Around Wilson’s theorem

Alain Connes

September 11, 2018

Abstract

We study the series $s(n,x)$ which is the sum for $k$ from 1 to $n$ of the square of the sine of the product $x \Gamma(k)/k$, where $x$ is a variable. By Wilson’s theorem we show that the integer part of $s(n,x)$ for $x = \pi/2$ is the number of primes less or equal to $n$ and we get a similar formula for $x$ a rational multiple of $\pi$. We show that for almost all $x$ in the Lebesgue measure $s(n,x)$ is equivalent to $n/2$ when $n$ tends to infinity, while for almost all $x$ in the Baire sense, $1/2$ is a limit point of the ratio of $s(n,x)$ to the number of primes less or equal to $n$.

1 Introduction

Let $\Pi(n)$ be the number of primes $p \leq n$. A slight improvement on a formula of Willans gives a simple formula for $\Pi(n)$ as the integer part of the sum

$$\sum_{k=1}^{n} \sin^2 \left( \frac{\pi \Gamma(k)}{2k} \right)$$

(1)

When one tries to compute naively the right hand side one finds that it requires an increasing precision on the numerical value of the number $\pi$ whose first 2500 decimals are needed to compute $\Pi(n)$ for $n$ of the order of a thousand. F. Villegas suggested to replace $\pi$ by a variable and analyse the dependence on $x$ in the above sequence. Thus for $n > 1$ an integer and $x \in \mathbb{R}$, let

$$s(n, x) := \sum_{k=1}^{n} \sin^2 \left( \frac{x \Gamma(k)}{k} \right)$$

(2)

Figure 1: The formula of Willans.
We shall show below that the dependence on \( x \in \mathbb{R} \) is quite interesting inasmuch as, due to the lacunary nature of the sequence \( \frac{\Gamma(n)}{n} \), the terms of the sum \( \sum \) are essentially independent random variables when suitably understood as functions on an almost periodic compactification \( G \) of \( \mathbb{R} \). This gives, by the proof of the strong law of large numbers, that for almost all \( x \in \mathbb{R} \) in the sense of the Lebesgue measure one has when \( n \to \infty \) that \( s(n, x) \sim \frac{\mu}{2} \). The interesting fact is that for the other natural notion of “generic” real number, namely the one provided by the Baire theory of dense countable intersections of open sets, it is a totally different behavior of the sequence \( s(n, x) \) which is generic: we show in Theorem \( 4.1 \) that for generic \( x \in \mathbb{R} \), the quotients \( \frac{s(n, x)}{\Pi(n)} \) get arbitrarily close to \( \frac{1}{2} \), i.e. \( \frac{1}{2} \) is a limit point of the sequence

\[
\frac{1}{2} \in \lim_{n \to \infty} s(n, x) \Pi(n).
\]

Generically this sequence will also have \( \infty \) as a limit point and will oscillate wildly. But for rational multiples of \( \pi \) the sequence \( s(n, x) \) behaves like the product of \( \Pi(n) \) by the rational number \( 1 - \frac{\mu(b)}{2\phi(b)} \), which only depends upon the denominator \( b > 1 \) of the irreducible fraction \( x = \frac{a}{b} \pi \) as a multiple of \( \pi \) (see Proposition 3.1).

## \( \Pi(n) \) and sum of squared sines

We start with the following variant of the formulas of Willans [4].

**Proposition 2.1.** Let \( n > 1 \) be an integer then \( \Pi(n) \) is the integer part of \( s(n, \frac{\pi}{2}) \).

**Proof.** For \( k > 4 \) not prime the quotient \( (k - 1)!/k \) is an even integer. Thus in that case one has

\[
\sin^2 \left( \frac{\pi \Gamma(k)}{2k} \right) = 0
\]

For \( k = p > 2 \) prime, the residue of \( (p - 1)! \) modulo \( 2p \) is \( p - 1 \) by Wilson’s theorem and the evenness of \( p - 1 \). Thus for \( p > 2 \) prime,

\[
\sin^2 \left( \frac{\pi \Gamma(p)}{2p} \right) = \sin^2 \left( \frac{\pi (p - 1)}{2p} \right) = \cos^2 \left( \frac{\pi}{2p} \right).
\]

One has

\[
1 \geq \cos^2(x) \geq 1 - x^2, \quad \forall x \in \mathbb{R}
\]

It follows that \( \delta(n) = s(n, \frac{\pi}{2}) - \Pi(n) \) is a decreasing function of \( n > 4 \) and that for \( m > n \),

\[
\delta(n) - \delta(m) = \sum_{\substack{p \text{ prime}, n < p \leq m \\atop \mu(b)}} \left( 1 - \cos^2 \left( \frac{\pi}{2p} \right) \right) \leq \frac{\pi^2}{4} \sum_{n < p \leq m} \frac{1}{p^2}
\]

The series \( \sum \frac{1}{p^2} \) is convergent and one has the bound \( \frac{\pi^2}{4} \sum_{p > 50} \frac{1}{p^2} < 0.0498448 \) for the (larger) sum over integers, while \( s(50, \frac{\pi}{2}) - \Pi(50) \sim 0.539005 \). It follows that

\[
s(n, \frac{\pi}{2}) \in \big[ \Pi(n) + 0.5 - 0.05, \Pi(n) + 0.6 \big] \subset \big[ \Pi(n), \Pi(n) + 1 \big)
\]

for any \( n > 50 \), and (see Figure 2) \( s(n, \frac{\pi}{2}) \in \big[ \Pi(n), \Pi(n) + 1 \big) \) for any \( n > 1 \) which gives the required result. \( \square \)

---

\( \mu \) is the Möebius function and \( \phi \) the Euler totient function.
The general term \( \sin^2 \left( \frac{x \Gamma(k)}{k} \right) \) of (2) depends on the knowledge of \( x \) up to an \( \epsilon \) of the order of

\[
dx \simeq \left( \frac{k}{\epsilon} \right)^{-k+2}
\]

Thus for instance to get the required precision around \( k = 945 \) one needs the first 2400 decimals of \( \pi \).

Figure 3: 2400 decimals of \( \pi \).

### 3 Rational multiples of \( \pi \)

We investigate the behavior of the sequence \( s(n, x) \) when \( x \) is a rational multiple of \( \pi \). We let \( x := \frac{\pi}{a} \) where \( a \) and \( b \) are relatively prime. For \( k \) large enough the term \( \Gamma(k)/k \) is divisible by \( b \) in the sense that the \( p \)-adic valuation of \( \Gamma(k)/k \) is larger than that of \( b \) for all prime divisors of \( b \). This can be seen using the Legendre formula for the \( p \)-adic valuation which gives

\[
v_p(\Gamma(k)/k) \geq \sum_{\ell \geq 1} \left( k - 1 \right)/p^\ell \geq \left\lfloor \log_p(k) \right\rfloor
\]

Since the prime factors of \( b \) are fixed there exists \( k_0 < \infty \) such that \( \Gamma(k)/k \) is divisible by \( b \) (in the above sense) for all \( k > k_0 \). Thus if \( k > k_0 \) is not prime the product \( \frac{x \Gamma(k)}{k} \) is an integer
multiple of \( \pi \) and \( \sin^2 \left( \frac{x \Gamma(k)}{k} \right) = 0 \). When \( k > k_0 \) is a prime, the integer \( c = \frac{\Gamma(k)}{b} \) is such that \( bc = -1 \) modulo \( k \) by Wilson’s theorem. One gets in this case

\[
\sin^2 \left( \frac{x \Gamma(k)}{k} \right) = \sin^2 \left( \frac{\pi a \Gamma(k)}{bk} \right) = \sin^2 \left( \frac{\pi ac}{k} \right)
\]

and the right side only depends upon the residue of \( a \) and of \( c \) modulo \( k \). Since \( b \) and \( k \) are relatively prime there exists \( u \in \{1, \ldots, b-1\} \) which is the inverse of \( k \) modulo \( b \). Thus let \( m \in \mathbb{N} \) such that \( ku - 1 = bm \). One then has \( bm = -1 \) modulo \( k \) and it follows that \( m = c \) modulo \( k \). Thus

\[
\sin^2 \left( \frac{\pi ac}{k} \right) = \sin^2 \left( \frac{\pi am}{k} \right) = \sin^2 \left( \frac{\pi a(ku - 1)}{bk} \right) = \sin^2 \left( \frac{\pi au}{b} \right) + \sin \left( \frac{a \pi}{bk} \right) \sin \left( \frac{a \pi}{bk} - \frac{2au \pi}{b} \right)
\]

Thus, since

\[
|\sin \left( \frac{a \pi}{bk} \right) \sin \left( \frac{a \pi}{bk} - \frac{2au \pi}{b} \right)| \leq |\sin \left( \frac{a \pi}{bk} \right)| = 0(1/k)
\]

the asymptotic behavior of \( s(n, x) \) only depends, up to a term of the order of \( \log \log n \) due to the \( \sum 1/k \) over primes less than \( n \), upon the residues of the primes modulo \( b \). Thus by Dirichlet’s theorem, in the strong form due to La Vallée Poussin (see \([2]\), V §7), one gets

**Proposition 3.1.** Let \( x \) be a rational multiple of \( \pi \), \( x = \frac{\pi a}{b} \) with \( a, b \) relatively prime, \( b > 1 \), then

\[
s(n, x) \sim \frac{\Pi(n)}{\phi(b)} \sum_{\nu \in (\mathbb{Z}/b\mathbb{Z})^*} \sin^2 \left( \frac{\pi \nu}{b} \right) = \Pi(n) \left( \frac{1}{2} - \frac{\mu(b)}{2\phi(b)} \right)
\]

**Proof.** It remains to show the second equality in (3). It follows from \( \sin^2(\nu) = \frac{1-\cos(2\nu)}{2} \) and the fact that the sum of primitive roots of unity of order \( b \) is the Möbius function \( \mu(b) \). \( \square \)

## 4 Generic behavior of \( s(n, x) \)

This suggests to investigate the general behavior of the sequence \((2)\). It is given by the following result which shows that almost everywhere in measure theory (for the Lebesgue measure) the behavior is Gaussian and \( s(n, x) \sim \frac{n}{2} \). But generically at the topological level which means on a dense countable intersection of open sets, the sum grows far more slowly and behaves like \( \frac{\Pi(n)}{2} \) in the weak sense that \( \frac{1}{2} \) is a limit point of the sequence \( \frac{s(n, x)}{\Pi(n)} \). In fact it oscillates wildly since \( \infty \) is also a limit point of that sequence. Note that the two behaviors are exclusive of each other but this is not a contradiction.

**Theorem 4.1.** (i) For almost all \( x \in \mathbb{R} \) the sequence \( s(n, x) \) of \((2)\) has the Gaussian behavior

\[
s(n, x) \sim \frac{n}{2}
\]

(ii) For generic \( x \in \mathbb{R} \) (i.e. on a dense countable intersection of open sets) one has

\[
\frac{1}{2} \in \lim_{n \to \infty} \frac{s(n, x)}{\Pi(n)}
\]
Proof. (i) Let $G$ be the compact group projective limit of the compact groups $G_n := \mathbb{R}/n\mathbb{Z}$ under the natural morphisms
\[
y_{n,m} : G_n \rightarrow G_m, \quad y_{n,m}(x + m\mathbb{Z}) = x + n\mathbb{Z}, \quad \forall n|m. \tag{4}
\]
One has a natural isomorphism, with $A_Q$ the adeles of the global field $Q$,
\[
G = \lim_{\leftarrow \lim} (\mathbb{Z}/n\mathbb{Z} \times \mathbb{R})/\mathbb{Z} = (\hat{\mathbb{Z}} \times \mathbb{R})/\mathbb{Z} = A_Q/(Q, +).
\]
The Pontrjagin dual of $G$ is identified with the discrete additive group $Q$ of rational numbers by associating to $r \in Q$ the character $\alpha_r$ of $G$ specified by its restriction to the dense subgroup $\mathbb{R}$, range of the homomorphism $\mathbb{R} \ni t \mapsto a(t) = (0, t) \in A_Q/(Q, +)$ of adeles with 0 non-archimedean component
\[
\alpha_r(a(t)) := e^{2\pi irt}.
\]
Next, using $1 - \cos(2x) = 2\sin^2(x)$ we get with $r = \frac{\Gamma(k)}{k} \in Q$ the equality
\[
\sin^2 \left( \frac{x\Gamma(k)}{k} \right) = \frac{1}{2} - \frac{1}{4} \alpha_r \left( \frac{x}{\pi} \right) - \frac{1}{4} \alpha_{-r} \left( \frac{x}{\pi} \right). \tag{5}
\]
Thus with the basic function defined on $G$ by
\[
X(x) := -\frac{1}{4} \alpha_1(x) - \frac{1}{4} \alpha_{-1}(x)
\]
we get that the general term of the sum $s(n, x)$ is simply $\frac{1}{2} + X_k \left( \frac{x}{n} \right)$ where
\[
X_k(x) := X(r(k)x), \quad r(k) := \frac{\Gamma(k)}{k} \in Q^x
\]
The multiplication by elements of $Q^x$ defines automorphisms of $G$. One has the orthogonality relation of characters which implies since the rationals $\frac{\Gamma(k)}{k}$ are distinct, they form (for $k > 1$) a strictly increasing sequence (the first ones are $\{1, \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \frac{9}{2} \}$) that the random variables $X_k$ on the probability space $G$ equipped with its normalized Haar measure, are equidistributed and essentially independent inasmuch as, $\int_G X_k(x)X_{k'}(x)dx = 0, \forall k \neq \ell$ and that one controls the 4-th moment as follows
\[
\int_G |\sum_{k=1}^{n} X_k(x)|^4 dx \leq Cn^2 \tag{6}
\]
since $|\sum_{k=1}^{n} X_k(x)|^4 = \sum_{1 \leq k_1 \leq n} X_{k_1}(x)X_{k_2}(x)X_{k_3}(x)X_{k_4}(x)$ and the number of solutions of the equation
\[
\pm r(k_1) \pm r(k_2) \pm r(k_3) \pm r(k_4) = 0, \quad k_j \in \{1, \ldots, n\} \tag{7}
\]
is of the order of $n^2$ due to the lacunary nature ([3]) of the sequence $r(k)$. Indeed for $k > 4$ one has $r(k+1) > 3r(k)$ and (7) is possible only if the largest $k_j$ appears at list twice (and with opposite signs) and the remaining $k_j$ are equal, which gives $n^2$ as a bound on the number of solutions.

Thus one has, for any $\epsilon > 0$ that
\[
\int_G \frac{1}{n} |\sum_{k=1}^{n} X_k(x)|^4 \leq C/n^2, \quad |\{x \in G | |\frac{1}{n} \sum_{k=1}^{n} X_k(x)| > \epsilon\}| \leq C/n^2 e^{-4}
\]
and the Borel-Cantelli lemma applies and shows that the subset $E \subset G$ defined by

$$E := \{ x \in G \mid \frac{1}{n} \sum_{k=1}^{n} X_k(x) \to 0 \}$$

is of measure 1. Since $\mathbb{R} \subset G$ is of measure 0 we cannot yet get (i) but it will follow from the invariance of $E$ under the translation by the subgroup $\hat{\mathbb{Z}} \subset G$. To see this we use the equality for $k$ non-prime

$$X_k(x + u) = X_k(x), \quad \forall u \in \hat{\mathbb{Z}}$$

which follows from the integrality of $\frac{n}{\Gamma(k)}$ and the periodicity of the cosine:

$$\cos \frac{2\pi(x + 1)\Gamma(k)}{k} = \cos \frac{2\pi x \Gamma(k)}{k}$$

Thus one gets the same equality for the closure $\hat{\mathbb{Z}} \subset G$. It follows that

$$\left| \frac{1}{n} \sum_{k=1}^{n} X_k(x + u) - \frac{1}{n} \sum_{k=1}^{n} X_k(x) \right| \leq \frac{\Pi(n)}{n}, \quad \forall u \in \hat{\mathbb{Z}}$$

and this suffices to show that $E$ is invariant under the translation by the subgroup $\hat{\mathbb{Z}} \subset G$. The image of $x \in G$ in the quotient $G/\hat{\mathbb{Z}} = \mathbb{R}/\mathbb{Z}$ thus suffices to decide if $x \in E$ and it follows that almost all elements of $\mathbb{R} \subset G$ are in $E$. Finally the Gaussian behavior follows from the results of [3] on lacunary trigonometric series.

(ii) When $x \in \pi \mathbb{Q}$ is a rational multiple of $\pi$ one applies Proposition 3.1. For $b \to \infty$ one has the equidistribution

$$\frac{1}{\phi(b)} \sum_{v \in (\mathbb{Z}/b\mathbb{Z})^*} \sin^2 \left( \frac{\pi v}{b} \right) = \frac{1}{2} - \frac{\mu(b)}{2\phi(b)} \to \frac{1}{2}.$$ 

This shows that, for $\epsilon > 0$, the following countable intersection of open sets is dense in $\mathbb{R}$

$$W(\epsilon) := \bigcap_{m} \bigcup_{n \geq m} \{ x \in \mathbb{R} \mid (1-\epsilon)\Pi(n)/2, (1+\epsilon)\Pi(n)/2 \}$$

and for $x \in W(\epsilon)$ one has $\lim_{n \to \infty} \frac{2\epsilon(n,x)}{\Pi(n)} \cap [1-\epsilon, 1+\epsilon] \neq \emptyset$ which gives the required conclusion after intersecting the $W(\epsilon)$ for $\epsilon = \frac{1}{a}, a \to \infty$ which still gives a dense countable intersection of open sets by Baire’s Theorem [1].

References

[1] R. Baire. *Sur les fonctions de variables réelles*. Ann. di Mat., 3:1–123, (1899).

[2] K. Prachar, *Primzahlverteilung*. (German) Springer-Verlag, Berlin-Göttingen-Heidelberg, (1957). x+415 pp.

[3] R. Salem, A. Zygmund, *On Lacunary Trigonometric Series*. Proceedings of the National Academy of Sciences of the United States of America Vol. 33, No. 11 (Nov. 15, 1947), pp. 333–338

[4] C. P. Willans, *On formulae for the nth prime number*. Math. Gaz. 48 (1964) 413–415.