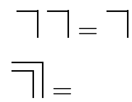

Laws Of Form and the Riemann Hypothesis

J. M. Flagg

(George Spencer-Brown)

Louis H. Kauffman

Department of Mathematics, Statistics, and Computer Science,
851 South Morgan Street,
University of Illinois at Chicago,
Chicago, IL 60607-7045
and
Department of Mechanics and Mathematics,
Novosibirsk State University
Novosibirsk,
Russia
kauffman@uic.edu

Divyamaan Sahoo

Sound Department,
School of the Art Institute of Chicago,
112 S Michigan Ave, Suite 512,
Chicago, IL 60603
dsahoo@artic.edu

RH

The Riemann Hypothesis (RH) is a conjecture first made by Bernhard Riemann in *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, "On the Number of Primes Less Than a Given Magnitude", 1859 [**42**]. The conjecture is about the "zeros" of the zeta function, ζ, whose domain is the complex numbers, $s \in \mathbb{C}$:

$$\zeta(s) = \sum_{1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + ...,$$

where $a, b \in \mathbb{R}$, and $i = \sqrt{-1}$.

By convention, the trivial zeros of ζ are the integers $-2, -4, -6, -8, ...$. These are understood to be zeros via analytic continuation[1]. The remaining zeros of ζ are called the non-trivial zeros. Riemann conjectured that the non-trivial zeros of zeta are always of the form $s = \frac{1}{2} + bi$, i.e., that the "real part" a of the non-trivial zero s, is $\frac{1}{2}$.

[1] In Section 1.4, "The Function $\zeta(s)$", *Riemann's Zeta Function* [**11**], H. M. Edwards writes: "The view of analytic continuation in terms of chains of disks and power series convergent in each disk descends from Weierstrass and is quite antithetical to Riemann's basic philosophy that analytic functions should be dealt with *globally*, not locally in terms of power series". Edwards notes that "Riemann does not speak of the "analytic continuation" of the function Σn^{-s} beyond the halfplane Re $s > 1$, but speaks rather of finding a formula for which it "remains valid for all of s". In *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse* [**42**], Riemann indeed writes: "The function of the complex variable s which is represented by these two expressions (*the two halves of the Euler product identity*), wherever they converge, I denote by $\zeta(s)$. Both expressions converge only when the real part of s is greater than 1; at the same time an expression for the function can easily be found which always remains valid" (transl. Wilkins, 1998 [**42**]).

Euler first studied the zeta function in the domain of the real numbers[2], $\mathbb{R}$. In *Variae observationes circa series infinitas* (1737) [**16**], he derived a fundamental relationship between the zeta function and the prime numbers:

$$\zeta(s) = \sum_n \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + ... = \frac{1}{(1-\frac{1}{2^s})(1-\frac{1}{3^s})(1-\frac{1}{5^s})...} = \prod_p \frac{1}{1-p^{-s}}.$$

Riemann extends ζ to the complex domain by using the factorial function, notated $\Pi(s)$ due to Gauss in *Disquisitiones generales circa seriem infinitam* (1813) [**18**]. Originally attributed to Euler in *De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt*(1730) [**17**][3], the factorial function for $n \in \mathbb{N}$,

$$n! = \int_0^\infty e^{-x} x^n dx$$

can be extended to the real numbers > -1. Due to Gauss [**18**] we have

$$\Pi(s) = \int_0^\infty e^{-x} x^s dx, \qquad (s > -1).$$

Abel's *Solution de quelques problemes a l'aide d'integrales définies* (1823) [**1**] and Chebyshev's *Sur la fonction qui détermine la totalité des nombres premiers inferieurs à une limite donnée* (1852) [**7**] give us the following identity for real values of s, $s > 1$:

$$\zeta(s) = \sum_1^\infty \frac{1}{n^s} = \frac{\Pi(-s)}{2\pi i} \int_{+\infty}^{+\infty} \frac{(-x)^s}{e^x - 1} . \frac{dx}{x}.$$

[2]In Section 1.2 "The Euler Product", *Riemann's zeta function* [**11**], H. M. Edwards writes, "it seems certain that Riemann's use of the Euler product formula was influenced by Dirichlet" and "Dirichlet, unlike Euler, used the formula (1) with s as a real variable and, also unlike Euler, he proved rigorously that (1) is true for all real $s > 1$" (Page 7, Edwards). Edwards provides Dirichlet's remarks in a footnote: "Since the terms p^{-s} are all positive, there is nothing subtle or difficult about this proof – it is essentially a reordering of absolutely convergent series – but it has the important effect of transforming (1) from a formal identity true for various values of s to an analytical formula true for all real $s > 1$" (Dirichlet, page 7, footnote by Edwards). We know, of course, that Riemann went even further, broadening the scope of ζ to include all complex s, $s \neq 1$. Here, formula (1) is the Euler product formula provided above.

[3]Euler wrote it as follows: $y = e^{-x}$ and $n! = \int_0^1 (log\frac{1}{y})^n dy$

The contour integral in $\frac{\Pi(-s)}{2\pi i}\int_{+\infty}^{+\infty}\frac{(-x)^s}{e^x-1}.\frac{dx}{x}$ clearly converges for all values of s, real or complex, since e^x grows faster than x^s as $x \to \infty$. Since $\Pi(s)$ is an analytic function of the complex variable s, which has simple poles at $s = -1, -2, -3, ...$, $\Pi(-s)$ has poles at $s = 1, 2, 3, ...$. We know $\zeta(s)$ converges for $s = 2, 3, 4, ...$ (à la Euler et al), so by Cauchy's theorem, the integral must have a zero which cancels the pole of $\Pi(-s)$ at $s = 2, 3, 4, ...$. Thus, $\zeta(s) = \sum_1^\infty \frac{1}{n^s} = \frac{\Pi(-s)}{2\pi i}\int_{+\infty}^{+\infty}\frac{(-x)^s}{e^x-1}.\frac{dx}{x}$ is analytic at all points of the complex s-plane except for a simple pole at $s = 1$ (Edwards, Section 1.4, "The Function $\zeta(s)$" [**11**]).

On page 2-3 of *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse* (transl. Wilkins, 1998) [**42**], Riemann continues, "thus a relation between $\zeta(s)$ and $\zeta(1-s)$, which through the use of known properties of the function Π, may be expressed as follows" and that "remains unchanged when s is replaced by $1-s$":

$$\Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}}\zeta(s) = \Pi\left(\frac{1-s}{2}-1\right)\pi^{-\frac{1-s}{2}}\zeta(1-s).$$

Riemann considers this symmetrical statement as the natural statement of the functional equation. The function $\Pi(\frac{s}{2}-1)\pi^{-\frac{s}{2}}\zeta(s)$ in symmetrical form, has poles at $s = 0$ and $s = 1$. Riemann then introduces ξ, an analytic function defined for all values of s, by multiplying the simpler left hand side of the functional equation of ζ by $\frac{s(s-1)}{2}$ to get[4]

$$\xi(s) = \Pi\left(\frac{s}{2}\right)(s-1)\pi^{-\frac{s}{2}}\zeta(s)$$

Given a convergent infinite product can only be zero if one of its factors is zero, $\zeta(s) = \Pi_p(\frac{1}{1-p^{-s}})^{-1}$ cannot have zeros for Re $s > 1$. Since the factors other than $\zeta(s)$ in $\xi(s) = \prod(\frac{s}{2})(s-1)\pi^{-\frac{s}{2}}\zeta(s)$ only have the simple zero at $s = 1$, we know that none of the roots ρ of $\xi(\rho) = 0$ lie in the Re $s > 1$. Due to the functional equation, we know ρ is a root *iff* $1-\rho$ is a root, so we also know that none of the roots ρ of $\xi(\rho) = 0$ lie in the half-plane Re $s < 0$. Hence, all the roots ρ of $\xi(\rho) = 0$ lie in the strip $0 \leq$ Re $\rho \leq 1$ (Edwards, Section 1.9 "The Roots ρ of ξ" [**11**]).

[4] Adopting Landau's notational convention for ξ, Edwards remarks "Actually Riemann uses the letter ξ to denote the function which it is now customary to denote by Ξ, namely, the function $\Xi(t) = \xi(\frac{1}{2}+it)$, where ξ is defined as above." (Edwards, footnote on page 16 [**11**])

The following product representation of $\xi(s)$ was proved by Hadamard in *Étude sur les Propriétés des Fonctions Entières et en Particulier d'une Fonction Considérée par Riemann* (1893) [**21**][5]:

$$\xi(s) = \xi(0) \prod_{\rho} \left(1 - \frac{s}{\rho}\right),$$

where ρ "ranges over" the roots of the equation $\xi(\rho) = 0$ and multiple roots are to be counted with multiplicities. Recall that a polynomial $p(s)$ can be expanded as a finite product:

$$p(s) = p(0) \prod_{\rho} \left[1 - \frac{s}{\rho}\right],$$

where ρ ranges over the roots of the equation $p(\rho) = 0$. Thus, $\xi(s)$ is "*like a polynomial of infinite degree*[6]" (H. M. Edwards, Section 1.10, "The Product Representation of $\xi(s)$" [**11**]).

[5] Since the function $\log \xi(s)$ has logarithmic singularities at the roots ρ of $\xi(s)$ and no other singularities, it has the same singularities as the formal sum $\sum_\rho \log(1 - \frac{s}{\rho})$. If this sum converges and is well behaved with respect to $\log \xi(s)$ near ∞, then $\sum_\rho \log(1 - \frac{s}{\rho})$ differs from $\log \xi(s)$ by at most an additive constant. Setting $s = 0$ gives $\log \xi(0)$, which upon exponentiating gives us the identity (H. M. Edwards, Section 1.10 "The Product Representation of $\xi(s)$" [**11**]).

For any fixed s, the ambiguity in the imaginary part of $\log\left[1 - \frac{s}{\rho}\right]$ disappears for large ρ. Hence, the sum $\sum_p \log\left(1 - \frac{s}{\rho}\right)$ is defined except for a finite multiple of $2\pi i$ which drops out when one exponentiates the ξ identity above. Furthermore, "one can ignore the imaginary parts altogether", the real parts of the terms of $\sum_p \log\left(1 - \frac{s}{\rho}\right)$ "are unambiguously defined and their sum is a harmonic function which differs from $\operatorname{Re} \log \xi(s)$ by a harmonic function without singularities, and if this difference function can be shown to be constant, it will follow that its harmonic conjugate is constant also" (Edwards, Section 1.10 "The Product Representation of $\xi(s)$" [**11**]).

There are two problems associated with $\sum_\rho \log(1 - \frac{s}{\rho})$, the first being the "determination of the imaginary parts of the logarithms it contains", and the second associated with its convergence, for the sum is "in fact a conditionally convergent sum, and the *order* of the series must be specified in order for the sum to be well determined", i.e., it suffices merely to stipulate that each term be paired with its "twin" $\rho \leftrightarrow 1 - \rho$. (Edwards, Section 1.10 "The Product Representation of $\xi(s)$" [**11**])

[6] The identity $\xi(s) = \sum_{n=0}^{\infty} a_{2n} (s - \frac{1}{2})^{2n}$ also suggests that $\xi(s)$ is a polynomial of infinite degree, and a finite number of terms of this series representation of $\xi(s)$, presented in Riemann's second proof of the functional equation, gives a very good approximation in any finite part of the plane.

The J function[7] can be introduced via Stieltjes integrals. Using the identity $\log(1-x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 ...$, the Euler product formula can be reformulated:

$$\log\zeta(s) = \sum_p \left[\sum_n \left(\frac{1}{n}\right)p^{-ns}\right] = \sum_p \sum_n \left(\frac{1}{n}\right)p^{-ns} = \int_0^\infty x^{-s}dJ(x)$$

"where $J(x)$ is the function which begins at 0 for $x = 0$ and increases by a jump of 1 at primes p, by a jump of $\frac{1}{2}$ at prime squares p^2, by a jump of $\frac{1}{3}$ at prime cubes, etc" (Edwards, Section 1.11, "The Connection Between $\zeta(s)$ And Primes" [**11**]). Assuming the notation π (attributed to Landau, *Handbuch*, 1909 [**28**]) for the number of primes less than a given magnitude, we have the following two identities:

$$J(x) = \pi(x) + \frac{1}{2}\pi(x^{\frac{1}{2}}) + \frac{1}{3}\pi(x^{\frac{1}{3}}) + ... + \frac{1}{n}\pi(x^{\frac{1}{n}}) + ...$$

$$\pi(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} J(x^{\frac{1}{n}}) = J(x) - \frac{1}{2}J(x^{\frac{1}{2}}) - \frac{1}{3}J(x^{\frac{1}{3}}) - \frac{1}{5}J(x^{\frac{1}{5}}) + \frac{1}{6}J(x^{\frac{1}{6}}) - ...$$

It is of special interest to note that $\dfrac{1}{\zeta(s)} = \prod_p \left(1 - \dfrac{1}{p^s}\right)$.

In the Euler product of $(1 - \frac{1}{p^s})$, we see the general term is (-1) raised to the number of factors, a product of distinct prime factors. This leads to the definition $\mu(n) = (-1)^k$, where k is the number of distinct prime factors of n, and n has no higher order of factors, with $\mu(n) = 0$ otherwise[8].

Thus, we have $\dfrac{1}{\zeta(s)} = \displaystyle\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$.

---

[7] "Riemann denoted this function $f(x)$, and most other writers denote it $\Pi(x)$. Since $f(x)$ now is commonly used to denote a generic function, I have taken the liberty of introducing a new notation $J(x)$ for this function" (Edwards, Section 1.11, footnote, page 22 [**11**]).

"Let $F(x)$ be equal to this number when x is not exactly equal to a prime number; but let it be greater by $\frac{1}{2}$ when x is a prime number, so that, for any x at which there is a jump in the value in $F(x)$, $F(x) = \frac{F(x+0)+F(x-0)}{2}$. If in the identity $\log\zeta(s) = -\sum\log(1-p^{-s}) = \sum p^{-s} + \frac{1}{2}p^{-2s} + \frac{1}{3}p^{-3s} + ...$ one now replaces

$$p^{-s} \text{ by } s\int_p^\infty x^{-s-1}ds, \quad p^{-2s} \text{ by } s\int_{p^2}^\infty x^{-s-1}ds, \quad ... \quad , \quad \text{one obtains}$$

$$\frac{\log\zeta(s)}{s} = \int_1^\infty f(x)x^{-s-1}dx, \text{ if one denotes } F(x) + \frac{1}{2}F(x^{\frac{1}{2}}) + \frac{1}{3}F(x^{\frac{1}{3}}) + ... \text{ by } f(x).$$

This equation is valid for each complex value $a + bi$ of s for which $a > 1$" (page 5-6, *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, Riemann, transl. Wilkins, 1998 [**42**]).

[8] The classical Möbius function found in *Uber ein besondere Art von Umkehrung der Reihen* (1832) [**35**] is defined as follows: $g(m) = \sum_{n|d} f(n)$ and $f(m) = \sum_{n|d} g(n)\mu(\frac{m}{n})$

In his second proof of the functional equation, Riemann performs the change of variable $x = n^2\pi x$ in Euler's integral for $\Pi(\frac{s}{2}-1)$. For Re $s>1$, he arrives at
$\Pi(\frac{s}{2}-1)\pi^{-\frac{s}{2}}\frac{1}{n^s} = \int_0^\infty e^{-n^2\pi x}x^{\frac{s}{2}}\frac{dx}{x}$. Summation over n (for Re $s>1$) gives:

$$\Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}}\zeta(s) = \int_0^\infty \psi(x)x^{\frac{s}{2}}\frac{dx}{x}$$

where $\psi(x) = \sum_{n=1}^\infty e^{-n^2\pi x}$. In order to prove that the function above is unchanged by the substitution $s = 1-s$, Riemann uses the form of the functional equation of Jacobi's theta function, referring to Section 65 of Jacobi's *Fundamenta Nova Theoriae Functionum Ellipticarum*, in particular "Suites des notices sur les fonctions elliptiques" (1828) [**27**]. The following form is attributed by Jacobi to Poisson:

$$\frac{1+2\psi(x)}{1+2\psi(\frac{1}{x})} = \frac{1}{\sqrt{x}}$$

We see that $\psi(x)$ approaches zero very rapidly as $x\to\infty$. This shows in particular that $\psi(x)$ is like $\frac{1}{2}(x^{-\frac{1}{2}}-1)$ for x near zero. Hence $\int_0^\infty \psi(x)x^{\frac{s}{2}}\frac{dx}{x}$ is convergent for $s>1$. Thus[9], for $s>1$:

$$\xi(s) = \Pi\left(\frac{s}{2}-1\right)\pi^{-\frac{s}{2}}\zeta(s) = \int_1^\infty \psi(x)[x^{\frac{s}{2}}+x^{\frac{1-s}{2}}]\frac{dx}{x} - \frac{1}{s(s-1)}.$$

Since $\psi(x)$ decreases more rapidly than any power of x as $x\to\infty$, the integral in this formula converges for all s. This gives, therefore, another formula for $\zeta(s)$ which is "valid for all s" other than $s=0,1$, i.e., it gives an alternate proof of the fact that $\zeta(s)$ can be analytically continued. Since both sides of the above equation are analytic, the equation holds for all s. Further, $\int_1^\infty \psi(x)[x^{\frac{s}{2}}+x^{\frac{1-s}{2}}]\frac{dx}{x} - \frac{1}{s(s-1)}$ is unchanged by the substitution $s=1-s$, which proves the functional equation of the zeta function (Edwards, Section 1.7. "Second Proof of the Functional Equation" [**11**]). We have the following series representation[10] for $\xi(s)$, for $s = \frac{1}{2}+bi$:

$$\xi(s) = \sum_{n=0}^\infty a_{2n}\left(s-\frac{1}{2}\right)^{2n}, \quad \text{where } a_{2n} = 4\int_1^\infty \frac{d[x^{\frac{3}{2}}\psi'(x)]}{dx}x^{-\frac{1}{4}}\frac{(\frac{1}{2}\log x)^{2n}}{(2n)!}dx.$$

[9]Riemann reformulates $\int_0^\infty \psi(x)x^{\frac{s}{2}}\frac{dx}{x} = \int_1^\infty \psi(x)[x^{\frac{s}{2}}+x^{\frac{1-s}{2}}]\frac{dx}{x} + \frac{1}{2}\int_1^\infty [x^{-\frac{s-1}{2}} - x^{\frac{-s}{2}}]\frac{dx}{x}$; since $\int_1^\infty x^{-a}\frac{dx}{x} = \frac{1}{a}$ for $a>0$, so $\frac{1}{2}\int_1^\infty [x^{-\frac{s-1}{2}} - x^{\frac{-s}{2}}]\frac{dx}{x} = \frac{1}{2}\left[\frac{1}{\frac{s-1}{2}} - \frac{1}{\frac{s}{2}}\right] = \frac{1}{s(s-1)}$.

[10]Riemann states that this series representation of $\xi(s)$ as an even function of $s-\frac{1}{2}$ "converges very rapidly".

Riemann applies Fourier[11] inversion to the formula (for Re $s > 1$):

$$\frac{\log \zeta(s)}{s} = \int_0^\infty J(x) x^{-s-1} dx$$

$$\text{to obtain: } J(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \log \zeta(s) x^s \frac{ds}{s}.$$

Here, it is understood that the improper integral above "means the limit as $T \to \infty$ of the integral over the vertical line segment from $a - iT$ to $a + iT$" (Edwards, page 23 [**11**]). The improper integral given above is only conditionally convergent and an "order of summation" must be specified. Conditionally convergent integrals and series are very common in Fourier analysis, and it is always understood that such integrals and series are summed in their "natural[12] order" (Edwards, Section 1.12 "Fourier Inversion" [**11**]).

On page 86 of Riemann's *Partielle Differentialgleichungen* (ed. Hattendorf) (1876) [**41**], Riemann's use of "Fourier's theorem" is revealed to be simply Fourier inversion. The reader is encouraged to review Paper XII of Riemann's Collected Papers, namely *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, vol. 13*, "On the representation of a function by a trigonometric series" [**43**]. A remark by Fourier threw new light on the topic of the analytic representation of arbitrary functions: for the trigonometric series $f(x)$ the coefficients a_n and b_n can be determined by the formula below and this method can also be applied if $f(x)$ is any arbitrary function:

$$f(x) = \begin{cases} a_1 \sin x + a_2 \sin 2x + ... \\ + \frac{1}{2} b_0 + b_1 \cos x + b_2 \cos 2x + ... \end{cases}$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx.$$

In Riemann's words, "the applications of Fourier series are not restricted to research in the physical sciences. They are now also applied with success in an area of pure mathematics, number theory" (Riemann, Collected Papers, page 229 [**43**]).

[11] A "master of evaluating and estimating definite integrals" (Edwards, Section 1.9), Riemann was also "a master of Fourier analysis and his work in developing this theory must certainly be counted among his greatest contributions to mathematics" (Edwards, Section 1.12 "Fourier Inversion" [**11**]).

[12] Edwards provides some examples: $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ means $\lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{inx}$, and $\int_{-\infty}^{\infty} f(y) e^{iyx} dy$ means $\lim_{T\to\infty} \int_{-T}^{T} f(y) e^{iyx} dy$.

On page 5 [**42**], Riemann states that the number of roots of $\xi(t) = 0$ whose real parts lie between 0 and T is equal to the integral of $\frac{\xi'(s)ds}{2\pi i\xi(s)}$ around the boundary of the rectangle $\{0 \leq \text{Re } s \leq 1, 0 \leq \text{Im } s \leq 1,\}$ and this integral is equal to $\frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi}$ with a relative error of $\frac{1}{T}$. The fact that the vertical density of the roots ρ is in some sense a constant times $\log(\frac{T}{2\pi})$ was stated by Riemann without proof (Edwards, Section 1.10 "The Product Representation of $\xi(s)$" [**11**]). A proof of Riemann's estimate for the number of roots of $\xi(t) = 0$ whose real parts lie between 0 and T was first provided by von Mangoldt in *Zur Verteilung der Nullstellen der Riemannschen Funktion* $\xi(t)$ (1905) [**33**].

For Riemann, $\xi(t)$ is $\xi(s)$ with s replaced by $\frac{1}{2}+ti$. On page 5 (transl. Wilkins, 1998), Riemann originally states: "The number of roots of $\xi(t) = 0$, whose real parts lie between 0 and T is approximately: $= \frac{T}{2\pi}\log\frac{T}{2\pi} - \frac{T}{2\pi}$ because the integral $\int d\log\xi(t)$, taken in a positive sense around the region consisting of the values of t whose imaginary parts lie between $\frac{1}{2}i$ and $-\frac{1}{2}i$ and whose real parts lie between 0 and T, is (up to a fraction of the order of magnitude of the quantity $\frac{1}{T}$) equal to $(T\log\frac{T}{2\pi} - T)i$; this integral however is equal to the number of roots of $\xi(t) = 0$ lying within this region, multiplied by $2\pi i$. One now finds indeed approximately this number of real roots within these limits, and it is very probable that all roots are real" (Riemann [**42**]). As Bombieri remarks in *Problems of the Millennium: The Riemann Hypothesis* [**3**], the Riemann Hypothesis is the statement:

All zeros of the function $\xi(t)$ *are real.*

The Riemann Hypothesis remains unproven to this day.

The conjecture of Gauss and Legendre[13] remained unproven till in 1896, Hadamard (in *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques* [**22**]) and C. J. de la Vallée Poussin (in *Recherches analytiques sur la théorie des nombres premiers* [**51**]), using complex analytic methods, independently proved what is now referred to as the "Prime Number Theorem" (PNT), i.e.,

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1$$

In 1949, Selberg (*An elementary proof of the prime number theorem* [**45**]) and Erdös (*On a new method in elementary number theory which leads to an elementary proof of the prime number theorem* [**14**]) discovered "elementary" proofs of the prime number theorem, without using $\zeta(s)$ or complex function theory.

The idea of furnishing an elementary proof of the prime number theorem "caused a sensation in the mathematical world" (Apostol, page 9 [**2**]). Before Selberg and Erdös published their findings, Hardy addressed the Mathematical Society of Copenhagen in 1921, ruling out the possibility of an elementary proof of the prime number theorem, asserting that "some theorems "lie deep" and others nearer to the surface", and that "if anyone produces an elementary proof of the prime number theorem, he will show that these views are wrong, that the subject does not hang together in the way we have supposed, and that it is time for the books to be cast aside and for the theory to be rewritten" (Hardy, as quoted by Goldfield in *The Elementary Proof of the Prime Number Theorem: An Historical Perspective*, 2003 [**20**]).

[13] In "Historical Introduction", *Introduction to Analytic Number Theory* [**2**], Apostol traces the history of $\pi(n)$ back to Legendre (in *Essai sur la Theorie des Nombres*, 1798 [**30**]) and Gauss (in "Letter to Encke", dated 24, December, 1849 [**19**]). Apostol provides a table with columns x, $\pi(x)$, $\frac{x}{\log x}$, and $\frac{\pi(x)}{\frac{x}{\log x}}$, suggesting that though the distribution of the prime numbers is irregular, "by examining large blocks of primes one finds that their average distribution seems to be quite regular" (page 8, Apostol [**2**]). Thus, Gauss and Legendre correctly conjecture that as $x \to \infty$, $\frac{\pi(x)}{\frac{x}{\log x}} \to 1$, but lack the means to provide a proof. In 1851, Chebyshev [**7**] proved that if the ratio does tend to a limit, then this limit must be 1.

In Landau's *Handbuch der Lehre von der Verteilung der Primzahlen* [**28**], or simply *"Handbuch"*, hailed by Hardy (in the obituary of Landau he wrote for the London Mathematical Society [**24**]) as the first time the analytic theory of numbers is presented "not as a collection of a few beautiful scattered theorems, but as a systematic science", one may flip through the first few pages to see that Landau begins this systematic science with two simple functions, namely the Möbius function, μ, and the Liouville function, λ.

The Möbius function, $\mu(n)$, is defined as $(-1)^k$, if n is the product of k *distinct* primes and 0 otherwise; the Liouville Function $\lambda(n)$, on the other hand, is defined as $(-1)^k$, if n is a product of k primes (including multiplicities). Remarkably, μ and λ are closely related to ζ. On page 12 of *The Theory of the Riemann Zeta-Function* [**49**], Titchmarsch provides the following relationships between ζ and μ (equation 1.1.4, Titchmarsch [**49**]), and ζ and λ (equation 1.2.11, Titchmarsch [**49**]), for $s = a + ib$, with $a > 1$, as follows, the second identity attributed to Lehman (*On Liouville's Function*, 1960 [**31**]):

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \qquad \frac{\zeta(2s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s}$$

In Chapter 3 "Averages of arithmetical functions", *Introduction to Analytic Number Theory* [**2**], Apostol says that the statement "$\mu(n)$ has average order 0", i.e., "$\sum_{n\to\infty} \frac{\mu(n)}{n}$ converges and has sum 0", is equivalent to the PNT. Further, as referenced by Spencer-Brown in Appendix 9 [**48**], by Edwards in Section 12.3, "Miscellany" [**11**], by Apostol in Exercise 4 of Chapter 13 (page 301) [**2**], and by Derbyshire as Theorem 15-2, "Big Oh[14] and Möbius Mu" [**10**], RH is equivalent to $M(x) = O(x^{\frac{1}{2}+\epsilon}), \quad \forall\epsilon > 0$, where $M(x) = \sum_{n\leq x} \mu(n)$.

Similarly, Theorem 1.4 in *The Riemann Hypothesis* (Borwein et al, 2006) [**4**], states that the PNT is equivalent to the statement $\lim_{n\to\infty} \frac{\lambda(1)+\lambda(2)+\ldots+\lambda(n)}{n} = 0$. Further, as also referenced by Spencer-Brown on page 209 [**48**], and Borwein et al on page 6 (Theorem 1.2) [**4**], RH[15] is equivalent to $\lim_{n\to\infty} \frac{\lambda(1)+\lambda(2)+\ldots+\lambda(n)}{n^{\frac{1}{2}+\epsilon}} = 0$, for any fixed $\epsilon > 0$. Both of these results are attributed to Landau in his dissertation (1899) [**29**].

[14] We say $f(x)$ is $O(g(x))$ or "big-O of $g(x)$" if there are positive constants C and x_0 such that $|f(x)|\leq C|g(x)|$ whenever $x > x_0$.

[15] Spencer-Brown reiterates here that upon subsituting n for $n^{\frac{1}{2}+\epsilon}$ one obtains the much weaker PNT.

Let dM be the Stieltjes measure such that if $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$, (for Re $s > 1$),

$$\frac{1}{\zeta(s)} = \int_0^{\infty} x^{-s} dM(x),$$

where $M(x) = \int_0^x dM$ is a step function which is zero at $x = 0$, which is constant except at positive integers, and which has a jump[16] of $\mu(n)$ at n (Edwards, Section 12.1, "The Riemann Hypothesis and the Growth of $M(x)$" [**11**]). Integration by parts for Re $s > 1$, gives:

$$\frac{1}{\zeta(s)} =$$

$$\int_0^{\infty} d[x^{-s}M(x)] - \int_0^{\infty} M(x)d(x^{-s}) = \lim_{X\to\infty}\left[X^{-s}M(X) + s\int_0^X M(x)x^{-s-1}dx\right]$$

$$= s\int_0^{\infty} M(x)x^{-s-1}dx$$

because the inequality $|M(x)| \leq x$ implies that $x^{-s}M(x) \to 0$ as $x \to \infty$ and $\int_0^{\infty} M(x)x^{-s-1}dx$ converges, both provided Re $s > 1$. If $M(x)$ grows less rapidly than x^a for some $a > 0$, then this integral for $\frac{1}{\zeta(s)}$ converges for all s in the halfplane $\{\text{Re } (a - s) < 0\} = \{\text{Re } s > a\}$. By analytic continuation, the function $\frac{1}{\zeta(s)}$ is analytic in this halfplane. Since $\frac{1}{\zeta(s)}$ has poles on the line Re $s = \frac{1}{2}$,

$M(x)$ does not grow less rapidly than x^a for any $a < \frac{1}{2}$.

Moreover, it shows that *in order to prove the Riemann hypothesis, it would suffice to prove that $M(x)$ grows less rapidly than $x^{\frac{1}{2}+\epsilon}$ for all $\epsilon > 0$* (Edwards, Section 12.1, page 260-261 [**11**]). Littlewood in *Quelques conséquences de l'hypothèse que la fonction $\zeta(s)$ n'a pas de zéros dans le demi-plan Re $(s) > \frac{1}{2}$* (1912) [**32**] proved that this sufficient condition for the Riemann hypothesis is also necessary, hence proving the following **theorem**, attributed to Hardy:

The Riemann Hypothesis (RH) is equivalent to the statement that

for every $\epsilon > 0$, $\dfrac{M(x)}{x^{\frac{1}{2}+\epsilon}}$ approaches zero as $x \to \infty$.

This theorem will be very important for the rest of the paper. On the one hand, it is the source of speculation for RH and coin-tossing and on the other hand, it is the central aspect with which Spencer-Brown writes Appendix 9 of Laws Of Form.

[16] The value of M at a jump is by definition $\frac{1}{2}[M(n-\epsilon) + M(n+\epsilon)] = \sum_{j=1}^{n-1} \mu(j) + \frac{1}{2}\mu(n)$.

An argument for the plausibility of RH can be found in Chapter 12, "Miscellany" of Edwards' *Riemann's Zeta Function*, in Section 12.3, "Denjoy's Probabilistic Interpretation[17] of the Riemann Hypothesis" [**11**]. On page 268, Edwards reasons as follows: "Suppose an unbiased coin is flipped a large number of times, say N times. By the Moivre-Laplace theorem the probability that the number of heads deviates by less than $KN^{\frac{1}{2}}$ from the expected number of $\frac{1}{2}N$ is nearly equal to $\int_{-(\frac{2K^2}{\pi})^{\frac{1}{2}}}^{(\frac{2K^2}{\pi})^{\frac{1}{2}}} \exp(-\pi x^2)dx$ in the sense that the limit of these probabilities as $N \to \infty$ is equal to this integral. Thus if the total number of heads is subtracted from the total number of tails, the probability that the resulting number is less than $2KN^{\frac{1}{2}}$ in absolute values is nearly equal to $2\int_0^{\frac{2K^2}{\pi}^{\frac{1}{2}}} \exp(-\pi x^2)dx$. The fact that this approaches 1 as $N \to \infty$ can be regarded as saying that *with probability one the number of heads minus the number of tails grows less rapidly than* $N^{\frac{1}{2}+\epsilon}$" (Edwards, page 268, Section 12.3 [**11**]).

If we consider a very large square-free integer n, then $\mu(n) = \pm 1$. "But then the evaluation of $M(x)$ would be like flipping a coin once for each square-free integer less than x and subtracting the number of heads from the number of tails. It was shown above that for any given $\epsilon > 0$ the outcome of this experiment for large[18] number of flips is, with probability nearly one, less than the number of flips raised to the power $\frac{1}{2} + \epsilon$ and *a fortiori* less than $x^{\frac{1}{2}+\epsilon}$" (Edwards, page 268, Section 12.3 [**11**]).

[17] In pages 195 - 196 of *Probabilités confirmant l'hypothèse de Riemann sur les zéros de* $\zeta(s)$ [**9**], we find Denjoy's original statement: "En outre $\mu(1) = 1$.

$$\text{Posons,} \quad \Delta(n) = \sum_1^n \mu(i).g(s) = \sum_1^\infty \Delta(n)[n^{-s} - (n+1)^{-s}] \sim s \sum \Delta(n) n^{-1-s},$$

$$\mid s \mid^{-1} \mid g(s) \mid \leq \sum \mid \Delta(n) \mid n^{-1-\sigma}$$

Il suffrait donc de prouver que $\overline{\lim}_{n\to\infty} \log \Delta(n)/\log n = \frac{1}{2}$ (l'inégalité $< \frac{1}{2}$ est impossible) pour que l'hypothèse de Riemann soit justifiée." Also, "Si $s = \sigma + it$, $\mid \sigma$ et t étant réels, la série et le produit infini convergent absolument pour $\sigma > 1$" (Denjoy, page 195 [**9**]).

[18] "The number of flips goes to infinity as $x \to \infty$ because, among other reasons, there are infinitely many primes, hence *a fortiori* infinitely many square-free integers (products of distinct primes)" (original footnote, Edwards, page 268).

On page 268, Section 12.3, "Denjoy's Probabilistic Interpretation of the Riemann Hypothesis" [**11**], Edwards, however, goes on to write, "it is perhaps plausible to say that successive evaluations of $\mu(n) = \pm 1$ are "independent" since knowing the value of $\mu(n)$ for one n would not seem to give any[19] information about its values for other values of n."

As pointed out by Spencer-Brown, it would be remiss to consider successive evaluations of $\mu(n) = \pm 1$ as independent of one another. In *Essai sur la théorie des nombres* [**30**], Legendre reformulates the prime counting function $\pi(n)$ as a function of $\pi(\sqrt{n})$ using the Möbius function μ. Spencer-Brown rectifies this formula for $d = 1$ and offers a simple recursive formula to compute $\mu(n)$, thus showing that subsequent values of $\mu(n) = \pm 1$ are **not** "independent".

The Spencer-Brown Formula[20] Möbius $Mu(n)$ is presented as follows: for $n > 1$,

$$\mu(n) = -\Big(n - 1 + \sum_{d=2}^{n-1} \mu(d) \left[\frac{n}{d}\right]\Big) = 1 - n - \sum_{d=2}^{n-1} \mu(d) \left[\frac{n}{d}\right],$$

where the bracketed variable $[\frac{n}{d}]$ denotes the integer part of this fraction.

[19]an exception to this statement is that for any prime p, $\mu(pn)$ is either $-\mu(n)$ or zero. However, this principle can only be applied once for any p because $\mu(p^2 n) = 0$ and this "information" really says little more than that μ is determined by a formula and is not, in fact, a random phenomenon (original footnote by Edwards).

[20]"There are two Spencer-Brown formulae $Mu(n)$. They are both algebraic and connected by the sign of equality: '$=$'. It should be observed that both formulae are sensitive to parentheses. If we remove the parentheses from the first formula, the minus sign distributes and converts the formula into the second. Conversely, if we add parentheses and a minus sign the second formula converts once more into the first. The formula was designed specifically to prove RH. It is a modification of procedures employed by Legendre and Sierpinski" (J. M. Flagg in *Conversations with GSB, Oct 8, 2012*).

The sixth English edition of "Laws of Form" (2015) [**48**] includes George Spencer-Brown's most recent version of "Appendix 9: A proof of Riemann's hypothesis using Denjoy's equivalent theorem". It begins with a "Description of the proof", which states that "The Riemann hypothesis is true if and only if the numbers of positive and negative signs of $\mu(n)$ are asymptotically[21] equal", referencing Denjoy's probabilistic interpretation[22], that "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 (different) prime divisors" (page 203, "Laws of Form", 2015 [**48**]). We note that Spencer-Brown is undoubtedly referring to the full technical statement that $M(x) = O(x^{\frac{1}{2}+\epsilon})$.

The Legendre/Spencer-Brown formula for the prime counting function $\pi(n)$ is a recursive function that calculates the number of primes $\leq n$ by referring to $\pi(\sqrt{n})$. The formula only uses primes up to $\sqrt{n}$, what Spencer-Brown calls "small" primes, to sieve the remaining primes up to n, what Spencer-Brown calls the "large" primes. The Legendre formula, rectified by Spencer-Brown, is presented as follows:

$$\pi(n) = \pi(\sqrt{n}) + \sum_{d=1}^{\sqrt{n}} \mu(d) \left[\frac{n}{d}\right].$$

Specifically for $d = 1$, Spencer-Brown rectifies the dividend to $\left[\frac{n-1}{d}\right]$. Note that the integer part of the function $\left[\frac{n}{d}\right]$ is often written as the "floor" function $\left\lfloor\frac{n}{d}\right\rfloor$. Every number up to $\sqrt{n}$ factors into a product of primes. We sum over all d up to $\sqrt{n}$ but $\mu(d) = 0$ when there are repeated prime factors. Therefore we need only sum over those d that have single prime factors up to $\sqrt{n}$ number of factors.

[21] Historically credited to Paul Bachmann in *Die Analytische Zahlentheorie* "Analytic Number Theory", 1894, the big O notation currently used to describe orders of magnitude was first clearly defined by Landau. "More specifically, the $f(x) \sim g(x)$ relation for two functions $f, g : \mathbb{R} \to \mathbb{R}$ means that $\lim_{x\to\infty}(\frac{f(x)}{g(x)}) = 1$. Under these conditions, the two functions are said to be *asymptotically* equivalent (or simply *asymptotic*). Alternatively, we say that $f(x)$ is $O(g(x))$ (pronounced "big-O of $g(x)$") if there are positive constants C and x_0 such that $|f(x)| \leq C|g(x)|$ whenever $x > x_0$. It is not too difficult to see that $f(x) \sim g(x)$ implies that $f(x)$ is $O(g(x))$ and $g(x)$ is $O(f(x))$" (Tou, in *Math Origins: Orders of Growth* [**50**]).

[22] *Probabilités confirmant l'hypothèse de Riemann sur les zéros de $\zeta(s)$. Note $(*)$ de M. Arnaud Denjoy, Membre de l'Académie, in "Conclusions tirées d'une forme donnée au développement de $\frac{\zeta(2s)}{\zeta(s)}$"* [**9**]

for $n = 100$,

$$\pi(100) = \pi(10) + \sum_{d=1}^{\sqrt{100}} \mu(d) \left[\frac{100}{d}\right].$$

The small primes are all primes up to $\sqrt{100} = 10$, which are $2, 3, 5$, and 7. They combine to form products that divide 100. They can be organized by collecting the doubles, triples, and so on:

$$\overbrace{2,3,5,7}, \overbrace{(2.3),(2.5),(2.7),(3.5),(3.7),(5.7)}, \overbrace{(2.3.5),(2.3.7),(2.5.7),(3.5.7)}, (2.3.5.7)$$

$$\text{Hence, } \sum_{d=1}^{\sqrt{100}} \mu(d) \left[\frac{100}{d}\right]$$

$$= \Bigg(\mu(1)\left[\frac{100-1}{1}\right] + \mu(2)\left[\frac{100}{2}\right] + \mu(3)\left[\frac{100}{3}\right] + \mu(5)\left[\frac{100}{5}\right] + \mu(7)\left[\frac{100}{7}\right]$$
$$+\mu(2.3)\left[\frac{100}{2.3}\right] + \mu(2.5)\left[\frac{100}{2.5}\right] + \mu(2.7)\left[\frac{100}{2.7}\right] + \mu(3.5)\left[\frac{100}{3.5}\right] + \mu(3.7)\left[\frac{100}{3.7}\right]$$
$$+\mu(5.7)\left[\frac{100}{5.7}\right] + \mu(2.3.5)\left[\frac{100}{2.3.5}\right] + \mu(2.3.7)\left[\frac{100}{2.3.7}\right] + \mu(2.5.7)\left[\frac{100}{2.5.7}\right]$$
$$+\mu(3.5.7)\left[\frac{100}{3.5.7}\right] + \mu(2.3.5.7)\left[\frac{100}{2.3.5.7}\right]\Bigg)$$

$$= \Bigg(\mu(1)\left[\frac{99}{1}\right] + \mu(2)\left[\frac{100}{2}\right] + \mu(3)\left[\frac{100}{3}\right] + \mu(5)\left[\frac{100}{5}\right] + \mu(7)\left[\frac{100}{7}\right]$$
$$+\mu(2.3)\left[\frac{100}{6}\right] + \mu(2.5)\left[\frac{100}{10}\right] + \mu(2.7)\left[\frac{100}{14}\right] + \mu(3.5)\left[\frac{100}{15}\right] + \mu(3.7)\left[\frac{100}{21}\right]$$
$$+\mu(5.7)\left[\frac{100}{35}\right] + \mu(2.3.5)\left[\frac{100}{30}\right] + \mu(2.3.7)\left[\frac{100}{42}\right] + \mu(2.5.7)\left[\frac{100}{70}\right]$$
$$+\mu(3.5.7)\left[\frac{100}{105}\right] + \mu(2.3.5.7)\left[\frac{100}{210}\right]\Bigg)$$

$$= \Bigg(+\left[\frac{99}{1}\right] - \left[\frac{100}{2}\right] - \left[\frac{100}{3}\right] - \left[\frac{100}{5}\right] - \left[\frac{100}{7}\right]$$
$$+\left[\frac{100}{6}\right] + \left[\frac{100}{10}\right] + \left[\frac{100}{14}\right] + \left[\frac{100}{15}\right] + \left[\frac{100}{21}\right] + \left[\frac{100}{35}\right]$$
$$-\left[\frac{100}{30}\right] - \left[\frac{100}{42}\right] - \left[\frac{100}{70}\right] - \left[\frac{100}{105}\right] + \left[\frac{100}{210}\right]\Bigg)$$

$$= +99 - 50 - 33 - 20 - 14$$

$$+16 + 10 + 7 + 6 + 4 + 2$$

$$-3 - 2 - 1 - 0 + 0$$

$$= 21$$

Hence, according to LSB (Legendre/Spencer-Brown), $\pi(100) = \pi(10) + 21$. But we know $\pi(10) = 4$, when we listed out all the small primes $2, 3, 5$, and 7.

$$\pi(100) = \pi(10) + 21$$

$$= 4 + 21 = \mathbf{25}.$$

The Legendre/Spencer-Brown formula can be tested for any n and the formula "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., such that $2 \leq q \leq n^{\frac{1}{2}}$" (page 203, Spencer-Brown, 2015 [**48**]). In every application of the Legendre procedure, the small primes, i.e., the primes up to $\sqrt{n}$, are needed to count the large primes, i.e., primes up to n. Two important observations:

When the divisor d exceeds n, the 'floored' quotient $\left[\frac{n}{d}\right]$ returns 0 and so $f(d) = \mu(d)\left[\frac{n}{d}\right] = \mu(d).0 = 0$. We saw this with divisors $3.5.7 = 105$ and $2.3.5.7 = 210$ in the Legendre/Spencer-Brown demonstration for $n = 100$ on the previous page.

When a square-free divisor d exceeds $\frac{n}{2}$ but remains $\leq n$, however, the quotient $\left[\frac{n}{d}\right]$ yields 1, and $f(d) = \mu(d)\left[\frac{n}{d}\right]$ is simply $\mu(d)$. This important observation leads Spencer-Brown to call the square-free numbers $\leq \frac{n}{2}$ the "upper section", and the square-free numbers $> \frac{n}{2}$ but $\leq n$ the "lower section". On page 205, Spencer-Brown writes, "in the lower section the sum of the terms is exactly the sum of the $\mu(d)$ in the section. Call a series of such terms an LSB series" (page 205, Appendix 9 [**48**]).

"The fact that with d unrestricted, the formula $\sum \mu(d)[\frac{n}{d}] = 1$ is true for all n was first noted by Meissel in *Observationem quaedam in theoria numerorum*, Berlin 1850, and proved by Sierpiński in *Elementary theory of numbers*, Warsaw 1964, pp 180 and 181" (Spencer-Brown, footnote on page 204 [**48**]). The point here is that instead of summing from 1 to $\sqrt{n}$ in the Legendre summation, one can sum from 1 to n so one has $\sum_{d=1}^{n} \mu(d)[\frac{n}{d}] = 0$, by using Spencer-Brown's adjustment. On page 204, he provides an illustration of the use of Legendre's formula to calculate $\pi(n)$ for $n = 20$:

d	$f(d) = \mu(d)[\frac{n}{d}]$
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	$\underline{-1}$
	$\sum +1$

$\sum \mu(d^*)[\frac{n}{d^*}] = 7$
$7 - 1 + \pi(n^{\frac{1}{2}}) = 8 = \pi(20)$

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 Spencer-Brown's words: "Since 1 is not struck out by the sieve of Eratosthenes and is also included in the count of "primes" calculated by the section $\sum \mu(d)[\frac{n}{d}]$, 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" (Spencer-Brown, on page 204). "Notice I have used all (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" (Spencer-Brown, on page 204 [**48**]).

"We thus further rectify the procedure by making the following change: use $f(d) = \mu(d)[\frac{n-1}{d}]$ for $d = 1$, and use $f(d) = \mu(d)[\frac{n}{d}]$ for all other values of d" (Spencer-Brown, page 205 [**48**]). This rectification ensures:

$$\sum_{d=1}^{n} f(d) = 0, \qquad \text{where } f(d) = \mu(d)\left\lfloor \frac{n}{d} \right\rfloor, \text{ and for } d = 1, f(d) = \mu(d)\left\lfloor \frac{n-1}{d} \right\rfloor.$$

"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 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 set of magnified differences tends to zero, then the average of the same set of differences unmagnified must also tend to zero" (page 214, Appendix 9 [**48**]).

The negative feedback property of $\mu(n)$ is used by Spencer-Brown to argue that the cumulative Möbius function varies around zero and that its asymptotics are sufficient, by Denjoy's interpretation, to prove the Riemann Hypothesis (RH), which is equivalent to the relation $M(x) = O(x^{\frac{1}{2}+\epsilon})$, $\forall \epsilon > 0$, where $M(x) = \sum_{n \leq x} \mu(n)$.
This argument of George Spencer-Brown deserves further investigation.

Appendix ±1

We are presented with The Spencer-Brown Formula[23] Möbius $Mu(n)$, which can be seen to follow as a consequence of the Legendre/Spencer-Brown formula for $\pi(n)$. The Spencer-Brown Formula Möbius $Mu(n)$ (J. M. Flagg, *Conversations with GSB, Oct 8, 2012)*:

$$\mu(n) = -\Big(n - 1 + \sum_{d=2}^{n-1} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor\Big) = 1 - n - \sum_{d=2}^{n-1} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor$$

$$= -\left(\sum_{d=1}^{n-1} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor \right),$$

since for $d = 1$, we make the Spencer-Brown adjustment, rectifying $[\frac{n}{d}]$ to $[\frac{n-1}{d}]$.

Start with $\mu(1) = +\mathbf{1}$.
$\mu(2) = -(\mu(1)[\frac{2-1}{1}]) = -\mathbf{1}$.

$\mu(3) = -(\mu(1)[\frac{3-1}{1}] + \mu(2)[\frac{3}{2}]) = -(2 - 1) = -\mathbf{1}$.

$\mu(4) = -(\mu(1)[\frac{4-1}{1}] + \mu(2)[\frac{4}{2}] + \mu(3)[\frac{4}{3}]) = -(3 - 2 - 1) = \mathbf{0}$.

$\mu(5) = -(\mu(1)[\frac{5-1}{1}] + \mu(2)[\frac{5}{2}] + \mu(3)[\frac{5}{3}] + \mu(4)[\frac{5}{4}]) = -(4 - 2 - 1 + 0) = -\mathbf{1}$.

$\mu(6) = -(\mu(1)[\frac{6-1}{1}] + \mu(2)[\frac{6}{2}] + \mu(3)[\frac{6}{3}] + \mu(4)[\frac{6}{4}] + \mu(5)[\frac{6}{5}]) = -(5 - 3 - 2 + 0 - 1) = +\mathbf{1}$.

23

$$\text{(for } n > 1\text{)}, \quad \mu(n) = -\Big(n - 1 + \sum_{d=2}^{n-1} \mu(d) \left[\frac{n}{d}\right]\Big) = 1 - n - \sum_{d=2}^{n-1} \mu(d) \left[\frac{n}{d}\right],$$

where the bracketed variable $\left[\frac{n}{d}\right]$ denotes the integer part of $\frac{n}{d}$.

Spencer-Brown's cascade (pages 215-216) [**48**] is a tabular arrangement: The first column lists the square-free numbers $d \leq n$. The second column lists $f(d) = \mu(d)\left\lfloor \frac{n}{d} \right\rfloor$ (where, for $d = 1$, $\frac{n}{d}$ is rectified to $\frac{n-1}{d}$). The third column lists $\sum f(d)$, or what Spencer-Brown calls the "running total".

for $n = 2$,

d	$f(d)$	$\sum f(d)$
1	$+1$	$+1$
$\mu(2) =$	-1	0

for $n = 3$,

d	$f(d)$	$\sum f(d)$
1	$+2$	$+2$
2	-1	$+1$
$\mu(3) =$	-1	0

for $n = 4$,

d	$f(d)$	$\sum f(d)$
1	$+3$	$+3$
2	-2	$+1$
3	-1	0
$\mu(4) =$	0	0

for $n = 5$,

d	$f(d)$	$\sum f(d)$
1	$+4$	$+4$
2	-2	$+2$
3	-1	$+1$
$\mu(5) =$	-1	0

for $n = 6$,

d	$f(d)$	$\sum f(d)$
1	$+5$	$+5$
2	-3	$+2$
3	-2	0
5	-1	-1
$\mu(6) =$	$+1$	0

On page 217, Spencer-Brown says: "All the cascade requires is what 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 in the running total with the sign reversed". Thus, the penultimate term in $\sum f(d)$ with reversed sign

$$= -\left(\sum_{d=1}^{n-1} f(d)\right) = f(n) \quad (= \text{the final } f(d) \text{ term} = \mu(n)).$$

For Spencer-Brown's cascade on pg. 215-217 [**48**], note that the final term of $\sum f(d) = 0$. In his Letter to Moshe Klein, Spencer-Brown provides a variation of the cascade on pg. 215-216 where the final term of $\sum f(d) = \mu(n)$. Consider,

d	$n = 2,$	
	-1	-1

$\therefore\ \mu(2) = -1$

d	$n = 3$	
	-2	-2
2	$+1$	-1

$\therefore\ \mu(3) = -1$

d	$n = 4$	
	-3	-3
2	$+2$	-1
3	$+1$	0

$\therefore\ \mu(4) = 0$

d	$n = 5$	
	-4	-4
2	$+2$	-2
3	$+1$	-1

$\therefore\ \mu(5) = -1$

d	$n = 6$	
	-5	-5
2	$+3$	-2
3	$+2$	0
5	$+1$	$+1$

$\therefore\ \mu(6) = +1$

We may now see how Spencer-Brown rectifies Legendre's method of counting large primes in n and then corrupts it to give the answer zero for all (n). By this he precisely means the cascading algorithm used in pg. 215-216 [**48**], in which the final term of the "running total" or $\sum f(d)$ is zero. Recall that Spencer-Brown splits $n \leq d$ into two sections at $\frac{n}{2}$, calling $d < \frac{n}{2}$ the upper section and $\frac{n}{2} < d \leq n$ the lower section, since for the lower section $\mu(d)\lfloor \frac{n}{d} \rfloor = \mu(d)$. Thereupon, Spencer-Brown says that "any difference between the upper and the lower sections can be seen to be self-annihilating, because of the negative feedback between the two sections" (page 206, [**48**]). For any square-free number n, such as $n = 20$, since the final $f(d)$ term, $f(n) = \mu(n)$ equals 0, if we use the cascading algorithm in Appendix 9,

$$\sum_{d=1}^{n} f(d) = 0.$$

By using the generalizable example of $n = 20$, and doubling n to $2n = 40$, and then doubling $2n$ to $4n = 80$, and so on, the lower sections eventually include all values of $\mu(d)$ for square-free $(d) > 1$" (page 214, Appendix 9 [**48**]), and 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" (page 214, Appendix 9, [**48**]), since the LSB series, or Spencer-Brown's cascade, establishes a negative feedback relation between the final $f(d)$ term and penultimate $\sum f(d)$ term (with reversed sign). Thus, if $\mu(n)$ is the n^{th} term of $f(d)$, and $\mu(n) = f(n) = \pm 1$, then the $(n-1)^{th}$ term of $\sum f(d) = \mp 1$, for square-free n.

Appendix 0

We first explore the identity that $(1+p_1)(1+p_2)\ldots(1+p_n) =$ the sum of all possible products of p_i with exactly one or no appearance of each p_i. This implies that the sum of their μ values is equal to $(1+(-1))^N = 0$ (here N is the number of primes being considered). Recall that the Legendre/Spencer-Brown (LSB) formula[24] requires the primes up to $\sqrt{n}$ to find the primes up to n. For the sake of computing $\pi(n)$, build all square-free divisors $d \leq n$, by starting with the first square-free divisor 1. Call:

$\{p_i\} = \{p_i | \text{"}p_i \text{ is prime" and "}1 \leq i \leq \pi(\sqrt{n})\text{"}\}$
$\{p_i.p_j\} = \{(p_i.p_j) | \text{"}p_i, p_j \text{ are prime" and "}1 \leq i < j \leq \pi(\sqrt{n})\text{"}\}$
$\{p_i.p_j.p_k\} = \{(p_i.p_j.p_k) | \text{"}p_i, p_j, p_k \text{ are prime" and "}1 \leq i < j < k \leq \pi(\sqrt{n})\text{"}\}$
...

Note how the small primes up to $\sqrt{n}$ combine uniquely to form a large (but finite) collection of square-free divisors $d = \{1, \{p_i\}, \{p_i.p_j\}, \{p_i.p_j.p_k\}, ...\}$. The last item in this list is the product of all the primes up to $\sqrt{n}$. By definition of the Möbius function, we know that $\mu(1) = +1$, $\mu(p_i) = -1$, $\mu(p_i.p_j) = +1$, $\mu(p_i.p_j.p_k) = -1$, ..., for all primes p_i, p_j, and p_k. We wish to keep track of the $+1$'s and -1's contributed by these square-free divisors. Let $\{p_i\}$ have m elements. What is the total number of elements of $\{p_i.p_j\}$?

Let $\#S =$ the number of elements of a set S. For $m = 5$,

[24]The Legendre formula, rectified by Spencer-Brown, is presented as follows:

$$\pi(n) = \pi(\sqrt{n}) + \sum_{d=1}^{\sqrt{n}} \mu(d) \left[\frac{n}{d}\right].$$

Specifically for $d = 1$, Spencer-Brown rectifies the dividend to $\left[\frac{n-1}{d}\right]$. We take the bracketed variable $\left[\frac{n}{d}\right]$ to mean the floored quotient $\left\lfloor\frac{n}{d}\right\rfloor$. Further, d is any number $\leq n$ that can be expressed as a product of primes up to $\sqrt{n}$.

$\#\{p_i\} = \#\{2, 3, 5, 7, 11\} = 5 = m$, in general.

$\#\{p_i.p_j\} = \#2\{3, 5, 7, 11\} = \#\{2.3, 2.5, 2.7, 2.11\} = 4 = (m-1)$ elements, in general.
$+\#3\{5, 7, 11\}, = \#\{3.5, 3.7, 3.11\} = 3$ elements
$+\#5\{7, 11\}, = \#\{5.7, 5.11\} = 2$ elements
$+\#7\{11\} = \#\{7.11\} = 1$ element.

Note that $\{p_i.p_j\}$ has $1+2+3+4$ elements, and in general, $\#\{p_i.p_j\} = 1+2+3+...+(m-1)$.

$$\#\{p_i.p_j.p_k\} = \#2\begin{cases} 3\{5,7,11\} & = \#\{2.3.5, 2.3.7, 2.3.11\} = 3 = (m-2) \text{ elements} \\ 5\{7,11\} & = \#\{2.5.7, 2.5.11\} = 2 \text{ elements} \\ 7\{11\} & = \#\{2.7.11\} = 1 \text{ element} \end{cases}$$
$$+\#3\begin{cases} 5\{7,11\} & = \#\{3.5.7, 3.5.11\} = 2 \text{ elements,} \\ 7\{11\} & = \#\{3.7.11\} = 1 \text{ element} \end{cases}$$
$$+\#5\begin{cases} 7\{11\} & = \#\{5.7.11\} = 1 \text{ element.} \end{cases}$$

Note that $\{p_i.p_j.p_k\}$ has $1 + (1+2) + (1+2+3)$ elements, and in general,
$\#\{p_i.p_j.p_k\} = 1 + (1+2) + (1+2+3) + ... + (1+2+3+...+(m-2))$.

Recall that $\#\{p_i.p_j\} = 1+2+3+...+(m-2)+(m-1)$. Notice how this sum, can be broken up into two parts, $1+2+3+...+(m-2)$, and $(m-1)$. Take the first part. It is equal to the largest collection in $\{p_i.p_j.p_k\}$ with $(1+2+3+...+(m-2))$ elements.

Similarly, recall that $\#\{p_i\} = m$. We can break this sum into two parts $(m-1)$ and 1. Take the first part. It equals the second part $(m-1)$ of $\{p_i.p_j\}$. The second part, 1 of $\{p_i\}$, however, is equal to the first square-free divisor 1.

As we collect all square-free products of $p_i, p_j, p_k, ...$ into $\{1, \{p_i\}, \{p_i.p_j\}, \{p_i.p_j.p_k\}, ...\}$, we see the 1 giving $\mu(1) = +1$, the $\{p_i\}$ giving $\mu(p_i) = -1$, the $\{p_i.p_j\}$ giving $\mu(p_i.p_j) = +1$, and so on, and we observe how one collection of $+1$'s cancels out a preceding or successive collection of -1's in our set of divisors.

$\{1\}$ has 1 element.

$\{p_i\}$ has m elements $= (m-1)+1$

$\{p_i.p_j\}$ has $1+2+3+\ldots+(m-2) \quad +(m-1)$ elements.

$$\{p_i.p_j.p_k\} \text{ has } \begin{matrix} 1 \\ +1+2 \\ +1+2+3 \\ +... \\ +1+2+3+\ldots+(m-3) \end{matrix} \quad + \quad 1+2+3+\ldots+(m-2) \quad \text{elements.}$$

$$\{p_i.p_j.p_k.p_l\} \text{ has } 1 + \begin{matrix} 1 \\ +1+2 \end{matrix} + \begin{matrix} 1 \\ +1+2 \\ +1+2+3 \end{matrix} + \ ...\ + \begin{matrix} 1 \\ +1+2 \\ +1+2+3 \\ +\ldots \\ +1+2+3+\ldots+(m-4) \end{matrix}$$

$$+ \begin{matrix} 1 \\ +1+2 \\ +1+2+3 \\ +\ldots \\ +1+2+3+\ldots+(m-3) \end{matrix} \quad \text{elements.}$$

Notice how the total number of divisors in each collection correspond to the counts of the vertical columns of Pascal's triangle:

1						
1	1					
1	2	1				
1	3	3	1			
1	4	6	4	1		
1	5	10	10	5	1	
1	6	15	20	15	6	1

We see how column 1 corresponds to the total number of divisors in $\{p_i\}$,
column 2 corresponds to the total number of divisors in $\{p_i.p_j\}$,
column 3 corresponds to the total number of divisors in $\{p_i.p_j.p_k\}$,
column 4 corresponds to the total number of divisors in $\{p_i.p_j.p_k.p_l\}$,
...

Column 1 $= 1+1+1+1+1+1+1+... = \#\{p_i\}$ with $\mu(d) = -1$.
Column 2 $= 1+2+3+4+5+6+... = \#\{p_i.p_j\}$ with $\mu(d) = +1$. If we look at the 6 in this list, it is the sum of all the entries of column 1 till 5, i.e., $1+1+1+1+1+1 = 6$.
Column 3 is the sum of *triangular* numbers $= 1+3+6+10+15+... = \#\{p_i.p_j.p_k\}$ with $\mu(d) = -1$. If we look at the 15 in this list, it is the sum of all entries of column 2 till 10, i.e., $1+2+3+4+5 = 15$.
Column 4 is the sum of *tetrahedral* numbers $= 1+4+10+20+35+... = \#\{p_i.p_j.p_k.p_l\}$ with $\mu(d) = +1$. If we look at the 20 in this list, it is the sum of all entries of column 3 till 10, i.e., $1+3+6+10 = 20$.
Column 5 is the sum of *pentatope* numbers $= 1+5+15+35+70+... = \#\{p_i.p_j.p_k.p_l.p_m\}$ with $\mu(d) = -1$. If we look at the 15 in this list, it is the sum of all entries of column 4 till 5, i.e., $1+4+10 = 15$.

Note: Pascal referred to this self reference as a “hockey stick pattern”. When the numbers are arranged in Pascal’s form, the rows and columns are interchangeable as follows:

1	1	1	1	1	1	1
1	2	3	4	5	6	...
1	3	6	10	15	...	
1	4	10	20	...		
1	5	15	...			
1	6	...				
1	...					

The “hockey stick pattern” that Pascal is referring to is identical to the dovetailing identities we saw in the previous page. If we look at the 6 in Column 2, it is the sum of all numbers in the previous Column 1 up to 6 itself. This gives us the identity $1+1+1+1+1+1=6$, which is the same as what we had on the previous page. Similarly, if we look at the 15 in Column 3, it is the sum of all numbers in the previous Column 2 up to 15 itself. This gives us the identity $1+2+3+4+5=15$, which is the same as what we had on the previous page. Similarly, if we look at the 20 in Column 4, it is the sum of all numbers in the previous Column 3 up to 20 itself. This gives us the identity $1+3+6+10=20$. The 15 in Column 5 is the sum of all numbers in the previous Column 4 up to 15 itself, giving us the identity $1+4+10=15$. The 6 in Column 6 is the sum of all numbers in the previous Column 5 up to 6 itself, giving us the identity $1+5=6$. Finally, the 1 in the last Column 7 equals the adjacent 1 in the previous Column 6:

$1+1+1+1+1+1=6$

$1+2+3+4+5=15$

$1+3+6+10=20$

$1+4+10=15$

$1+5=6$

$1=1$

Pascal said[25],"In every arithmetical triangle each cell is equal to the sum of all the cells of the preceding row from its column to the first, inclusive". Since the rows and columns are interchangeable, we also have "In every arithmetical triangle each cell is equal to the sum of all the cells of the preceding column from its row to the first, inclusive".

In collecting all square-free divisors from a set of prime numbers $\{p_i\}$ of size m, we notice that the count of all possible combinations of square-free divisors generated by the set $\{p_i\}$, when collected combinatorially in collections of $\{p_i\}, \{p_i.p_j\}$, $\{p_i.p_j.p_k\}, \dots$, correspond exactly to the counts in the columns of Pascal's triangle. If the columns of Pascal's triangle are assigned $-$ or $+$ in an alternating way, i.e., if Column 1 is assigned $-$, Column 2 is assigned $+$, Column 3 is $-$, Column 4 is $+$,..., then not only do these columns correspond to the counts of square-free divisors for each collection $\{p_i\}, \{p_i.p_j\}, \{p_i.p_j.p_k\}, ...,$ but also each column now corresponds to the cumulative sum of Möbius values of the square-free divisors that correspond to that column, contributing a positive Möbius sum for columns $2, 4, 6, 8, ...$ or negative Möbius sum for columns $1, 3, 5, 7, ...$.

Column 1, which now represents the singular prime count up to m, gives $\mu(d) = -1$ throughout the column. So, if we sum all members of Column 1, we are summing the Möbius values of each of the divisors that Column 1 corresponds to, which is simply $\#\{p_i\}$. Column 2, which represents all possible $(p_i.p_j)$ combinations, gives $\mu(d) = +1$. So, if we sum all members of Column 2, we are summing the Möbius values of each of the divisors that Column 2 corresponds to, which is simply $\#\{p_i.p_j\}$. Column 3, which represents all possible $(p_i.p_j.p_k)$ combinations, gives $\mu(d) = -1$. So, if we sum all members of Column 3, we are summing the Möbius values of each of the divisors that Column 3 corresponds to, which is simply $\#\{p_i.p_j.p_k\}$.

Pascal's "hockey stick pattern" shows us that the alternating sum of all columns of a finite Pascal's triangle equals -1. One may verify this for a Pascal's triangle of any size. For a Pascal's triangle of size $m = 7$, the alternating sum of all 7 columns
$= -(1+1+1+1+1+1+1) + (1+2+3+4+5+6) - (1+3+6+10+15) + (1+4+10+20) - (1+5+15) + (1+6) - 1 = \mathbf{-1}$.

Column 1 has one remaining element p_* that survives as all the other columns cancel each other out in the alternate sign summation of all columns of Pascal's triangle. We know $\mu(p_*) = -1$. We see that p_* could be none other than the first prime 2, as shown in the first Pascal

[25]https://www.cut-the-knot.org/arithmetic/combinatorics/PascalTriangleProperties.shtml

arrangement, or the m-th prime as seen in Pascal's original arrangement[26]. This taken together with $\mu(1) = +1$ will always give us a total Möbius sum of zero for all square-free divisors generated by a finite set of primes, including the set of primes and the first square-free divisor 1.

THEOREM 1. *A. Let n be a natural number. Let $p_1, p_2, p_3, ...,$ enumerate the primes up to $\sqrt{n}$.*
Let d be a square-free divisor in $\{1, \{p_i\}, \{p_i.p_j\}, \{p_i.p_j.p_k\}, ...\}$, where

$\{p_i\} := \{p_i | \text{“}p_i \text{ is prime” and “}1 \leq i \leq \pi(\sqrt{n})\text{”}\}$,

$\{p_i.p_j\} := \{(p_i.p_j) | \text{“}p_i, p_j \text{ are prime” and “}1 \leq i < j \leq \pi(\sqrt{n})\text{”}\}$

$\{p_i.p_j.p_k\} := \{(p_i.p_j.p_k) | \text{“}p_i, p_j, p_k \text{ are prime” and “}1 \leq i < j < k \leq \pi(\sqrt{n})\text{”}\}$, ...

Then,

$$\sum_d \mu(d) = 0.$$

B. Let n be a natural number. Let $p_1, p_2, p_3, ..., p_n$ be any collection of n primes.
Let d be a square-free divisor in $\{1, \{p_i\}, \{p_i.p_j\}, \{p_i.p_j.p_k\}, ...\}$, where

$\{p_i\} := \{p_i | \text{“}p_i \text{ is prime” and “}1 \leq i \leq n\text{”}\}$,

$\{p_i.p_j\} := \{(p_i.p_j) | \text{“}p_i, p_j \text{ are prime” and “}1 \leq i < j \leq n\text{”}\}$

$\{p_i.p_j.p_k\} := \{(p_i.p_j.p_k) | \text{“}p_i, p_j, p_k \text{ are prime” and “}1 \leq i < j < k \leq n\text{”}\}$, ...

Then,

$$\sum_d \mu(d) = 0.$$

We see how any collection of n primes can form divisors d in $\{1, \{p_i\}, \{p_i.p_j\}, \{p_i.p_j.p_k\}, ...\}$ and that $\#\{p_i\}, \#\{p_i.p_j\}, \#\{p_i.p_j.p_k\}, ...$ correspond exactly to the columns (or rows) of Pascal's triangle. The "alternating sums of subsequent columns (or rows) in Pascal's triangle" is $\{1, 0, 0, 0, , ...\}$, which gives a total Möbius sum of -1. Since $\mu(1) = +1$,

$$\sum_d \mu(d) = \mu(1) - 1 = 0.$$

[26]http://oeis.org/wiki/Pascal_triangle

The Liouville function λ maps to values $+1$ or -1. If the sum of the powers of the prime components of a number n is even, $\lambda(n) = +1$, and if the sum of the powers of the prime components of n is odd, $\lambda(n) = -1$.

Consider, the set of primes $\{2, 3\}$. Write the "Pascal form" or "triangular form" of the collection of divisors of the maximal product of elements in that set, as: $\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix}$. This is the collection of all square-free products generated by $\{2, 3\}$.

$$\text{Now consider, } \begin{pmatrix} \mathbf{2}(2) & \mathbf{3}(3) \\ & \mathbf{2.3}\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2^2 & 3^2 \\ & \begin{pmatrix} 2^2.3 & 2.3^2 \\ & 2^2.3^2 \end{pmatrix} \end{pmatrix}.$$

We see how a simple product of each entry of this Pascal form (with itself or its corresponding Pascal form), helps us generate all the possible combinations of $\{2, 3\}$ of degree 2. Call the degree[27] of a number the degree of the highest-degree-prime in the number's prime factorization. For instance, the degree of $45 = 3^2.5$ is 2.

Now consider Degree 3 collections:

$$\begin{pmatrix} 2(2(2)) & 3(3(3)) \\ & 2.3\begin{pmatrix} 2(2) & 3(3) \\ & 2.3\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2^3 & 3^3 \\ & \begin{pmatrix} 2^3.3 & 2.3^3 \\ & \begin{pmatrix} 2^3.3^2 & 2^2.3^3 \\ & 2^3.3^3 \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

Degree 4 collections:

$$\begin{pmatrix} 2(2(2(2))) & 3(3(3(3))) \\ & 2.3\begin{pmatrix} 2(2(2)) & 3(3(3)) \\ & 2.3\begin{pmatrix} 2(2) & 3(3) \\ & 2.3\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 2^4 & 3^4 \\ & 2.3\begin{pmatrix} 2^3 & 3^3 \\ & 2.3\begin{pmatrix} 2^2 & 3^2 \\ & 2.3\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

[27] This type of definition has a similar sense to the "degree" of a polynomial.

$$= \begin{pmatrix} 2^4 & 3^4 \\ & \begin{pmatrix} 2^4.3 & 2.3^4 \\ & \begin{pmatrix} 2^4.3^2 & 2^2.3^4 \\ & \begin{pmatrix} 2^4.3^3 & 2^3.3^4 \\ & 2^4.3^4 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

Degree 5 collections:

$$\begin{pmatrix} 2(2(2(2(2)))) & 3(3(3(3(3)))) \\ 2.3 & \begin{pmatrix} 2(2(2(2))) & 3(3(3(3))) \\ 2.3 & \begin{pmatrix} 2(2(2)) & 3(3(3)) \\ 2.3 & \begin{pmatrix} 2(2) & 3(2) \\ 2.3 & \begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 2^5 & 3^5 \\ 2.3 & \begin{pmatrix} 2^4 & 3^4 \\ 2.3 & \begin{pmatrix} 2^3 & 3^3 \\ 2.3 & \begin{pmatrix} 2^2 & 3^2 \\ 2.3 & \begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 2^5 & 3^5 \\ & \begin{pmatrix} 2^5.3 & 2.3^5 \\ & \begin{pmatrix} 2^5.3^2 & 2^2.3^5 \\ & \begin{pmatrix} 2^5.3^3 & 2^3.3^5 \\ & \begin{pmatrix} 2^5.3^4 & 2^4.3^5 \\ & 2^5.3^5 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

If we were to replace every divisor in these collections with o for "odd", when $\lambda(d) = -1$ and e for "even", when $\lambda(d) = +1$, we have:

Degree 0 : e

Degree 1: $\begin{pmatrix} o & o \\ & e \end{pmatrix}$

Degree 2: $\begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix}$

Degree 3: $\begin{pmatrix} o & o \\ & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix} \end{pmatrix}$

Degree 4: $\begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$

and Degree 5: $\begin{pmatrix} o & o \\ & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix}$

We wish to sum the λ values over these collections:

Degree 0 is simply e or $+1$, and Degree 1, by Theorem 1, sums to o or -1. Degree 2 sums to e, Degree 3 sums to o, Degree 4 sums to e, Degree 5 sums to o, ...

Consider, now, the set of primes $\{2, 3, 5\}$. The degree 1 divisors are

$$\begin{pmatrix} 2 & 3 & 5 \\ & 2.3 & 2.5 \\ & & 3.5 \\ & & 2.3.5 \end{pmatrix}$$

,

the degree 2 divisors are: $$\begin{pmatrix} 2^2 & 3^2 & 5^2 \\ & 2.3\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} & 2.5\begin{pmatrix} 2 & 5 \\ & 2.5 \end{pmatrix} \\ & & 3.5\begin{pmatrix} 3 & 5 \\ & 3.5 \end{pmatrix} \\ & & 2.3.5\begin{pmatrix} 2 & 3 & 5 \\ & 2.3 & 2.5 \\ & & 3.5 \\ & & 2.3.5 \end{pmatrix} \end{pmatrix},$$

and the degree 3 divisors are:

$$\begin{pmatrix} 2^3 & 3^3 & 5^3 \\ & 2.3\begin{pmatrix} 2^2 & 3^2 \\ & 2.3\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} \end{pmatrix} & 2.5\begin{pmatrix} 2^2 & 5^2 \\ & 2.5\begin{pmatrix} 2 & 5 \\ & 2.5 \end{pmatrix} \end{pmatrix} \\ & & 3.5\begin{pmatrix} 3^2 & 5^2 \\ & 3.5\begin{pmatrix} 3 & 5 \\ & 3.5 \end{pmatrix} \end{pmatrix} \\ & & 2.3.5\begin{pmatrix} 2^2 & 3^2 & 5^2 \\ & 2.3\begin{pmatrix} 2 & 3 \\ & 2.3 \end{pmatrix} & 2.5\begin{pmatrix} 2 & 5 \\ & 2.5 \end{pmatrix} \\ & & 3.5\begin{pmatrix} 3 & 5 \\ & 3.5 \end{pmatrix} \\ & & 2.3.5\begin{pmatrix} 2 & 3 & 5 \\ & 2.3 & 2.5 \\ & & 3.5 \\ & & 2.3.5 \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

If we were to replace every entry in these collections with o for "odd", when $\lambda(d) = -1$ and e for "even", when $\lambda(d) = +1$, we have:

Degree 0 : e

Degree 1: $\begin{pmatrix} o & o & o \\ & e & e \\ & & e \\ & & o \end{pmatrix}$

Degree 2:
$$\begin{pmatrix} e & e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} & \begin{pmatrix} o & o \\ & e \end{pmatrix} \\ & & \begin{pmatrix} o & o \\ & e \end{pmatrix} \\ & & \begin{pmatrix} e & e & e \\ & o & o \\ & & o \\ & & e \end{pmatrix} \end{pmatrix}$$

Degree 3:
$$\begin{pmatrix} o & o & o \\ & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix} & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix} \\ & & \begin{pmatrix} e & e \\ & \begin{pmatrix} o & o \\ & e \end{pmatrix} \end{pmatrix} \\ & & \begin{pmatrix} o & o & o \\ & \begin{pmatrix} e & e \\ & o \end{pmatrix} & \begin{pmatrix} e & e \\ & o \end{pmatrix} \\ & & \begin{pmatrix} e & e \\ & o \end{pmatrix} \\ & & \begin{pmatrix} o & o & o \\ & e & e \\ & & e \\ & & o \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

We wish to sum the λ values over these collections: Degree 0 is just simply e or $+1$, and Degree 1, by Theorem 1, sums to o or -1. Degree 2 sums to e, Degree 3 sums to o, ...

The degree[28] of a number is defined as the degree of the highest-degree-prime in the number's prime factorization. We observe that the parity of the degree of a number satisfies the following theorem[29]:

THEOREM 2. *Let $\{p_i\}$ be the finite set of primes of size n, with $i \leq n$. Define the "degree of a number" k as the degree of the highest-degree-prime in the number's prime factorization. Call d_0 a degree 0 number, d_1 a degree 1 number, and d_k a degree k number, formed only by primes from $\{p_i\}$. Then,*

$$A. \sum_{d_1} \lambda(d) = -\lambda(d_0).$$

$$B. \sum_{d_k} \lambda(d) = (-1)^k.$$

(A) Begin with the finite set of primes of size n. The Degree 0 collection always consists of a single element, the number 1, which has 0 number of prime divisors, and since 0 is an even number, $\lambda(1) = +1$. The Degree 1 collection corresponds to all the square-free numbers generated by our initial set of generators, which, in this case, is the set of the first n primes. Since we know that all possible distinct combinations of a set of primes taken distinctly can be given by the Pascal or "triangular" form generated by the set, and that $\lambda(d) = \mu(d)$ for each d in the degree 0 and degree 1 collections, so by Theorem 1,

$$\sum_{d_1} \lambda(d) = -\lambda(d_0).$$

(B) Now, let us consider the collection of divisors of degree k. In our concrete demonstrations for prime generators $\{2, 3\}$ and $\{2, 3, 5\}$, we offered a recursive method to list all possible combinations of these prime generators for *any* degree. The method employed is simple. In order to construct the collection of all possible divisors of degree n generated by a set of primes, we must first write out the degree 1 collection, which is simply the Pascal or "triangular" form of the collection of square-free numbers generated by that given set of primes. Then, we multiply each entry of this representation with its corresponding Pascal form, i.e., we multiply each

[28]As a reminder, this sense of "degree" is similar in the context of polynomials

[29]The two theorems mentioned in Appendix 0 are original and are attributed to the co-authors of this paper, J. M. Flagg, Louis H. Kauffman, and Divyamaan Sahoo.

entry p_i with its Pascal form p_i, we multiply each entry $p_i.p_j$ with $\begin{pmatrix} p_i & p_j \\ & p_i.p_j \end{pmatrix}$, we multiply $p_i.p_j.p_k$ with $\begin{pmatrix} p_i & p_j & p_k \\ & p_i.p_j & p_i.p_k \\ & & p_j.p_k \\ & & p_i.p_j.p_k \end{pmatrix}$, and so on.

This recursive method, applied *once*, gives us all possible degree 2 numbers whose prime components are restricted by the starting set of primes of size n, i.e., the degree 2 collection. When this recursive method is applied *twice*, we get the degree 3 collection, and when applied $(k-1)$ *times*, we get the degree k collection.

Note that the degree k collection will have k containers[30] by the end of this recursive method, i.e., the depth of degree k arrangement is k. Further, by design, we see that a degree k collection is essentially a degree 1 collection of degree $(k-1)$ collections, since each entry of the degree 1 collection combines with its corresponding Pascal form and begins nesting within itself $(k-1)$ times in order to exhaustively list the entire degree k collection. In essence, a degree k collection is a degree 1 colljection, and we know, by Theorem 2.A, $\sum_{d_1} \lambda(d) = -1$, regardless of the parity of n.

Starting from the innermost depth, we will be able to replace degree 1 collections with $\sum_{d_1} \lambda(d) = -1$, since we wish to sum all the λ values in the degree k collection. As we move inside-out, we subsequently remove containers and simplify the arrangement, and are finally left with a degree 1 arrangement with n entries in the first row, namely, the single primes (from the finite set of primes of size n) raised to the k^{th} power. Now consider, n can be even or odd, but the degree k collection will always begin with a row of n numbers whose λ values are all $+1$, if k is even and -1, if k is odd. From our knowledge of the Pascal triangle, only one entry of the first row will survive the alternate sign summation of the columns of Pascal's triangle, which contributes λ of $+1$ if k is even and -1 if k is odd. Hence,

$$\sum_{d_k} \lambda(d) = (-1)^k$$

[30]The reader is encouraged to make note of this in the concrete demonstrations provided for example and perform investigations for large n on a large piece of paper.

Bibliography

[1] Abel, N. H. *Solution de quelques problemes a l'aide d'integrales définies.* Mug. Nuturvidemkuberne 2, 1823.

[2] Apostol, T. M. *Introduction to Analytic Number Theory* Springer, New York, 1976.

[3] Bombieri, E. *Problems of the Millennium: The Riemann Hypothesis*, 2000. Retrieved October 5, 2020, from https://www.claymath.org/sites/default/files/official˙problem˙description.pdf

[4] Borwein, P. Choi, S. Rooney, B. Weirathmueller, A. *The Riemann Hypothesis: A Resource for the Afficionado and Virtuoso Alike.* Canadian Matematical Society. Springer, New York, 2006.

[5] Borwein, P. Ferguson, R. Mossinghoff, M. J. *Sign changes in sums of the Liouville function.* Math. Comp.77 (2008), 1681-1694.

[6] Caldwell, J. H. "The distribution of prime numbers", *Topics in Recreational Mathematics*, Cambridge University Press, 1966.

[7] Chebyshev, P. L. *Sur la fonction qui détermine la totalité des nombres premiers inferieurs à une limite donnée.* J. Math. Pures Appl. [1] **17**, 1852.

[8] Conrey, J. B. *The Riemann Hypothesis*, Notices of the AMS, March 2003.

[9] Denjoy, A. *L'Hypothese de Riemann sur la distribution des zéros de* $\zeta(s)$*, reliéea la théorie des probabilités.* Comptes Rendus Acad. Sci. Paris, 1931.

[10] Derbyshire, J. *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*, Joseph Henry Press, 2003.

[11] Edwards, H. M. *Riemann's Zeta Function.* Academic Press, New York, 1974.

[12] Eisner, J. *Computation of* $\pi(n)$ *by Sieving, Primality Testing, Legendre's Formula and Meissel's Formula.* Spring, 1993.

[13] Elliott, J. *Asymptotic expansions of the prime counting function.* arXiv:1809.06633v4. 27, July 2019.

[14] Erdös, P. *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem.* PNAS 35 (7) 374-384; July 1, 1949 Retrieved October 6, 2020 from https://doi.org/10.1073/pnas.35.7.374

[15] Eswaran, K. *The Dirichlet Series for the Liouville Function and the Riemann Hypothesis.* arXiv:1609.06971v10. 8 October, 2017.

[16] Euler, L. *Variae observationes circa series infinitas*, 1737. Retrieved October 5, 2020, from http://eulerarchive.maa.org/backup/E072.html

[17] Euler, L. *De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt.* Commentarii academiae scientiarum Petropolitanae, Volume 5, pp. 36-57, 1730.

[18] Gauss, C. F. *Disquisitiones generales circa seriem infinitam*, Commentationes Societatis Regiae Scientiarum Gottingensis, Tom. II, Classis mathematicae, 1812.

[19] Gauss, C. F. *Letter to Enke dated 24 December 1849.* “Werke,” Vol. II, pp.444-447. Königlichen Gesellschaft der Wissenschaften zu Göttingen.

[20] Goldfield, D. *The Elementary Proof of the Prime Number Theorem: An Historical Perspective*, 2003. Retrieved October 6, 2020 from https://www.math.columbia.edu/ goldfeld/ErdosSelbergDispute.pdf

[21] Hadamard, J. *Étude sur les Propriétés des Fonctions Entières et en Particulier d'une Fonction Considérée par Riemann.* Journal de Mathématiques Pures et Appliquées, (pp. 171-216), 1893.

[22] Hadamard, J. *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques* Bulletin de la Société Mathématique de France, Volume 24 (p. 199-220), 1896.

[23] Halberstam, H. Richert, H. E. *Sieve Methods.* Academic Press, London. 1974.

[24] Hardy, G. H. Heilbronn, H. “Edmund Landau” *Obituary.* Reproduced from J. London Math. Soc. 13 (1938) 302-310 with the permission of the London Mathematical Society. Retrieved October 6, 2020 from http://www.numbertheory.org/obituaries/LMS/landau/

[25] Havil, J. *Gamma: Exploring Euler's Constant*, Princeton University Press, 2003.

[26] Hayes, B. *The Spectrum Of Riemannium* pp. 296-300, American Scientist, Volume 91, Number 4, July - August 2003.

[27] Jacobi, C. G. J. *Fundamenta Nova Theoriae Functionum Ellipticarum*, “Suites des notices sur les fonctions elliptiques”. J. Reine. Angew. Math. 3, 303-310, 1828.

[28] Landau, E. *Handbuch der Lehre von der Verteilung der Primzahlen.* Zweiter Band. Leipzig und Berlin. 1909.

[29] Landau, E. *Neuer Beweis der Gleichung* $\sum_{k=1}^{\infty} \frac{\mu(k)}{k} = 0$. Inaugural-Dissertation, Berlin, 1899.

[30] Legendre, A. M. *Essai sur la théorie des nombres*, Duprat, 1797-1798.

[31] Lehman, R. S. *On Liouville's Function*, 1960. Retrieved October 6, 2020 from American Mathematical Society (AMS)

[32] Littlewood, J. E. *Quelques conséquences de l'hypothèse que la fonction $\zeta(s)$ n'a pas de zéros dans le demi-plan Re $(s) > \frac{1}{2}$.* C.R. Acad. Sci. Paris **154**, 263-266, 1912.

[33] von Mangoldt, H. *Zur Verteilung der Nullstellen der Riemannschen Funktion $\xi(t)$* Math. Ann. 60, 1-19, 1905.

[34] Mazur, B. Stein, W. *Prime numbers and the Riemann hypothesis.* Cambridge University Press, 2016.

[35] Möbius, A. F. *Uber ein besondere Art von Umkehrung der Reihen*, Crelle's Journal d. M. Bd. IX. Hft. 2., 1832.

[36] n-Category, Café. *The Riemann Hypothesis.* Part 1. Retrieved August 2, 2020, from https://golem.ph.utexas.edu/category/2019/09/the˙riemann˙hypothesis˙part˙1.html.

[37] n-Category, Café. *The Riemann Hypothesis.* Part 2. Retrieved August 2, 2020, from https://golem.ph.utexas.edu/category/2019/09/the˙riemann˙hypothesis˙part˙2.html.

[38] n-Category, Café. *The Riemann Hypothesis.* Part 3. Retrieved August 2, 2020, from https://golem.ph.utexas.edu/category/2019/09/the˙riemann˙hypothesis˙part˙3.html.

[39] Nagasawa, T. Akiyama, K. *On Legendre's formula and distribution of prime numbers.*

[40] Polyà, G. *Verschiedene Bemerkungen zur Zahlentheorie*, Jahresbericht der deutschen Math-Vereinigung, Vol. 28 (pp. 31-40), 1919.

[41] Riemann, B. Hattendorf, K. (ed.) *Partielle Differentialgleichungen und ihre Anwendungen auf physikalische Fragen.* Vieweg, Braunschweig, 1876.

[42] Riemann, B. (transl. Wilkins, 1998) *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsberichte der Berliner Akademie, November 1859.

[43] Riemann, B. (transl. Baker, R. Christenson, C. Orde, H.), *Collected Papers* from *Gesammelte Mathematische Werke'*, 2nd edn., edited by Heinrich Weber, Teubner, Leipzig, 1892.

[44] Sarnak, P. *Problems of the Millennium: The Riemann Hypothesis*, 2004. Retrieved October 5, 2020, from https://www.claymath.org/sites/default/files/sarnak˙rh˙0.pdf

[45] Selberg, A. *An Elementary Proof of the Prime-Number Theorem*. Annals of Mathematics, Second Series, Vol. 50, No. 2 (pp. 305-313), Apr., 1949.

[46] Sharon, A. *Analytic Continuation of $\zeta(s)$ Violates the Law of Non-Contradiction (LNC).* arXiv:1802.08062v7, July 30, 2019.

[47] Spencer-Brown, G. *Probability and Scientific Inference*. Longmans, Green and Co., London, 1957.

[48] Spencer-Brown, G. "Appendix 9. A proof of Riemann's hypothesis via Denjoy's equivalent theorem." *Laws of Form*, Bohmeier Verlag, Germany, 2015.

[49] Titchmarsh, E.C. Heath-Brown, D.R.(ed) "The Theory of the Riemann Zeta-Function". Oxford University Press, New York, 1988.

[50] Tou, E. R. *Math Origins: Orders of Growth*. Retrieved on October 6, 2020 from MAA Publications, Periodicals, "Convergence", January 2018.

[51] de la Vallée Poussin, C. J. *Recherches analytiques sur la théorie des nombres premiers*. Bruxelles. Hayez, Imprimeur de L'Académie Royale de Belgique, 1897.

[52] Veisdal, J. *Prime numbers and the Riemann zeta function*. Trondheim. 09.05.13

[53] Voronin, S. M. Karatsuba, A. A. *The Riemann Zeta-Function*. Walter de Gruyter, May 3, 2011.

Acknowledgement: Kauffman's work was supported by the Laboratory of Topology and Dynamics, Novosibirsk State University (Contract no. 14.Y26.31.0025 with the Ministry of Education and Science of the Russian Federation).