Convergence of Random Fourier Series

RYAN MCCORVIE
Classical Harmonic Analysis (Math 258)
May 14, 2015

Summary

This essay examines the relationship between various modes of convergence for Fourier series whose coefficients are symmetric independent random variables, along with applications of those series to harmonic analysis

Introduction

This essay examines the convergence properties of the random trigonometric series such as

$$\sum_{n=0}^{\infty} \xi_n e^{i\omega_n} e^{itn} \tag{1}$$

Here the $\xi_m e^{i\omega_n}$ are independent and symmetric complex random variables. This was first investigated by Paley and Zygmund in series of papers in the 1930's [1]. To give some context, Payley and Zygmund published their article just a few years after Komolgorov proved the famous three-series criterion for the convergence of random series, and after Weiner discovered a random Fourier series for Brownian motion. This essay closely follows the approach of Kahane in [2], which is the classic reference on this material.

In addition to their intrinsic interest, random methods allow for insights into harmonic analysis more generally. Random constructions sometimes yield examples of series with certain properties which are hard to construct explicitly.

Our primary goal is to show that given the independent, symmetric random Fourier coefficients $\xi_n e^{i\omega_n}$, several convergence properties of (1) are related. Our main result shows that (1) represents a function almost surely, if and only if it converges almost surely, if and only if it converges in $L^p(\mathbb{T})$ for all $p \in [0,1)$. Also (1) represents a continuous function almost surely if and only if it is in $L^\infty(\mathbb{T})$ almost surely.

Wiener Process

Before moving on, let's consider one motivating example. The Wiener process is the stochastic process defined by three properties: $W(0) = 0$, $W(t)$ is almost surely continuous everywhere, and $W(t)$ has independent increments where $W(t) - W(s)$ has a normally distributed with mean 0 and variance $t - s$. There exists a unique random process with these properties (for a proof see [3]). The Wiener process playes a central role in probability theory, as well as in statistics and applied mathematics.

Here we provide a heuristic construction, which also shows a representation of the Wiener process as a random Fourier series. Let e_k be any orthonormal basis of $L^2(\mathbb{T})$ and let ζ_k be independent Gaussian random variables with mean 0 and variance 1. Consider the sum $W(t) = \sum_k \zeta_k \int_0^t$ is a Weiner process. Its Gaussian since the ζ_k, and the sample paths are continuous since fixed values of η_k, each $\int_0^t e_k(t)$ are continuous. To see it satisfies the covariance property, note that $1_{[0,t)} = \sum_k e_k \int_{\mathbb{T}} e_k 1_{[0,t)} = \sum_k e_k \int_0^t e_k$. Therefore by Parseval's thoerem

$$\sum_k \int_0^s e_k \int_0^t e_k = \int_{\mathbb{T}} 1_{[0,t)} 1_{[0,s)} = s \wedge t \tag{2}$$

The expression the left is what you get calcuating $EW(s)W(t)$ term by term, and the expression on the right is the desired expression for the covariance. Thus, $W(s)$ represents a Weiner process.

If we take the Fourier basis given by $e_0(x) = \frac{1}{\sqrt{2\pi}}$, $e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$ for $n \neq 0$, then

$$W(t) = \frac{\zeta_0}{\sqrt{2\pi}} t + \sum_{n \neq 0} \frac{\zeta_n}{\sqrt{2\pi} ni} (e^{int} - 1) \tag{3}$$

This is called the Fourier-Wiener series, and is itself in the form of a random Fourier series. This example also shows the interplay of $L^2(\Omega)$ and $L^2(\mathbb{T})$ for functions defined by random Fourier series.

1 Random Series

Before wading into the properties of random Fourier series, first lets collect some basic facts about probability and convergence.

Definition. A random variable is *symmetric* if the ξ and $-\xi$ have the same distribution. If ξ takes values in $\mathbb{R}$ then $E\xi = 0$.

This study will concern itself only with symmetric random variables. These will turn out to have nice properties which greatly facilitate the study of random series. Here are some of the particular series which we will examine in more detail.

Definition. Let u_k be non-random elements of a Banach space. A *Rademacher series* is given by $\sum_k \epsilon_k u_k$ where ϵ_k are independent random variables which takes on values ± 1 with equal probability. A *Steinhaus series* is given by $\sum_k e^{i\omega_k} u_k$ where ω_k are independent uniform random variables in $[0,1]$. A *Gaussian series* is given by $\sum_k \zeta_k u_k$ where ζ_k are normal random variables.

It will turn out that the Rademacher series is a fundamental object for understanding the convergence of general symmetric random series. To understand why, first note that a general symmetric random variable can be written $\xi = \epsilon\nu$ where $\nu \geq 0$ and ν is independent of ϵ. Therefore statements about $\sum_k \xi_k u_k = \sum_k \epsilon_k \nu_k u_k$ can be analyzed by first conditioning on the values ν_k then analyzing the Rademacher series with coeffients given by $\nu_k u_k$. Essentially this technique is Fubini's thoerem, $Ef(\xi,\eta) = \int f(s,t)\mu \otimes \nu(ds\,dt) = \int \mu(ds) \int f(s,t)\nu(dt) = \int g(s)\mu(ds) = Eg(s)$ where $g(s) = Ef(s,\eta)$. A sufficient condition for this technique to work is that $E|f(\xi,\eta)| < \infty$.

One situation where this approach can be strikingly successful is in proving and event happens almost surely. Let A be an event, and $f = 1_A$. The analysis in the preceding paragraph can be summarized this way: if a random series has a certain property almost surely for every Rademacher series formed by fixed values of ν_k, then it has the property almost surely. Kahane [2] calls this the "principle of reduction", Kallenberg [3] calls this "conditioning". This is useful since Rademacher series often allow for sharp quantitative analysis which generic symmetric series resist.

Given their centrality, its interesting to provide a concrete model for this probability space of Rademacher functions. Take $\Omega = [0,1]$ and endow it with the Lesbegue measure. In this setting, we can represent the nth Rademacher function $r_n(x) : \Omega \to \{\pm 1\}$ in terms of the b_n, the nth binary digit of x, by $r_n(x) = 2b_n(x) - 1$. Then any statement about the joint distribution of some ϵ_k can be translated in to statements about subsets of $[0,1]$ and the behavior of the r_n.

The properties we will study, such as convergence, boundedness and continuity, do not depend on any finite set of the ξ_k. These properties are unchanged if a finite number of the values of ξ_k are adjusted, the events under consideration are independent of any finite number of ξ_k. However, the events are in the probability space generated by the values of the ξ_k. This leads to a curious situation. For the events under consideration, the probabilities are trivial in the sense that they are zero or one. This idea is captured in the following theorem.

1.1 Theorem (Komolgorov 0-1 law). *Let A be an event which depends on independent random variables ξ_k, but does not depend on the particular values of any finite subcollection. Then $P(A) = 1$ or $P(A) = 0$. In particular, $\sum_k \xi_k$ converges almost surely or almost never.*

Proof. (sketch) The event A is independent of any event which is defined by the values of $\xi_1, \ldots, \xi_n$. But then A is independent of any event defined by $\xi_1, \xi_2, \ldots$. Since A is such an event, by independence, $P(A) = P(A \cap A) = P(A)^2$. □

1.1 Concentration inequalities

This section records some of the quantitative bounds which allows us to analyze convergence. As a benchmark, its worth comparing these with teh Markov inequality which states that for any positive function ϕ, $P\{\phi(\xi) > rE(\phi)\} \le \frac{1}{r}$.

1.2 Theorem (Paley-Zygmund inequality). *Given ξ is a $\mathbb{R}_+$-valued random variable, $P(\xi > rE\xi) \ge (1-r)_+^2 E\xi^2$. For independent symmetric ξ_k,*

$$P\{(\sum_k \xi_k)^2 \ge r \sum_k \xi_k^2\} \ge (1-r)^2/3 \tag{4}$$

Proof. Start with $E\xi = E\xi 1_{\{\xi \le r\}} + E1_{\{\xi > r\}}$. We can bound the first term by $rE\xi$ and use Cauchy-Schwartz to bound the second term

$$E\xi 1_{\{\xi > rE\xi\}} \le (E\xi^2)^{1/2}(E1^2_{\{\xi > rE\xi\}})^{1/2} \tag{5}$$

which gives the first inequality.

For the second inequality, first consider the case where $\xi_k = \epsilon_k a_k$ for non-random a_k. Then

$$E(\sum_k \xi_k)^4 = \sum_k a_k^4 + 6\sum_{j<k} a_j^2 a_k^2 \le 3(\sum_k a_k^2)^2 = 3E\sum_k \xi_k^2 \tag{6}$$

since for distinct i, j, k, l,

$$E\xi_i\xi_j^3 = E\xi_i\xi_j\xi_k^3 = E\xi_i\xi_j\xi_k\xi_l = 0 \tag{7}$$

Now apply the above inequality to $\xi = (\sum_k \xi)^2$.

For the general symmetric ξ_k write $\xi_k = \epsilon_k \eta_k$ and use the principle of reduction. Let $A = \{(\sum_k \xi_k)^2 \ge r\sum_k \xi_k^2\}$. The argument above shows that $P(A) \le\ge E_\eta(1-r)^2/3$ when η is fixed. Integrating over the distribution for η shows the total probability satisfies the same inequality. □

The previous result gives a lower bound on the probability of large sums. The next gives an upper bound.

1.3 Theorem (Komolgorov inequality). *Let ξ_k be independent random variables with mean zero and let $S_n = \xi_1 + \cdots + \xi_n$.*

$$P\{\sup_n |S_n| > r\} \le r^{-2}\sum_n E\xi_n^2, \qquad r > 0 \tag{8}$$

Proof. We may assume $E\xi_n^2 < \infty$, since otherwise the inequality is trivial. Let $\tau = \inf\{n : |S_n| > r\}$. Now τ depends only on $\xi_1, \ldots, \xi_k$ so $S_n - S_k = \xi_{k+1} + \ldots \xi_n$ so its independent of the event $\{\tau = k\}$. Therefore

$$\sum_{k\le n} E\xi_k^2 = ES_n^2 \ge \sum_{k\le n} E[S_n^2; \tau = k] \ge \sum_{k\le n} E[S_k^2; \tau = k] + 2E[S_k(S_n - S_k); \tau = k]$$

By independence $E[S_k(S_n - S_k); \tau = k] = E[S_n - S_k]E[S_k; \tau = k] = 0$. Also $E[S_k^2; \tau = k] \ge r^2 P\{\tau \le n\}$. So as $n \to \infty$

$$\sum_k E\xi_k^2 \ge r^2 P\{\tau < \infty\} = r^2 P\{\sup_k |S_k| > r\} \tag{9}$$

□

Next is a result which sometimes allows us to bootstrap probability estimates for random symmetric series into much stronger ones.

1.4 Theorem (Reflection inequality)**.** *Let ξ_k be independent symmetric random variables, and $S_n = \xi_1 + \cdots + \xi_n$. If $S_n \to S$ converges almost surely then*

$$P\{\sup_n |S_n| > r\} \leq 2P\{|S| > r\} \tag{10}$$

Proof. Let $\tau = \min\{n : |S_n| > r\}$. Then $T = S - S_n = \sum_{k \geq n+1} \xi_k$ is symmetric, and independent of $\{\tau = n\}$ and S_n (since those depend only on $\xi_1, \ldots, \xi_n$. Now $|S_n| \leq \frac{1}{2}(|S_n + T| + |S_n - T|)$ so $\max(|S_n + T|, |S_n - T|) \geq |S_n|$. By symmetry $\{|S_n + T| \geq |S_n - T|\}$ has the same probability as $\{|S_n + T| \leq |S_n - T|\}$, and thus it has probability at least $\frac{1}{2}$. Therefore

$$P\{|S| > r, \tau = n\} \geq P\{|S| \geq |S_n|, \tau = n\} \geq \frac{1}{2} P\{\tau = n\}. \tag{11}$$

Summing over n results in $P\{S > r\} \geq \frac{1}{2} P\{\tau < \infty\} = \frac{1}{2} P\{\max|S_n| > r\}$. □

Each of the prior two theorems use a stopping-time argument to get an upper bound for a maximal inequality. The techniques is reminiscent of the approach used to prove the Calderón-Zygmund decomposition, though they concern the measures of certain sets rather than the values of integrals.

1.2 Convergence

As a warm up, let's use the Komolgorov inequality to show convergence for mean-zero series in $L^2(\Omega)$.

1.5 Proposition. *Suppose ξ_k are independent with mean 0. Then $\sum_k \xi_k$ converges a.s. if $\sum_k E\xi_k^2 < \infty$.*

Proof. Let $S_n = \sum_{k=1}^n \xi_k$. By theorem 1.3, $P(\sup_{k \geq n} |S_n - S_k| > r) < r^{-2} \sum_{k \geq n} E\xi_k^2$. Taking the limit $n \to \infty$, $P(\lim_{k\to\infty} \sup_{n \geq k} |S_n - S_k| > r) = 0$. Taking a sequence of $r \to 0$, ths becomes $P(\lim_{k\to\infty} \sup_{n\geq k} |S_n - S_k| > 0) = 0$. Therefore S_n satisfies a Cauchy criterion for convergence a.s.. □

Next is the main theorem which characterizes convergence of symmetric series.

1.6 Theorem (Series with symmetric terms)**.** *Suppose ξ_k are independent symmetric random variables. The following are equivalent (a) $\sum_k \xi_k$ converges a.s. (b)$\sum_k \xi_k^2$ converges a.s. (c) $\sum_k E\xi_k^2 \wedge 1 < \infty$*

Proof. Let $\eta_k = \xi_k 1_{\{|\xi_k| \leq 1\}}$ and assume $\sum E\eta_k^2 < \infty$. Then thoerem 1.5 implies that $\sum_k \eta_k$ converges a.s.. Also, $\sum_k P\{|\xi_k| > 1\} \leq \sum_k E\xi_k^2 \wedge 1 < \infty$ implies that $\sum_k 1\{|\xi_k| > 1\} < \infty$ a.s. (otherwise the expectation of this sum would be infinite). Thus ξ_k and η_k differ by at most finitely many terms a.s.. So $\sum_k \xi_k$ converges whenever $\sum_k \eta_k$ does, namely almost surely. Therefore (c) implies (a).

Conversely, assume $\sum E\eta_k^2 = \infty$. Then by thoerem 1.2, $P(|\sum_{k=1}^n \eta_k| > r(\sum_{k=1}^n E\eta_k^2)^{1/2}) \geq (1 - r^2)^2/3 > 0$. Letting $n \to \infty$ this means $\sum_{k=1}^n \eta_k = \pm\infty$ diverges with positive probability. By the zero-one law, $\sum_{k=1}^n \eta_k$ diverges a.s., and therefore so does $\sum_{k=1}^n \xi_k$. This shows that (a) implies (c).

Next, note that for positive series, deterministic or random, $\lim_n \sum_{k=1}^n a_k = \sup_n \sum_{k=1}^n a_k$ so showing convergence is same as showing $\sum_{k=1}^\infty a_k < \infty$. If $\sum E\xi^2 \wedge 1 < \infty$, then by Fubini's theorem $E \sum_k (\xi_k^2 \wedge 1) < \infty$, and therefore $\sum_k (\xi_k^2 \wedge 1) < \infty$ a.s.. Furthermore, $\sum_k 1\{\xi_k^2 > 1\} < \sum_k (\xi_k \wedge 1) < \infty$ a.s.. When the later sum is finite then $\xi_k < 1$ for only finitely many values of k. Replacing the finitely many η_k^2 with ξ_k^2 does not affect convergence, so $\sum_k \xi_k^2 < \infty$ a.s.. This shows (c) implies (b).

Finally, to see that (b) implies (c), compare the series term-by-term to conclude $\sum(\xi_k^2 \wedge 1) \leq \sum \xi_k^2 < \infty$. So, replacing ξ_k^2 with $\xi_k^2 \wedge 1$ if necessary, assume without loss of generality assume that $\xi_k^2 \leq 1$ for all k. In this case $1 - \xi_k^2 \leq e^{-\xi_k^2} \leq 1 - a\xi_k^2$ where $a = 1 - e^{-1}$. Therefore

$$0 < E\exp\left(-\sum_k \xi_k^2\right) = \prod_k Ee^{-\xi_k^2} \leq \prod_k (1 - aE\xi_k^2) \leq \prod_k e^{-aE\xi_k^2} = \exp\left(-a\sum_k E\xi_k^2\right) \tag{12}$$

which shows $\sum_k E\xi_k^2 < \infty$.

$\square$

While its beyond the scope of this essay, the convergence of random series general random variables can be analyzed by symmetrization. If ξ_k are independent random variables (not necessarily symmetric), then take ξ'_k to be random variables with the same distribution as ξ_k. the random variables $\xi_k - \xi'_k$ are symmetric. This type of reasoning leads to the Komologorov three-series theorem which characterizes convergence in general. See [3] for more details.

Next we consider summability methods for random series. Summability matrices are key for analyzing the pointwise convergence properties of Fourier series. However, unlike for deterministic series, summability does not improve the convergence characteristics of random symmetric series.

Definition. A *summation matrix* (a_{mn}) is any series of numbers satisfying $\lim_{m\to\infty} a_{mn} = \infty$. Given a series $\sum_n v_n$ consider the series $w_m = \sum_n a_{mn} v_n$. If each w_m and $w = \lim_m w_m$ exists, then the series is said to be *a-summable* and w is the *a-sum*. If w_n is bounded then the series is *a-bounded*

Example. The Césaro method uses the summation matrix $a_{mn} = 0 \wedge (1 - \frac{n}{m})$. The Abel-Poisson method uses the summation matrix $a_{mn} = r_m^n$. When the $a_{mn} = 1_{n\leq m}$, the w_m are just the partial sums.

1.7 Proposition. *Let ξ_k be independent symmetric random elements and let a_{mn} be a summation matrix. If the series $\sum_n \xi_n$ is a.s. a-summable, it converges a.s..*

Proof. Our first task is to choose a matrix b_{mn} which resembles the partial sums such that the series is b-summable. Choose a sequence $\varepsilon_p \downarrow 0$. Because $\lim_m a_{mn} \to 1$, for fixed p and large enough m,

$$P\left(\left|\sum_{n\leq p}(1 - a_{mn})\xi_n\right| > \varepsilon_p\right) < \varepsilon_p \tag{13}$$

Say this happens for $m > m_p$. Furthermore, since $\sum_{m_p n} \xi_n$ converges a.s., the tail converges to 0 in probability, so for some q_p

$$P\left(\left|\sum_{n\geq q_p} a_{mn}\xi_n\right| > \varepsilon_p\right) < \varepsilon_p \tag{14}$$

Therefore define a new summation matrix b_{pn} by $b_p n = 1$ if $n \leq p$ and $b_{pn} = 0$ if $n > q_p$ and $b_{pn} = a_{m_p p}$ for the values $p < m \leq q_p$. Therefore

$$P\left(\left|\sum_n (b_{pn} - a_{m_p n})\xi_n\right| > \varepsilon_p\right) < 2\varepsilon_p \tag{15}$$

So if the series is a-summable, its also b-summable.

Next we modify the terms of ξ_k without affecting convergence. Let $s_n = \pm 1$ be any deterministic choice of signs. Because the ξ_n are symmetric, the random variables $\xi'_n = s_n \xi_n$ have the same joint distribution as ξ_n, and therefore $\sum_n s_n\xi_n$ has the same probability of converging. If both series converge, then so do $\frac{1}{2}\sum_n(\xi_n + \xi'_n)$ and $\frac{1}{2}\sum_n(\xi_n - \xi'_n)$. Conversely, if adjusted series converge, then so does $\sum_n \xi_n$. However, $\xi_n + \xi'_n = 0$ for all n where $s_n =$. By these considerations, we can "zero out" an infinite number of terms without affecting convergence. More precisely, given $I \subset \mathbb{N}$, if $\sum_{n\in I} \xi_n$ and $\sum_{n\in I^c} \xi_n$ converge, then so does the original series.

Recursively define a sequence of indices $p_1 = 1$ and $p_{j+1} = q_{p_j}$, and suppose that $\xi_k = 0$ for $p_j < k \leq q_{p_j} = p_{j+1}$. Then for this series, $\sum_n b_{p_j n}\xi_n = \sum_{n=1}^{p_{j+1}} \xi_n$. In other words, the b-sum for p_j is just the same as the partial sum $\sum_{n=1}^{p_j} \xi_n$. So let I be the union of an infinite collection of the intervals $\{p_j + 1, \ldots, p_{j+1}\}$ and let I^c be the same.

From what we've previously shown, we know $\eta_j = \sum_{n=1}^{p_j} \xi_n$ is a sequence which converges a.s.. Now we use 1.4 to extend over convergence for all partial sums. Since the η_j converge almost surely, they converge in probability, so given $\varepsilon > 0$ its possible to find an index $j(\varepsilon)$ such that for $k > j$

$$P(|\eta_k - \eta_j| > \varepsilon) = P\left(\left|\sum_{p_j < n \leq p_k}\right| > \varepsilon\right) < \varepsilon \tag{16}$$

By theorem 1.4, this means

$$P\left(\sup_{p_j < l \leq p_k}\left|\sum_{p_j < n \leq l} \xi_k\right| > \varepsilon\right) < 2\varepsilon \tag{17}$$

Given ε, let $j_0 = j$ and for $\nu = 1, 2, \ldots$, choose $j_\nu = j(\varepsilon 2^{-\nu})$ and $k_\nu = j_{\nu+1}$. Summing over inequalities like the above, we get

$$P\left(\sup_{p_j < l}\left|\sum_{p_j < n \leq l} \xi_n\right| > \varepsilon\right) < 2\varepsilon \sum_{\nu=0}^{\infty} 2^{-\nu} = 4\varepsilon \tag{18}$$

Since we can make this probability as small as desired, this shows that the partial sums converge. □

1.8 Corollary. *Let be ξ_n symmetric and independent. If $\xi_1 + \cdots + \xi_n$ converges in probability, it converges almost surely.*

Proof. If a series converges in probability, there is a subsequence which converges almost surely. But then the reasoning in the previous proof shows that the series converges almost surely. □

1.3 Rademacher functions

Rademacher functions $R = \sum_k \epsilon_k u_k$ have a particularly simple convergence properties. By 1.6, they converge if and only if $ER^2 = \sum_k u^2 < \infty$.

1.9 Lemma (Large deviations). *Let $S_n = \sum_k \epsilon_k u_k$ and $S = \sum_k \epsilon_k u_k$. If $P(|S| > r) < \lambda$, then $P(|S| > 2r) < 4\lambda^2$. Let $M = \sup_n |S_n|$. If $P(M < r) < \lambda$ then $P(M < 2r) < 2\lambda^2$*

Proof. Let $\tau = \inf\{n : |S_n| > r\}$. Let $C_n = \{|S - S_n| > r\}$. If $|S_{n-1}| < r$ and $|S| > 2r$, then $|S - S_n| > r$. Furthermore, the event $\{|S - S_n| > r\}$ is independent of the event $\{\tau = n\}$. This is because $\epsilon_1, \ldots, \epsilon_n$ is independent of $\epsilon_n\epsilon_{n+1}, \epsilon_n\epsilon_{n+2}, \epsilon_n\epsilon_{n+3}, \ldots$ and $|S - S_n|$ depends only on these values. Therefore

$$P\{|S| > 2r \text{ and } \tau = n\} \leq P\{|S - S_n| > r \text{ and } \tau = n\} = P\{|S - S_n| > r\}P\{\tau = n\}$$

Summing over n gives

$$P\{|S| > 2r\} < \sum_n P(\tau = n)P(|S - S_n| > r) \leq P\{\tau < \infty\} \sup_n P\{|S - S_n| > r\}.$$

By proposition 1.4, $P\{\tau < \infty\} = P(M > r) \leq 2\lambda$. Since the terms $\epsilon_1 u_1, \ldots, \xi_n u_n$ are symmetric, an argument similar to 1.4 shows $P(|S - S_n| > r) \leq 2P(|S| > r) = 2\lambda$. Therefore $P\{|S| > 2r\} < 4\lambda^2$.

The argument for M is similar. Let $D_n = \{\sup_{m \geq n} |S_m - S_n| > r\}$ which is independent of the event $\tau = n$. Since $M > 2r$ and $\tau = n$ implies $D_n > r$, $P\{M > 2r\} \leq P\{\tau = n\}PD_n$. Summing over n gives $P\{M > 2r\} \leq P\{M > r\} \sup_n P\, D_n$. Now again by symmetry, $PD_m \leq 2P\{M > r\}$, so $P\{M > 2r\} < 2\lambda^2$. □

1.10 Proposition. *The Rademacher series $R = \sum_k \epsilon_k u_k$ converges a.s.. iff $\xi \in L^p(\Omega)$ for any $p \in [1, \infty)$. Moreover, $\exp(\alpha R) \in L^1(\Omega)$ for suitable $\alpha > 0$.*

Proof. Let ϕ be a positive monotonic function. Let $p(t) = P(|R| > t)$. Then $E\phi(|R|) = -\int_0^\infty \phi(t)\,dp(t)$. Suppose $p(r) < \lambda$. Therefore, divide $[0,\infty)$ into the disjoint segments $I_0 = [0,r)$ and $I_k = [2^k r, 2^{k+1} r)$, and calculate, using

$$E\phi(|R|) = -\sum_k \int_{I_k} \phi(|R|)\,dp(t) \le \sum_k \phi(2^{k+1}r)p(2^k r) \le \frac{1}{4}\sum_k \phi(2^{k+1}r)(2\lambda)^{2^k}$$

So from this point its just a matter of testing various ϕ for suitable choices of λ and r. In particular, $\phi(x) = e^{\alpha x}$, if $\gamma = 2\alpha r + \log(2\lambda) < 0$ then $Ee^{\alpha|R|} < \sum_k e^{2^k\gamma} < \infty$.

For $\phi(x) = |x|^p$, one can do a similar calculation, or note that $E|R|^p \le 1 + E|R|^p \vee 1 \le 1 + E|R|^k \vee 1$ where $k = \lceil p \rceil$. Marjorizing $|R|^k$ by the term in the taylor expansion of $e^{\alpha|R|}$ gives $E|R|^k \vee 1 \le k!\,\alpha^{-k} E\exp(\alpha R \vee 1) < \infty$ since $E\exp(\alpha R \vee 1) \le e^a + E\exp(\alpha R) < \infty$ □

By considering R as the sum of a finite bounded function $\sum_{k\le n} \epsilon_k u_k$ and the tail $\sum_{k>n} \epsilon_k u_k$ its possible to show $e^{\lambda R} \in L^1(\Omega)$ for any $\alpha > 0$. More strongly, its possible to show $e^{\lambda^2 R^2} \in L^1(\Omega)$. This property is shared by Gaussian random variables, so random variables with this property are called sub-Gaussian. Even more can be said about the $L^p(\Omega)$ properties of R.

1.11 Proposition (Khinchine's Inequality). *Let $\xi = \sum_k \epsilon_k u_k$ be a Rademacher series. Then $\|\xi\|_{L^p} \le c_{pq}\|\xi\|_{L^q}$ where the constant c_{pq} depends only on p and q and not on ξ.*

Proof. Hōlder's inequality gives $\|\xi\|_{L^1} \le \|\xi\|_{L^q}\|1\|_{L^r} = \|\xi\|_{L^q}$ where $\frac{1}{q} + \frac{1}{r} = 1$, so it suffices to show the case when $p > 1$ and $q = 1$. Scaling ξ by a constant if necessary, assume $\|\xi\|_{L^1} = 1$.

Let $j \ge 1$ be the unique integer $2^{j-1} < p \le 2^j$. By Markov's inequality, $rP\{|\xi| > r\} \le E|\xi| = 1$, so

$$\begin{aligned}
E|\xi|^p &= \int_0^\infty pt^{p-1}P\{|\xi| > t\}\,dt \le 2^{jp}\int_0^\infty pr^{p-1}P\{|\xi| > 2^j r\}\,dr \\
&\le 2^{jp}4^{2p-1}\int_0^\infty pr^{p-1}(P\{|\xi| > r\})^{2^j}\,dr \\
&\le (2p)^p 4^{2p-1}\int_0^\infty pr^{p-1}(P\{|\xi| > r\})^p \\
&\le (2p)^p 4^{2p-1} p\int_0^\infty P\{|\xi| > r\} = (2p)^p 4^{2p-1} p
\end{aligned}$$

□

Much work has been done to determine the optimal constant $C_{p,q}$. For $p \le q$ the best possible is $C_{p,q} = 1$, the argument above is far from optimal for other values. See [4] for a full accounting.

1.4 Contraction

Generally, its difficult to put precise conditions on when a Rademacher series converges almost surely or is almost surely bounded. Paley and Zygmund were only able to show that $\sum_k u_k^2(\log n)^{1+\varepsilon} < \infty$ is a sufficient condition, though this has been refined over time.

Instead, of explicit criterion, its often possible to proceed by comparison. The following shows that shrinking the Rademacher coefficients only improves the convergence behavior of the function.

1.12 Proposition (Contraction). *Let λ_n be a bounded sequence of complex numbers. Let $R = \sum_k \epsilon_k u_n$ and $R' = \sum_k \lambda_k \epsilon_k u_k$. If R converges a.s. so does R'. If R is bounded then so is R'. If $R \in L^p(\Omega)$ then $R' \in L^p(\Omega)$*

Proof. First, let's assume λ_n is real and $\lambda_n \in [0,1]$. To begin with by applying the theorem with the multipliers $\operatorname{Re}\lambda_n$ and $\operatorname{Im}\lambda_n$ extends it to the complex case, and applying the theorem to $\lambda_n/\sup|\lambda_n|$ to the case when $\sup|\lambda_n| > 1$.

Suppose $\lambda_n \in \{0,1\}$. Then $\sum_k \lambda_k\epsilon_k u_k = \frac{1}{2}(\sum_k \epsilon_k u_k + (2\lambda_n - 1)\epsilon_k u_k)$ and, by the symmetry of ϵ_k each of the terms in the parenthesis have the same distribution. Thus the convergence of R implies the convergence of R'. Let $M' = \sup_n |\sum_{k=1}^n \lambda_k \epsilon_k u_k|$ be the maximum of the partial sums of R'. By 1.10 we know $M \in L^p(\Omega)$, so this equation implies $EM' \leq EM$, and therefore $M < \infty$ a.s..

For general λ_k, expand each in binary to get $\lambda_k = \sum_{n\geq 1} 2^{-k}\lambda_{nk}$ where $\lambda_{nk} \in \{0,1\}$ and note $R' = \sum_n 2^{-k} \sum_k \lambda_{nk}\epsilon_k u_k$.

This approach of representing R' as an average of series with the same distribution as R also shows that $R' \in L^p(\Omega)$ whenever R is, because $\|\sum_k \alpha_k \xi\|_p \leq \sum_k \alpha_k \|\xi_p\|$. □

1.13 Proposition. *Let $R = \sum_k \epsilon_k u_k$ be the Rademacher series and $S = \sum_k e^{in\omega_n} u_k$ be the Steinhaus series. The series S converges a.s. iff R converges a.s.. The series S is bounded a.s. iff R is bounded.*

Proof. Let $T = \sum_k \epsilon_k e^{i\omega_k} u_k$. If R converges, then by the conditioning principle, for each value of ω_n, we can apply proposition 1.12 with $\lambda_n = e^{i\omega_n}$. This shows that T converges almost surely. But T and S have the same distribution since S is symmetric. If one converges then the other one does to. The converse applies this argument in reverse with $\lambda_n = e^{-i\omega_n}$. The argument for when S or R is bounded is the same. □

2 Paley-Zygmund Theorem

Here we collect a few more basic results on Fourier series.

2.1 Lemma. *If $\sum_n a_n^2 = \infty$ then $\sum_n a_n^2 \cos^2(nt + \phi_n) = \infty$ almost everywhere*

Proof. If the conclusion is false, then $\sum_n a_n^2 \cos^2(nt+\phi_n) < b$ on some set E with $|E| > 0$. Therefore $\int_E \sum_n a_n^2 \cos^2(nt+\phi_n) < b|E|$. On the other hand $\int_E \cos^2(nt+\phi_n) \to \frac{1}{2}|E|$ by the Reimann-Lesbegue theorem. Therefore for n large enough, say $n > n_0$, $\int_E \cos^2(nt+\phi_n) > \frac{1}{3}|E|$. But then we have $\sum_{k\geq n_0} a_n^2 \leq \frac{1}{3} b|E|$ contrary to assumption. □

This is a basic result on the convergence of Fourier series, for a proof see [5] theorem 2.10.

2.2 Theorem. *Let $f \in L^1(\mathbb{T})$ and let $S = \sum_n \hat{f}(n)e^{inx}$ be its Fourier series. For almost every point $x \in \mathbb{T}$, the series S is Césaro-summable and Poisson-summable and the sum is $f(x)$.*

Next is our first interesting result on random Fourier series. It immediately gives a non-constructive example of a Fourier series which does not represent a $L^1(\mathbb{T})$ function. Almost all choices for $\hat{f}(n) = \pm 1/\sqrt{n}$ will due, though any particular choice may not.

2.3 Proposition. *Let $f(t) = \sum_k \epsilon_k a_k \cos(nt + \phi_k)$. If $\sum_k a_k^2 = \infty$ then almost surely $f(t)$ does not represent the Fourier series any function in L^1.*

Proof. By lemma 2.1, $Ef(t)^2 = \sum_k a_k^2 \cos^2(nt+\phi_k)\infty$ for almost all t. Thus by theorem 1.6, $f(t)$ diverges almost surely, almost everywhere. In particular, by 1.7, $f(t)$ is not Poisson-summable almost surely almost everywhere. Therefore, $f(t)$ does not represent the Fourier transform of any measure. □

As p increases, $L^p(\mathbb{T})$ gets smaller. However, convergent random Fourier series are in all $L^p(\mathbb{T})$ for all $p \in [1,\infty)$. So long as $(a_k) \in \ell^2$, the Rademacher series in the following proposition gives a non-constructive proof of a Fourier series which converges to a function in $\cap_p L^p(\mathbb{T})$ a.s..

2.4 Proposition. *Let $f(t) = \sum_k \epsilon_k a_k \cos(nt+\phi_k)$. If $s = \sum_k a_k^2 < \infty$ then $\int_0^{2\pi} e^{\lambda f(t)^2} < \infty$ a.s.. Consequently, $f(t) \in L^p(\mathbb{T})$ for $p \in [1,\infty)$*

Proof. Let $b_k(t) = a_k \cos(nt+\phi_k)$ so $f(t)$ is a Rademacher series with terms $b_k(t)$. Note

$$Ee^{\alpha f(t)} = E\exp\left(\alpha \sum_k b_k(t)\epsilon_k\right) = \prod_k E\exp(\alpha b_k(t)\epsilon_k) = \prod_k \cosh(\alpha b_k(t)) \quad (19)$$

Now $\cosh(\alpha x) \le e^{\alpha^2 x^2/2}$ (this can be verified by looking at the Taylor series expansions of each). Therefore, putting

$$Ee^{\alpha f(t)} \le \prod_k e^{\alpha^2 b_k(t)^2/2} \le e^{\alpha^2 s/2} \tag{20}$$

This is enough to show that $e^{\alpha f(t)} \in L^1(\mathbb{T})$, but we will now strengthen the inequality.

Let $m_k = Ef(t)^k$. Now $m_k = 0$ for k odd, by symmetry, and m_k is dominated by the corresponding term in the Taylor series for k even

$$m_{2k} \le \frac{(2k)!}{\alpha^k} Ee^{\alpha f(t)} \le \frac{(2k)!}{\alpha^k} e^{\alpha^2 s/2} \tag{21}$$

we minimize this term by choosing $\alpha^2 = 2k/r$ to find $m_{2k} \le Ck!(2r)^n$. Therefore for $\lambda < 1/(2r)$

$$Ee^{\lambda f(t)^2} = \sum_k \frac{\lambda^n}{n!} m_{2n} \le C \sum_k (2\lambda r)^n = b < \infty \tag{22}$$

In fact for arbitrary $\lambda > 0$, we could instead make this argument starting with $\tilde{f}(t) = \sum_{k \ge n_0} b_k(t)\epsilon_k$ where n_0 is chosen so that $2\lambda \sum_{k \ge n_0} a_k^2 < 1$. This would result in the conclusion that $Ee^{\alpha \tilde{f}^2(t)}$ is almost surely bounded. Then since $f(t) = \tilde{f}(t) + \sum_{k<n_0} b_k(t)\epsilon_k$ and the latter terms are bounded, the property holds for f as well.

Therefore $Ee^{\lambda f^2(t)} \le b < \infty$ almost everywhere so $\int_0^{2\pi} Ee^{\lambda f^2(t)} \le 2\pi b < \infty$. In particular, since $e^{\lambda f^2(t)}$ is positive, this means that $e^{\lambda f^2(t)} < \infty$ everywhere. □

We come now to the main result. Note that the fact that the statement that $S(t) \in L^p(\mathbb{T})$ a.s. does not necessarily imply that $S(t) \in L^p(\Omega)$. While $\|S(t)\|_p < \infty$ a.s., it may be that $E\|S(t)\|_p = \infty$.

2.5 Theorem (Paley-Zygmund). *Let ξ_k be real valued symmetric random variables and let ϕ_n be random variables on $\mathbb{T}$. Let $S(t) = \sum_k \xi_k e^{i\phi_n} e^{int}$.Suppose $\sum E\xi_k^2 \wedge 1 < \infty$. Then $S(t)$ converges a.s., a.e. in $\mathbb{T}$. Furthermore, $S(t) \in L^p(\mathbb{T})$ a.s., for $p \in [1, \infty)$. On the other hand, if $\sum \xi_k^2 \wedge 1 = \infty$ then a.s. $S(t)$ is not the Fourier series of any function in $L^1(\mathbb{T})$.*

Proof. These equivalences follow from conditioning and contraction. First assume that $\phi_n = 0$ and write $\xi = \epsilon_k \eta_k$ where $\eta_k \ge 0$ and ϵ_k, η_k are independent. For fixed values of η_k, the stated properties hold a.s. by 1.6, 2.4, 2.3. For example, to apply 2.4 we must know that $\sum_k \eta_k^2 < \infty$, but guaranteed by 1.6. For the case when $\phi_n \ne 0$, condition on the values of ϕ_n and use 1.12. □

For comparison, I state a related result without proof

2.6 Theorem (Billard). *Let $S(t) = \sum_k \xi_k e^{i\phi_n} e^{int}$. Then the following have the same probability (0 or 1): (a) $S(t) \in L^\infty(\mathbb{T})$, (b) $S(t) \in C$, (c) $S(t)$ converges uniformly.*

Extensions and Discussion

The main theorems of this essay explore the relationships of various modes of convergence for random Fourier series. In the case of a.s. convergence, the qualitative results were paired with a quantitative criterion (namely, $\sum_k E\xi_k^2 \wedge 1 < \infty$). This was not the case for a.s. uniform convergence. The various sufficient conditions were given in [1] and [2], and a precise necessary and sufficient condition was found in [6].

The case of non-symmetric random variables is taken up by various authors. In [7], Cuzick explores the case of scaled independent identical random variables $\xi_k = \eta_k u_k$ where the η_k are iid. Convergence conditions are related to the asymptotic behavior of a_k and the tail distribution of X_k. In [8], Talagrand considers $\xi_k = \eta_k u_k$ where the η_k are iid with mean zero to find a fairly simple condition for a.s. uniform convergence. Finally, in [9]Cohen shows that a.s. uniform convergence is not equivalent to a.s. boundedness.

I'd like to thank Steven Evans for some discussions while I was preparing this essay.

References

1. Raymond Paley and Antoni Zugmund. On some series of functions i. *Proc. Camb. Phil. Soc.*, 26:337–57, 1930.

2. Jean-Pierre Kahane. Some random series of functions. Cambridge University Press, 2nd edition, 1985.

3. Olav Kallenberg. *Foundations of Modern Probability*. Springer Science & Business Media, 2nd edition, 2002.

4. Uffe Haagerup. The best constants in the khintchine inequality. *Studia Mathematica*, 70(3):231–283, 1981.

5. Juan Arias de Renya. *Pointwise convergence of fourier series*, volume 1785 of *Lecture Notes in Mathematics*. Springer–Verlag Berline Heidelberg, 2002.

6. Micheal B Marcus and Gilles Pisier. *Random Fourier Series with Applications to Harmonic Analysis*. Princeton University Press, 1981.

7. Jack Cuzick and Tze Leung Lai. On random fourier series. *Transactions of the American Mathematical Society*, 261(1):53–80, 1980.

8. Michel Talagrand et al. A borderline random fourier series. *The Annals of Probability*, 23(2):776–785, 1995.

9. Guy Cohen and Christophe Cuny. On billard's theorem for random fourier series. *Bull. Pol. Acad. Sci. Math*, 53:39–53, 2005.