system, and its divisors must for this reason fall into the appropriate category without having to be counted or even observed at all.

All this seems to be astonishing only because arithmetical text books hitherto have been propagating a myth, notably that there exist two different entities, the numbers themselves and their ordinal places in the system.

According to this myth we think we have to analyse the number itself to discover its properties, decompose it into primes, look to see if it is square or triangular, does it have a repeated divisor, etc etc etc. Thinking about it this way, as an isolated phenomenon, we are naturally surprised to see that many of its properties can be ascertained simply from its place in the system.

In fact *all* of its properties must be evident from its place in the system, because what we were taught to think of as two different entities, the number itself and its place in the system, are one and the same.

A number *is* its place in the system, no more and no less: so there are not two different entities, numbers and the places where they live. There are just places, the places are the numbers and the numbers are the places.

G Spencer-Brown
England 2009 04 17 the 40th anniversary of the publication of this book.