A criterion arising from explorations pertaining to the Oesterle-Masser conjecture

Chris Goddard
May 6, 2017

Abstract

Using an extension of the idea of the radical of a number, as well as a few other ideas, it is indicated as to why one might expect the Oesterle-Masser conjecture to be true. Based on structural elements arising from this proof, a criterion is then developed and shown to be potentially sufficient to resolve two relatively deep conjectures about the structure of the prime numbers. A sketch is consequently provided as to how it might be possible to demonstrate this criterion, borrowing ideas from information theory and cybernetics.

1 Forward

1.1 A generalised zeta function, and some notions regarding radicals

Definition 1. Define

$$\zeta_a^{(2)}(z,u) := \sum a_{nm} n^{-z^2} m^{-u^2} (n+m)^{-uz}$$

as the generalised Riemann Zeta function, where n, m are taken over the natural numbers.

Remark. As an interesting curiosity, associated broadly within the remit of M-theory, much as the previous can be associated with concepts from L-theory, note that we can extend this further:

$$\zeta_a^{(3)}(u,v,w) := \sum a_{nmp} n^{-u^3} m^{-v^3} p^{-w^3} (nmp)^{-uvw} (nm+p)^{-(u+v)w} (mp+n)^{-(v+w)u} (pn+m)^{-(w+u)v} (n+m+p)^{-uvw}$$

The following is a standard definition.

Definition 2. (Radical). The radical of an integer is the product of its prime factors without multiplicity.

We now would like to extend this idea to the real numbers. We do this by observing that the radical of an integer can be equivalently computed by viewing it as an ideal in the ring of integers, and then computing its corresponding radical, as the intersection of its prime ideals.

Now consider the mapping between the ideals of the reals and the real numbers as $K : Ideals(R) \to R$. But ideals under multiplication in the reals encompass all the reals, so these are trivial. So in order to proceed, we must take a slightly different approach.

So let us consider instead the rational numbers. Suppose that $\frac{p}{q}$ is the reduced representation of a rational, k, such that p and q are coprime. Then we might, for instance, define the radical of k to be the tuple $(rad(p), rad(q))$.

To extend this general premise to the real numbers, we consider the continued fraction expansion for a real number, a. Then this will be of the form $(a_0, a_1, ..., a_n, ...)$ where the a_i are integers.

Consequently we can now define a sensible notion for the radical of a real number.

Definition 3. (Radical of a real number). Suppose that a is a real number with continued fraction expansion $\{a_i\}_{i \in N}$. Then its associated *radical sequence* is given by the tuple $\{rad(a_i)\}_{i \in N}$. Define then the radical of a to be the real number represented as a continued fraction expansion by its radical sequence.

Note that $rad(rad(x)rad(y)) = rad(xy)$. This follows when we observe that any term in the product of $x = (u_0, ..., u_k, ...)$ and $y = (v_0, .., v_k, ...)$ will be of the form $\sum a_{ijk} u_i v_j$ on the left hand side, and $\sum b_{ijk} rad(u_i) rad(v_j)$ on the right hand side, i.e.

$xy = (\Sigma a_{ij0} u_i v_j, \Sigma a_{ij1} u_i v_j, ...)$ and $rad(x)rad(y) = (\Sigma b_{ij0} rad(u_i) rad(v_j), ...)$.

Then suppose by an inductive hypothesis we have that $rad(\Sigma a_{ijN} u_i v_j) = rad(\Sigma b_{ijN} rad(u_i) rad(v_j))$. Then I claim that it is possible to extend this to $N + 1$ if the a_{ijk} and b_{ijk} are in reduced form, i.e., no extraneous factors - and using the property of the radical over the integers.

Theorem 1.1. *In particular, we have that this is the unique function that has the properties that 1) that rad(n) is the radical of n for n an integer, and 2) that*

$$rad(rad(x)rad(y)) = rad(xy),\ rad(x rad(x)) = rad(x),\ \text{and}\ rad(rad(x)) = rad(x)$$

Proof. Existence is clear from the above. We prove uniqueness.

Suppose we had a continuous function $P: R \to R$ such that $P(n) = rad(n)$ for n an integer, and such that $P(P(x)P(y)) = P(xy)$ for all x, y in R. Suppose that Q is another such function. If P and Q are not the same function, there is a t for which $P(t) \neq Q(t)$.

Note furthermore that $[P, Q] = 0$, since by minor abuse of pre-image notation, $PPQ = PQQ$ or $PQ = QQ$, and also $PQQ = PPQ$, or $PQ = PP$.

Similarly $QP = QQ$, hence $[P, Q] = 0$.

We would like to construct an integer T such that $P(T) \neq Q(T)$, contradiction since P, Q should match on all integers.

Certainly $1 = P(1) = P(P(t)P(1/t))$, and $1 = Q(1) = Q(Q(t)Q(1/t))$. By symmetry it is clear that $Q(1/t) \neq P(1/t)$. Also $P(tP(t)) \neq Q(tP(t))$, and $P(tQ(t)) \neq Q(tQ(t))$.

This follows since $P(tP(t)) = P(P(t)P(P(t))) = P(P(t)P(t)) = P(t)$, so if $P(t) = P(tP(t)) = Q(tP(t))$, then $PP(t) = PQ(tP(t)) = QP(tP(t)) = QP(t)$. Hence then $P(P(t)) = Q(P(t)) = PQ(t)$. Then taking the pre-image of both sides we have that $P(t) = Q(t)$, contradiction.

We hence have two operations on our original number t we can perform: we can multiply by $P(t)$, or we can invert the number and calculate the reciprocal. This generates a set of numbers S that has the property that $P(s) \neq Q(s)$ for all s in this set. ie.

$$S = \{s | P(s) \neq Q(s)\}$$

and we have that if $s \in S$, then $P(t)s \in S$ and $s^{-1} \in S$.

Therefore we can construct a sequence $\{s_i\}$ such that s_i converges to some integer T.

But then by continuity we have that $P(T) = limP(s_i) \neq limQ(s_i) = Q(T)$, contradiction.

□

Note that the mapping $rad: R \to R$ is smooth, as this trivially follows from the convergence properties of the continued fraction expansion to a real number - i.e., as one gets closer and closer to the number in question, there will be more and more terms of the expansion in a neighbourhood about it that become fixed with increasing proximity.

1.2 Extending the Euler product formula to pseudo-primes

Theorem 1.2. *Let*

$$S(x) := \{(a,b)|a+b=x, gcd(a,b)=1 \text{ and } rad(ab(a+b)) < x\}$$

where all numbers are understood to be naturals.
With

$$\zeta^{(2)}(z,u) := \frac{1}{2^{-uz}} \sum n^{-u^2} m^{-z^2} (n+m)^{-uz}$$

we have that $\zeta^{(2)}$ *satisfies the identity*

$$\zeta^{(2)}(z,u) = \Pi_{p\,prime} \Pi_{(a,b)\in S(p)} (1 - {}_2F_1 \left[\begin{matrix} B(a,u), B(b,z) \\ B(a+b, u+z) \end{matrix} ; -(u+z) \right])^{-1} \quad (1)$$

$$= \Pi_{x\in N} \Pi_{(a,b)\in S(x)} (1 - \circ(\Gamma; u+z)\{B(a,b)^{u+z}\})^{-1} \quad (2)$$

where $B(a,b)$ *is the* a^{th} *Bernoulli polynomial evaluated at* b *and* $\circ(\Gamma; u+z)$*, the iterated composition of* Γ*, can be computed as* $I_\circ(\star(\sigma;\tau))$*, with* $I_\circ$ *the information for a particular type of exotic geometry [ref Go2, pp ...], with* $\Gamma = I(\sigma)$*, and* $u+z = I(\tau)$ *for metrics* σ*,* τ *and the fisher information* I *over a Riemannian manifold.*

Remark. (Aside). For tetration, $\wedge(f;g) := f \wedge f \wedge \cdots \wedge f$ (g times), we have that $\wedge(f;g) = I_\wedge(\star(\sigma;\tau))$. Here $I_\wedge$ is another information for a different exotic geometry [ref Go2, pp ...].

Remark. (Euler's Product Formula). Note that this is in a way a natural extension / generalisation of an analogous identity for $\zeta^{(1)}(z) := \sum n^{-z}$:

$$\zeta^{(1)}(z) = \Pi_{p\in\mathcal{P}} (1 - p^{-z})^{-1}$$

where $\mathcal{P}$ is the set of primes.

Remark. This also relates to the formula for the Selberg zeta function on page 28 of Daniel Bump's notes [ref] regarding the Selberg trace formula:

$$Z(s) = \Pi_{\{N_0\}} \Pi_{k=0}^{\infty} (1 - N_0^{-s-k})$$

where the first product is indexed by the geodesics γ over the complex numbers that are prime with respect to some natural choice of hyperbolic group, so that in essence the numbers traced by each geodesic form prime ideals with respect to the global group. $N_0(\gamma)$ is the length of said geodesic.

Proof. (sketch).

The general idea is to note that

$$\Pi_{p\,prime}\Pi_{(a,b)\in S(p)}(1 - {}_2F_1\left[\begin{matrix} B(a,u), B(b,z) \\ B(a+b,u+z) \end{matrix}; -(u+z)\right])^{-1} \tag{3}$$

expands to a sum of expressions of the form

$${}_2F_1\left[\begin{matrix} B(a_{p_1},u), B(b_{p_1},z) \\ B(p_1,u+z) \end{matrix}; -(u+z)\right]^{m_1} \times ... \times {}_2F_1\left[\begin{matrix} B(a_{p_n},u), B(b_{p_n},z) \\ B(p_n,u+z) \end{matrix}; -(u+z)\right]^{m_n} \tag{4}$$

If we write

$$\Pi_i\{\Pi_{(a,b)\in S(a_{p_i})}\{\Gamma(a+b)B(a,b)\}^{m_i}\} = n,\ and \tag{5}$$

$$\Pi_i\{\Pi_{(a,b)\in S(b_{p_i})}\{\Gamma(a+b)B(a,b)\}^{m_i}\} = m \tag{6}$$

where $(a_{p_i}, b_{p_i}) \in S(p_i)$, then

$$\Pi_i\{\Pi_{(a,b)\in S(a_{p_i}+b_{p_i}=p_i)}\{\Gamma(a+b)B(a,b)\}^{m_i}\} = n+m.$$

Furthermore, since $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, these are all integers.

I then claim that the above product reduces to an expression of the form

$$n^{-u^2}m^{-z^2}(n+m)^{-uz}$$

with n, m defined as above.

Using the nature of the set $S(p)$, I claim moreover that this factorisation of the above expression is unique, and always exists for any $(n,m) \in N \times N$. We are then done.

□

1.3 Estimating the size of $|S(x)|$

We are now ready to state the main claim of this section.

Theorem 1.3. *(Growth of the size of S(x) with x). Let x be a natural number. Then*

$$|S(x)| \sim \frac{ln(x)}{|ln({}_2F_1\left[\begin{matrix} B(x/2,x/2),B(x/2,x/2) \\ B(x,x) \end{matrix}; -x\right])|} \; as\; x \to \infty. \tag{7}$$

In particular, $lim_{x\to\infty}|S(x)| = 0$, since if x is large, ${}_2F_1\left[\begin{matrix} B(x/2\ x/2)\ B(x/2\ x/2) \\ B(x\ x) \end{matrix}; -x\right]$ is very small, and therefore $ln({}_2F_1\left[\begin{matrix} B(x/2\ x/2)\ B(x/2\ x/2) \\ B(x\ x) \end{matrix}; -x\right])$ is very large and negative.

Remark. This can be viewed as a form of generalisation of a well known result regarding the size of the prime counting function, originally a conjecture due to Gauss and Legendre, and demonstrated by Hadamard and Poisson. This can be stated as follows:

"Denote the number of primes less than n as $\pi(x)$. Then $\pi(x) \sim \frac{x}{ln(x)}$."

Proof. (Sketch). The proof is similar to that for the prime number theorem above, but instead of using Euler's identity for the zeta function, we use the extension of this identity demonstrated in the previous theorem for $\zeta^{(2)}$, together with our lift of the concept of the radical of a natural number to the complex line - which was defined at the start of this paper.

I follow Newman's short proof of the prime number theorem.

The key quantities we are interested in are:

$$\zeta^{(2)}(u,z) = \Sigma_{n,m} n^{-u^2} m^{-z^2} (n+m)^{-uz}, \tag{8}$$

$$\Phi(u,z) = \Sigma_{p,q} ln({}_2F_1\left[\begin{matrix} B(p,u), B(q,z) \\ B(p+q, u+z) \end{matrix}; -(u+z)\right]) p^{-u^2} q^{-z^2} (p+q)^{-uz}, \tag{9}$$

$$\Theta(x) = \Sigma_{p\le x} \Sigma_{(a,b)\in S(p)} ln({}_2F_1\left[\begin{matrix} B(a), B(b) \\ B(a+b) \end{matrix}; -p\right]) \tag{10}$$

The goal of the first part of the proof will be to demonstrate that $\Theta(x) \sim x$.

Lemma 1.4. *(Euler's product formula for the zeta function $\zeta^{(2)}$) As above.*

Lemma 1.5. *(Holomorphicity of $\zeta^{(2)}$ minus correction term)*

Proof. Uses an argument similar to that in Newman's short proof of the prime number theorem.

□

Lemma 1.6. *(First estimate of Θ)* $\Theta(x) \sim O(x)$.

Proof. Follows from the fact that

$$\Sigma_{(a,b)\in S(p)} ln({}_2F_1\left[\begin{matrix} B(a) \; B(b) \\ B(a+b) \end{matrix}; -p\right]) \sim ln(p)$$

Then this lemma follows from the equivalent lemma in Newman's short proof of the prime number theorem.

□

Lemma 1.7. *(Holomorphicity of Φ minus correction term) $\Phi(u,z)$ minus correction term is holomorphic for an appropriate u, z region (eg $Re(u+z) \geq 1$). Also $\zeta^{(2)}$ is nonzero in this interval.*

Proof. Uses 1.5, 1.4 and an argument similar to Newman's short proof of the prime number theorem.

□

Lemma 1.8. *(Convergence of an integral involving Θ) $\int_1^\infty \frac{\Theta(x)-x}{x^2}dx$ is a convergent integral.*

Proof. Follows from the fact that Θ here is the same theta as in Newman's short proof of the prime number theorem per 1.6. We also need to use 1.7 to ensure that the integral exists.

□

Lemma 1.9. *(Sharpened estimate of Θ) $\Theta(x) \sim x$.*

Proof. Follows from 1.8 and Newman's short proof of the prime number theorem.

□

Now we can use our sharpened estimate of Θ to understand $|S(x)|$. In particular,

$$\Theta(x) \leq \Pi(x)|S(x)||ln({}_2F_1\left[\begin{matrix} B(x/2\ x/2)\ B(x/2\ x/2) \\ B(x\ x) \end{matrix}; -x\right])|$$

So, if $\Theta(x) \sim \frac{x}{\Pi(x)}$, then $|S(x)| \leq \frac{ln(x)}{|ln({}_2F_1\left[\begin{matrix} B(x/2,x/2),B(x/2,x/2) \\ B(x,x) \end{matrix}; -x\right])|}$ as to first order we know that $\Pi(x) \sim \frac{x}{ln(x)}$.

The inequality also extends in the other direction, since for any $\epsilon > 0$,

$$\Theta(x) \geq \Sigma_{x^{(1-\epsilon)} \leq p \leq x x} \Sigma_{(a,b) \in S(p)} ln({}_2F_1\left[\begin{matrix} B(a), B(b) \\ B(a+b) \end{matrix}; -p\right]) \tag{11}$$

$$\geq \Sigma_{x^{(1-\epsilon)} \leq p \leq x} \Sigma_{(a,b) \in S(p)} (1-\epsilon)|ln({}_2F_1\left[\begin{matrix} B(x/2, x/2), B(x/2, x/2) \\ B(x,x) \end{matrix}; -x\right])| \tag{12}$$

$$= (1-\epsilon)|ln({}_2F_1\left[\begin{matrix} B(x/2, x/2), B(x/2, x/2) \\ B(x,x) \end{matrix}; -x\right])|[\Pi(x)|S(x)| + O(x^{1-\epsilon})] \tag{13}$$

Hence by elementary application of the sandwich inequality theorem we are done.

□

Corollary 1.10. *(Oesterle-Masser conjecture). There are finitely many naturals a, b, and c such that a, b coprime satisfy $a + b = c$, and $c > rad(abc)$.*

Proof. We are done if we can demonstrate that $\Sigma_{n \in N}|S(n)|$ is finite. But this clearly follows from the estimate above.

□

1.4 Estimating the first order correction term for $|S(x)|$

Now things start to get interesting. In particular, we might ask ourselves in an analogous manner to the question that Riemann posed in his original paper [10] (regarding the prime counting function $\pi(x)$), what is the structure of the first order correction term in estimating the size of $S(x)$ with x? And this leads us to posit the following conjecture, based on structure hinted at in the resolution of the previous proof:

Conjecture 1.11. *(Generalised Zeta Criterion).* $\zeta_a^{(2)}(z,u)$ *has no zeroes on the critical plane* $2(z+u)=1$.

Corollary 1.12. *(GC). The Goldbach conjecture follows from the Generalised Zeta Criterion.*

Proof. (Sketch). Given the veracity of this statement, it is incumbent upon us to demonstrate that there is no even number n such that there does not exist a pair of primes p, q, with p + q = n (the Goldbach conjecture) if we wish to prove this claim.

We demonstrate this by recalling the Hardy Littlewood circle method.

Let $S(x,\alpha) = \Sigma_{p\leq x} exp(\alpha p)$, where the sum is understood to be over primes p less than x.

Let $f_2(x)$ be the number of ways to add n primes to get x.

Then $f_n(x) = \oint_{S^1} S(x,\alpha)^n exp(\alpha x) d\alpha$. If this is greater than zero for $n=2$, we have demonstrated the Goldbach conjecture. Assuming GRH, Hardy and Littlewood were able to demonstrate this for $n=3$. Later, Vinogradov was able to relax this requirement on the GRH. I would contend, however, that one still requires an inequality based on an information.

So, let $S(x,y|\mathbf{A}) = \Sigma_{(p,q)=\underline{p}\leq(x,y)=\underline{x}} exp(\underline{p}\mathbf{A}\underline{p}^T)$, where the sum is over primes $p\leq x$ and $q\leq y$.

We have that $f(\underline{x}) = \oint_{\underline{\tau}\in\mathbf{T}} \oint_{\mathbf{A}\in\Sigma(\mathbf{T})} S(\underline{x}|\mathbf{A}(\underline{\tau})) exp(-\underline{x}\mathbf{A}(\underline{\tau})\underline{x}^T) d\mathbf{A}(\tau) d\underline{\tau}$ is the number of ways to add two primes together to get $x+y$, where $\mathbf{T}$ is the Torus and $\Sigma(\mathbf{T})$ is the set of analytic metrics on said space. Note then that if we can prove this is positive in general, the Goldbach conjecture follows as a corollary, by setting $x=y$.

Also intriguingly note that $S(\underline{x}|\mathbf{A})$ is very similar to $\zeta^{(2)}(\underline{x})$, since both have sums that are quadratic in $\underline{x}$.

Therefore, there are two naive pieces of intuition we could use. The first is that $f(x)$ could be the information of a symmetric or antisymmetric $\mathbf{A}$. Or alternatively it could be the information of a subset of the reals, wherein then it would suffice to show that the complement is also an information.

But neither of these gives us what we need, which is to demonstrate that $f(\underline{x})$ is positive. Well, certainly, by the Cramer-Rao inequality, if we have an information, we know that it is greater than or equal to zero. So, consequently, if we can show that this is somehow a 'noise' of a higher order information on top of a base information, with the base optimised by the higher order not, then we will have that the inequality is not critical and we will be done.

But this should follow if we can demonstrate that, assuming criticality of the information associated to $\zeta^{(2)}(\underline{x})$, that the information corresponding to $S(\underline{x}|\mathbf{A})$ will not be critical. But it is clear that these are mutually exclusive conditions. We know that the former is true since the primes are a critical set (and such will be covered in more detail later).

We also know that $f(\underline{x}) \geq 0$.

So it remains to construct a functional such that we have $f(\underline{x})$ is an information of a biased estimator, since then we have that it will always be positive definite.

I would argue that whereas the signal function for $\zeta^{(2)}$ is of the form

$$\hat{f}(m,n,a) = \int_{Cyb(M)} F(m,n,b)\delta(\kappa(m,n,b)-a)db$$

we have for f that it is of the form

$$g(m,n,a) = \mathcal{F}(\int_{U\subset Cyb(M)} F(m,n,b)\delta(\kappa(m,n,b)-a)db)$$

for some set U, $\mathcal{F}$ the Fourier transform operator, since S is essentially the same of $\zeta^{(2)}$, but has missing information (it is a sum over primes, rather than over naturals). To understand this more clearly, note that $x_i A_{ij} x_j^T \sim a_{nm} n^{-z^2} m^{-u^2} (n+m)^{-uz}$ with $x_i = (u,z)$. In other words, f is almost a distorted discrete Fourier transform with missing information of $\zeta^{(2)}$.

But note that if $\hat{f}$ is the signal function for $\zeta^{(2)}$, then $\mathcal{F}(\zeta^{(2)})$ will have a signal function $g = \mathcal{F}(\hat{f})$ which will also form an information. If this is taken over a restricted set, then the estimator will be biased. I moreover claim that $f = I(\mathcal{F}(\hat{f})|_{primes})$

So since $I(\hat{f}) \geq 0$ and $I(g) \geq 0$, and $f \neq g$ a.e., then $I(\hat{f}) = 0$ implies $I(g) > 0$, QED.

□

Corollary 1.13. *(TPC). The twin prime conjecture follows from the Generalised Zeta Criterion.*

Proof. (Sketch). We seek to demonstrate that, if the above holds, that we can establish that for every prime p there will always exist a prime q with $p = q+2$, or $p = q-2$ (in other words, that the twin prime conjecture holds).

Consider as before the relation

$$f(\underline{x}) = \oint_{\underline{\tau}\in\mathbf{T}} \oint_{\mathbf{A}\in\Sigma(\mathbf{T})} S(\underline{x}|\mathbf{A}(\underline{\tau}))exp(-\underline{x}\mathbf{A}(\underline{\tau})\underline{x}^T)d\mathbf{A}(\tau)d\underline{\tau}$$

being the number of ways to add two primes p and q to equal $x+y$, for any natural numbers x, y. From the previous result, we know that this is positive.

Let

$$\hat{S}(x,\alpha) = \Sigma_{p\leq x} exp(p\alpha),$$

where the sum is over primes. Following Tao, we know that

$$\hat{f}_2(x) = \oint_{\alpha\in S^1} |\hat{S}(x,\alpha)|^2 exp(-2\alpha)d\alpha$$

will give the number of ways to represent 2 as the difference $p_1 - p_2$ of two primes with $p_1, p_2 \leq x$, and that if we can prove that this is positive definite for all $x \geq 2$, we will have demonstrated TPC.

Suppose then that we consider once again $f(\underline{x})$, but restrict to metrics $\mathbf{A}$ on $\mathbf{T}$ such that $S((x,x)|\mathbf{A}(\tau)) \geq 0$ for all $\tau \in \mathbf{T}$. Call this set $\Sigma_{\geq}(\mathbf{T}) \subset \Sigma(\mathbf{T})$. But these sets are the same if $\mathbf{A}$ is Kähler, which is naturally what we would study (note that we can view $(x,y)\mathbf{A}(x,y)^T := \langle x,y\rangle_{\mathbf{A}(\tau)}$ as an inner product of x and y with respect to $\mathbf{A}$).

Then we have an integral relation

$$f_{\geq 0}(\underline{x}) = \oint_{\underline{\tau}\in\mathbf{T}} \oint_{\mathbf{A}\in\Sigma(\mathbf{T})} S(\underline{x}|\mathbf{A}(\underline{\tau}))exp(-tr\mathbf{A}(\underline{\tau}))d\mathbf{A}(\tau)d\underline{\tau}$$

By similar arguments to before, we can conclude that this dominates an information, and hence must be positive definite. It is moreover clear that $\hat{f}_2(x)$ is positive definite follows as a corollary, since $dim(\mathbf{A}) = 2$, QED.

Note that the integrand of $f(\underline{x})$, $S(\underline{x}|\mathbf{A}(\underline{\tau}))exp(-\langle x,y\rangle_{\mathbf{A}(\underline{\tau})})$, is quite analogous to the Perelman entropy $(R_\sigma + (\nabla_\sigma F)^2)exp(-F)$, with $F = \langle x,y\rangle_{\mathbf{A}(\underline{\tau})}$.

Indeed, per [13], we have that, for

$$\begin{cases} \frac{\partial}{\partial t} g = 2Ric \\ \frac{\partial}{\partial t} u + \Delta u = 0 \end{cases}$$

then if $X_t(x)$ is a $g(t)$-Brownian motion on M starting from x, and $p(t,x,y)$ is the density of $X_t(x)$ with respect to $vol_{g(t)}$, then for

$$m_t(dy) := p(t,x,y)vol_{g(t)}(dy) = P\{X_t(x) \in dy\}$$

being the heat kernel measure, we have that

$$\int_M u(t,y)m_t(dy) = \int_M u(t,y)p(t,x,y)vol_{g(t)}(dy) = E[u(t,X_t(x))]$$

is constant under the flow, since $u(t, X_t(x))$ is a martingale.
We have moreover that the entropy of $\mu_t := u(t,.)dm_t = u(t,X_t(x))dP$ is

$$\varepsilon(t) = \int_M (uln(u))(t,y)p(t,x,y)vol_{g(t)}(dy) \tag{14}$$
$$= E[(uln(u))(t,X_t(x))] \tag{15}$$

Finally we have that the first derivative of the entropy is

$$\varepsilon'(t) = E[(|\nabla ln(u)|^2 u)(t,X_t(x))] \tag{16}$$
$$= E[\mathcal{I}(u)(t,X_t(x))] \tag{17}$$

Moreover, if we denote $E_{t,x} := E_{P_t(x)}$ as the expectation wrt $P_t(x)$, we have that

$$\varepsilon'(t) = E_{t,x}[((R + |\nabla ln(u)|^2)u)(t,X_t(x))] \tag{18}$$

where $\mathcal{I}(u)$ is the Fisher information of u.

Suppose $\varepsilon'(t) = 0$ as it will be when a point of stability is reached (the entropy no longer changes). We know, for instance, that this must be true for the primes. Then we have that

$$\int_M Rum_t(dy) = -\int |\nabla ln(u)|^2 um_t(dy) \tag{19}$$

If we set $u(\tau|x,y) = exp(-\langle x,y\rangle_{\mathbf{A}(\tau)})$, then if we can demonstrate that $S(x,y|\mathbf{A}(\tau))$ can be a curvature of a metric then we are done. But this is quite easy to see.

Hence we roughly have that $f_{\geq 0}(\underline{x})$, or at least its negative, is an Information - as is $f(\underline{x})$.

$\square$

Corollary 1.14. *(RH). If the Generalised Zeta Criterion is true, the Riemann Hypothesis follows.*

Proof. Suppose that $u = 0$. Then $\zeta_a^{(2)}(z,0) = \zeta_a(z^2)$ and the previous statement reduces to

$\zeta_b(z^2)$ has no zeroes on the critical line $2Re(z^2) = 1$, or $Re(z^2) = 1/2$.

But this is trivially equivalent to the statement

$\zeta_b(z)$ has no zeroes on the critical line $2Re(z) = 1$, or $Re(z) = 1/2$.

as $z \mapsto z^2$ is a bijection. □

The remainder of this paper will focus on sketching why one might believe this criterion to hold, and thereby attempt to upgrade it from a conjecture to a theorem. The main approach used to seek a sufficient proof in this instance will be to construct an appropriate information functional over an exotic geometry of adequate complexity, and apply the Cramer-Rao inequality to same. Hence, in essence, this mirrors (or, rather, extends) the approach taken in [5] towards an investigation of why one might expect the Riemann Hypothesis to be true.

2 Towards the delivery of a sharper estimate for the size of $|S(x)|$

2.1 Naive approach

$$f(m,a) = \int_{C\times C\to C} F(m,b,c)\delta(\sigma(m,b,c)-a)\delta(\tau(m,c)-b)dcdb$$

is a natural signal function, where σ is a Riemann-Cartan metric.
It then follows that:

$$I(f) = \int_M \int_{C\times C} \|\partial f\|^3/f^2 = \int_M \int_{C\times C} f\|\partial lnf\|^3$$

is a natural information, where $\psi = grad_\Lambda f$.

$$I(f) \geq 0$$

by the Cramer-Rao inequality for this information.

Now, by abuse of notation, if we think of f roughly represented as $F\delta$, and that the statistical distributive terms will take care of themselves, then the above information is essentially:

$$I(f) = \int_M \int_{C\times C}((\partial^3 F)\delta + F(\partial^3\delta)) + \int_M \int_{C\times C}(\text{cross terms})$$

It can be demonstrated that, if one assumes that F is asymptotically flat that the boundary terms vanish due to the holographic principle. So then

$$I(f) = \int_M \int_{C\times C}((\partial^3 F) + R^{(2)}(\sigma))\delta = \int_M \int_C (h''' + R^{(2)}(\sigma)\delta)e^h$$

where $F = e^h$, and $R^{(2)}(\sigma)$ is some appropriate geometric invariant. But is this the right invariant, or how things should be defined? Perhaps not.

2.2 Slightly less naive approach

So instead let us consider a more general approach to the matter or triumvirate structures, following on from the consequences of the paper [7].

$$f(m,a) = \int_{Cyb(M)} F(m,b)\delta(\kappa(m,b) - a)db$$

is a natural signal function, where κ is the 6-tensor associated to a first order cybernetic structure, and $Cyb(M) := \{(\mathcal{J}M)^3 \to \mathcal{J}M\}$ is the natural space of distributions, for $\mathcal{J}M$ being the first jet bundle for M.

The corresponding information is:

$$I = \int_M \int_{Cyb(M)\times Cyb(M)} f(\partial log f)^3 dmdVdW$$

with $Cyb(M) = \{f|f : C \times C \to C\}$.

This invariant is positive as a consequence of the Cramer-Rao inequality for a first order cybernetic statistical structure.

Following an analogous argument to the naive approach and eliminating cross terms due to the holographic principle, we have that

$$I = \int_M \int_{Cyb(M)\times Cyb(M)}((\partial^3 F) + S(\kappa))\delta = \int_M \int_{Cyb(M)}(h''' + S(\kappa)\delta)e^h$$

where

$$S = \kappa_{ijklmn}\Gamma_{ijabcde}\Gamma_{klfgabc}\Gamma_{mndefg}$$

where

$$\Gamma_{pijklmn} := \langle E_{ij}, \partial_p E_{kl}, E_{mn}\rangle$$

Then we have that for the information to be critical,

$$h''' + S\delta = 0$$

Hence, $h(z,a) = A(a)z^2 + B(a)z + C + G(z,a)H\frac{H}{\delta}$. This follows since $\delta = \frac{H''}{\delta}$. Therefore $e^h = H(\gamma)\hat{F}$, where $\hat{F} = e^{Az^2+Bz+C}$ as $e^{H/\delta} = H$.

We can furthermore demonstrate that $\gamma = \gamma(rad(2z) - ln(\zeta_a(z)))$, where $rad(z)$ is the extension of the idea of the radical of an integer to a general number.

We can then show via a generalisation of the Riemann-Roch theorem that

$$\gamma = C(rad(2z^2) - ln(\zeta_a(z))).$$

Consequently, if the information is critical (as it will be for the prime numbers),

$$0 = \int_M \int_{Cyb(M)} F(z,a) = \int_M \int_A H(ln(\zeta_a(u)) - rad(2z^2))e^{A(a)z^2+B(a)z+C(a)}$$

$$= \int_M \int_A H(ln(\zeta_a(z)) - rad(2z^2))(\zeta_a^{(2)}(z,z))$$

It then follows that

$$0 = \int_{Re(rad(2z^2)) \geq Re(ln(\zeta_a(z)))} (\zeta_a^{(2)}(z,z))dzdu$$

which provides what appears to be a partial result.

2.3 Demonstration of the second order zeta criterion

We wish to construct an information that is multi variable in our 'physical space', but is not too high order.

Consider then the following signal function:

$$f(m,n,a) = \int_{Cyb(M)} F(m,n,b)\delta(\kappa(m,n,b)-a)db$$

Then we have that for the information to be critical, for $F = e^h$,

$$h''' + S\delta = 0$$

Hence, $h(z,u,a) = E(a,u)z^2 + F(a,z)u^2 + A(a)zu + B(a,u)z + C(a,z)u + D(a) + G(z,u,a)H\frac{H}{\delta}$. This follows since $\delta = \frac{H''}{\delta}$.

Therefore $e^h = H(\gamma)\hat{F}$, where $\hat{F} = e^{Az^2+Bzu+Cu^2+Dz+Eu+F}$ as $e^{H/\delta} = H$.

Lemma 2.1. *(Form of the Heaviside argument, part 1)*

We can demonstrate that

$$\gamma = \gamma(2(z+u)-b)$$

where $rad(z)$ is the extension of the idea of the radical of an integer to a general number.

Proof. This relies on a generalisation of the analogous part of the proof in [5]. Basically, we have that $H_b^{1/3}\hat{\psi} = H^{1/3}\hat{\psi}_b$, and also from $\delta I = 0$ for

$$I = \int H(\gamma)\|\hat{\psi}\|^3$$

and $\partial = \frac{\partial}{\partial z} + \frac{\partial}{\partial u} + \frac{\partial}{\partial b}$

we conclude that

$$\int (H_u^{1/3} + H_z^{1/3} + 2H_a^{1/3})(\partial\gamma)^{-1}\psi^2(H^{1/3}\hat{\psi}) = 0$$

which clearly implies that

$$\gamma_z + \gamma_u + 2\gamma_b = 0$$

which completes the proof.

$\square$

Lemma 2.2. *(Form of the Heaviside argument, part 2)*

We can then show again via a generalisation of the arguments in [Go] that for $t = 2(z+u) - a$, we can demonstrate that $\gamma = Ct + D$ for constants C and D. It then follows via the Riemann mapping theorem that WLOG we can have $\gamma = Ct$.

Proof. The information is

$$I = \int H(\gamma)\|\psi\|^3$$

Computing the first variation with respect to γ and setting this to zero gives

$$\delta(\gamma(t))\tfrac{d\gamma'}{d\gamma}|_{t=a-2z-2u}\|\hat{\psi}\|^3 = 0$$

But this implies that $\frac{d\gamma'}{d\gamma} = 0$, or γ' is constant, hence $\gamma = Ct + D$, and the rest of the proof follows.

$\square$

Consequently, if the information is critical (as it will be for the prime numbers), and for $\kappa(u, z, a) = rad(2z^2) - exp(u) - b)$,

$$0 = \int_M \int_M \int_{Cyb(M)} F(z, a)$$
$$= \int_M \int_M \int_A H(2(z+u) - b) e^{E(b)z^2 + F(b)u^2 + A(b)zu + B(b)z + C(b)u + D(b)}$$

We can then narrow the concern of our endeavours, and choose E, F, A, B, C for $b = (b_1, b_2)$ such that

$$e^{E(b)z^2 + F(b)u^2 + A(b)zu + B(b)z + C(b)u} = b_1^{-z^2} b_2^{-u^2} (b_1 + b_2)^{-uz}$$

Set

$$D(b) = \sum_{n,m} \delta(b_1 - n)\delta(b_2 - m)a_{nm}$$

Then the above reduces to

$$0 = \int_M \int_M \int_A H(2(z+u) - b) \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{nm} n^{-z^2} m^{-u^2} (n+m)^{-uz} \delta(b - (n, m))$$

or

$$0 = \int_{Re(2(z+u))\geq 1} \zeta_a^{(2)}(u,z)$$

which demonstrates the second order zeta criterion - and which in turn, by the considerations above, resolves Hilbert's 8th problem.

3 Further work

It is intriguing to revisit an expression from earlier in this paper

$$\zeta_a^{(3)}(u,v,w) := \frac{1}{3!^{-uvw}} \sum a_{nmp} n^{-u^3} m^{-v^3} p^{-w^3} (nmp)^{-uvw} (nm+p)^{-(u+v)w} (mp+n)^{-(v+w)u} (pn+m)^{-(w+u)v} (n+m+p)^{-uvw}$$

and ask whether phenomena of number associated to this type of quantity might be better understood through examination of a related expression

$$\hat{\zeta}_\epsilon^{(3)}(\underline{u}) := \sum_{\underline{p}\, primes} exp(-\epsilon_{ijk}(\underline{u}) p_i p_j p_k)$$

where $\underline{p} = (p_0, p_1, p_2)$ is a 3-tuple of primes, and ϵ is a 3-tensor.

It then seems like a natural approach to consequently consider studying $\hat{\zeta}^{(3)}(\underline{u})$ and its truncations

$$\hat{\zeta}_\epsilon^{(3)}(\underline{u}|\underline{x}) := \sum_{\underline{p}\leq\underline{x}} exp(-\epsilon_{ijk}(\underline{u}) p_i p_j p_k)$$

and seek to investigate whether these functions, together with deeper ideas from information theory, could shed light on more esoteric areas of number theory.

Of course, this suggests further that there might be a reverse transform for said expression, with

$$\bar{\zeta}_\tau^{(3)}(\underline{u}) = \sum_{\underline{n}} exp(\tau_{ijk}(\underline{u}) n_i n_j n_k)$$

such that $\bar{\zeta}_\tau^{(3)}(\hat{\zeta}_\epsilon^{(3)}(\underline{u})) = \underline{u}$. Indeed, this would be more natural to study, since $\underline{n} \in N^3$ is a more natural space than 3-tuples of primes.

Questions that might benefit from these investigations could include the Collatz Conjecture (otherwise known as the 3n + 1 or Ulam conjecture), and the Erdös-Strauss conjecture. In particular, these are:

Conjecture 3.1. *(Collatz Conjecture). Let n be any natural number. If it is odd, multiply it by 3 and add 1. If it is even, divide it by 2. Continue this process. Then after a finite number of iterations, this arithmetic progression will eventually hit 1; ie, there is some k such that 3 times the kth iterate plus 1 will be some power of 2.*

Conjecture 3.2. *(Erdös-Strauss Conjecture). For every integer $n \geq 2$, there exist positive (possibly identical) integers x, y, z such that:*

$$\tfrac{4}{n} = \tfrac{1}{x} + \tfrac{1}{y} + \tfrac{1}{z}$$

Also, one might be interested in proving that this criticality result holds:

Conjecture 3.3. *(3rd zeta criticality criterion). The only zeroes of $\zeta^{(3)}(u, v, w)$ lie on the critical manifold $2(u + v + w) = 1$.*

Towards obtaining an analogue of the Euler product formula for this 3rd zeta function, one might be interested in studying the set

$$S^{(2)}(x) := \{(a, b, c)|a + b + c = x, gcd(a_i, a_j) = 1, rad(abc(a + b)(a + c)(b+ c)(a + b + c)) < x\}$$

or some related set

$$S^{(2)}(x, y) := \{(a, b, c)|a + b + c = x + y, Taxicab(rad^{(2)}(abc(a + b)(a + b+ c), abc(a + c)(a + b + c), abc(b + c)(a + b + c))) < Taxicab(x, y, x + y)\}$$

where $rad^{(2)} : R^3 \to R^3$ is some natural lift of rad from R to R^3, or from C to the quaternions Q for the associated analytic lift. (Note that this may not simply correspond to applying rad to each term in the vector, as we wish $rad^{(2)}$ to satisfy naturality properties:

$$rad^{(2)}(xy) = rad^{(2)}(rad^{(2)}(x)rad^{(2)}(y)), \text{ for instance.)}$$

But of course, the Quaternions are but a special case of the Cayley-Dickson construction, which, beyond the complex numbers, allows representations of elements of such algebras as matrices. So instead, let us define $rad^{(2)} : GL2(R^2) \to GL2(R^2)$

as a map between 2 by 2 matrices with the naturality properties we require. I claim that it is possible to create such a map, and that it is unique. Also note that there is a natural analytic extension to $GL2(C^2) \rightarrow GL2(C^2)$.

But then $Taxicab(a, b, c)$, being the number that has c representations as the sum of b ath powers of natural numbers, no longer has a sufficiency of arguments. We wish to extend the idea of generalised Taxicab numbers, perhaps, to allow it to take a matrix argument in a natural way.

Consider then $Taxicab(a, B, c)$ as the matrix that has c representations as the sum of B ath powers of natural numbers, where B is a 2 by 2 matrix, ie, if

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

then we are interested in determining the matrix $M = Taxicab(a, B, c)$ such that

$$M = \begin{bmatrix} \Sigma_{i=1}^{b_{11}} n_{11i}^a & \Sigma_{i=1}^{b_{12}} n_{12i}^a \\ \Sigma_{i=1}^{b_{21}} n_{21i}^a & \Sigma_{i=1}^{b_{22}} n_{22i}^a \end{bmatrix}$$

Then it might be sensible to define

$$S^{(2)}(x, y) := \{(a, b, c)|a + b + c = x + y, A = rad^{(2)}(\begin{bmatrix} a(b + c) & c(a + b) \\ b(a + c) & (a + b + c) \end{bmatrix}),$$
$$det(Taxicab(xy, A, x + y)) < tr(Taxicab(xy, diag(xy, x + y), x + y))\} \quad (20)$$

Then investigate finding a triple product identity perhaps of the form

$$\zeta^{(3)}(u, v, w) = \Pi_{(p,q)\ prime} \Pi_{(a,b,c) \in S^{(2)}(p,q)} (1 - \mathcal{F}(a, b, c, u, v, w))^{-1} \quad (21)$$

for some mysterious function $\mathcal{F}$.

There is a natural question as to whether it is indeed $Taxicab$ that we would like to consider in the definition of $S^{(2)}(x, y)$, or rather perhaps instead an invariant associated to the lengths of sequences of Aliquot numbers, or perhaps instead some form of generalisation of the Partition function. Recall that $p(n)$ is the number of ways to construct a number as the sum of some number of positive integers. But this is an inverse of $\sum_m Taxicab(1, m, p(n))$, as $Taxicab(1, m, p(n, m))$ is the number that has $p(n,m)$ ways to be constructed from the sum of m integers, where $p(n,m)$ is the number of ways to add m integers together to get n.

Certainly if y and c are zero, then the expression above reduces to $S^{(1)}(x)$, as we must necessarily have that $rad^{(1)}(ab(a+b)) < x$. Whether or not this is a sensible definition, of course, is another matter entirely.

As to our function $\mathcal{F}$, a natural choice might well be

$$F(u,v,w;a,b,c) := {}_3F_1^{(2)}\left[\begin{matrix} B(u,a), B(v,b), B(w,c) \\ B(u+v+w, a+b+c) \end{matrix}; -alt^{(1)}(u,v,w), -alt^{(2)}(u,v,w)\right] \tag{22}$$

where

$$alt^{(1)}(u,v,w) := (u+v)rad^{(1)}(w) + (v+w)rad^{(1)}(u) + (w+u)rad^{(1)}(v) \tag{23}$$

and

$$alt^{(2)}(u,v,w) := (u-v)rad^{(1)}(w) + (v-w)rad^{(1)}(u) + (w-u)rad^{(1)}(v) \tag{24}$$

Here

$${}_3F_1^{(2)}\left[\begin{matrix} a,b,c \\ d \end{matrix}; u,v\right] := \Sigma_{n,m} \frac{(a)_{n,m}(b)_{n,m}(c)_{n,m}}{(d)_{n,m}} \frac{u^{n^2}v^{m^2}(u+v)^{nm}}{n!!m!!} \tag{25}$$

where

$$(q)_{n,m} := \Pi_{k=0}^{n-1}\Pi_{l=0}^{m-1}(q+k+l) \tag{26}$$

is the generalised Pochhammer symbol, or the rising double factorial, for $n,m > 0$. (For the edge cases, if $m=0$, $(q)_{n,0} = (q)_n$, similarly $(q)_{0,m} = (q)_m$.)

Of course, we might ask to consider the generalised Beta function

$$B(x,y;z) := \int_{t^2+s^2=0}^{1} (t^2)^{x-1}(s^2)^{y-1}(1-t^2-s^2)^{z-1}$$

or equivalently

$$B(x,y;z) := \int_{\|(t,s)\|=0}^{1} (t^2)^{x-1}(s^2)^{y-1}(1-t^2-s^2)^{z-1} \tag{27}$$

in the above. We then might instead posit that a natural $\mathcal{F}$ could take the form

$$ {}_3F_1^{(2)}\left[\begin{matrix} B(u,v+w;a), B(v,w+u;b), B(w,u+v;c) \\ B(u+v+w,uvw;a+b+c) \end{matrix}; -alt^{(1)}(u,v,w), -alt^{(2)}(u,v,w)\right] \tag{28}$$

Irrespective of this, there is definitely more to explore and investigate further in this direction. Indeed, there is fertile ground here for future work.

References

3.1 Core References

[1] "Spectral Theory and the Trace Formula (Expanded Text)", Daniel Bump

[2] "A generalisation of the Bernoulli numbers from the discrete to the continuous", Donal F. Connon, 16 May 2010

[3] "Spectral Theory for SL2ZH", Emanuel Geomin, September 9, 2013

[4] "Weil's Conjecture for Function Fields", Dennis Gaitsgory and Jacob Lurie

[5] "A treatise on information geometry", http://vixra.org/abs/0908.0073, 2010 Chris Goddard.

[6] "Advanced topics in information dynamics", http://www.rxiv.org/abs/1002.0033, 2010 Chris Goddard.

[7] "Optimisation of Dynamical Systems Subject to Meta-Rules", http://vixra.org/abs/1604.0148, 2015 Chris Goddard

[8] "Some Problems of 'Partitio Numerorum'; III: On the Expression of a Number as a Sum of Primes", G. H. Hardy and J. E. Littlewood, 1922

[9] "Extremal Combinatorics", 2001 S. Jukna

[10] "On the number of primes less than a given magnitude", 1859 Bernhard Riemann.

[11] "254A, Notes8: The Hardy-Littlewood circle method and Vinogradov's theorem", Terry Tao, 30 March 2015, blog entry

[12] "Heuristic limitations of the circle method", Terry Tao, 20 May 2012, blog entry

[13] "Brownian motion, moving metrics and entropy formulas", Anton Thalmaier, Université du Luxembourg, link

[14] "Notes on the Hardy-Littlewood Circle Method", Jacques Verstraëte, University of California, San Diego

[15] "On Partitions of Goldbach's Conjecture", Max See Chine Woon, Cambridge University, Cambridge

[16] "Newman's Short Proof of the Prime Number Theorem", D. Zagier, The American Mathematical Monthly, Vol. 104, No. 8 (Oct., 1997), pp. 705-708

3.2 Mochizuki's program

[17] "Effectivity in Mochizuki's work on the abc-conjecture", Vesselin Dimitrov, arxiv.org/abs/1601.03572

[18] "Inter-universal Teichmüller Theory I", Shinichi Mochizuki, link

[19] "Inter-universal Teichmüller Theory II", Shinichi Mochizuki, link

[20] "Inter-universal Teichmüller Theory III", Shinichi Mochizuki, link

[21] "Inter-universal Teichmüller Theory IV", Shinichi Mochizuki, link

[22] "A Panoramic Overview of Inter-universal Teichmüller Theory", Shinichi Mochizuki, link