A combinatorial view of the orthogonality of Rademacher functions

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

December 11, 2016

1 Introduction

The first chaper of Mark Kac's deservedly famous book [5] explores in a fundamental way how independence lurks in the binary representation of real numbers t in [0,1]. It was Borel who researched the fundamental probabilistic aspects of pure number theory circa 1909 [2] and in so doing came up with the concept of "normal" numbers. In the context of the binary representation of numbers on [0,1] this means that there is an equal number of 0s and 1s in the binary representation.

Kac's book is like a finely crafted degustation menu - there are many delicious courses one could consume, but for the purposes of this article I only want to focus on one issue which strikingly exposes the independence angle. What Kac described can be summarised in part as follows. We have that for any t in [0,1] there is a unique binary representation thus:

$$t = \frac{\epsilon_1(t)}{2} + \frac{\epsilon_2(t)}{2^2} + \dots \tag{1}$$

where $\epsilon_k(t) = 0$ or 1 for all k

The $\epsilon_k(t)$ are step functions and $\epsilon_1(t)$, for instance, is set out in Figure 1.

The introduction of the Rademacher functions $r_k(t)$, which are defined as follows, enables us to develop some rather surprising relationships out of very little:

$$r_k(t) = 1 - 2\epsilon_k(t) \tag{2}$$

for $k = 1, 2, 3, \dots$

Actually, Rademacher defined his functions as $-r_k(t)$ (see [8], page 130).

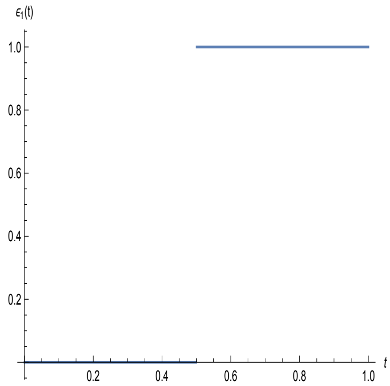

Figure 1

Using (1) and (2) the following relationship exists:

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k} \tag{3}$$

For each k the Rademacher functions are step functions which alternate between +1 and −1. . For instance, $r_1(t)$ is set out in Fugure 2.

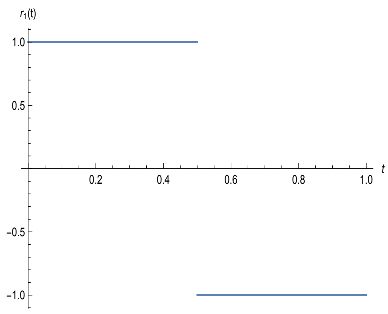

Figure 2

There are various ways to represent $r_n(t)$. For instance one can define the functions as follows for $n = 1, 2, \ldots$:

$$\begin{aligned} r_1(t) = 1 \quad (0 \leq t < \frac{1}{2}) \\ r_1(t) = -1 \quad (\frac{1}{2} \leq t < 1) \\ r_1(t+1) = r_1(t) \end{aligned} \tag{4}$$

One can exploit the periodicity of these functions to write the n^{th} function in terms of the first as follows:

$$r_n(t) = r_1(2^{n-1}t) \tag{5}$$

We could also give a formula for the general Rademacher function as follows:

$$r_n(t) = \sum_{k=1}^{2^n} (-1)^{k-1} \mathcal{X}_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(t) \tag{6}$$

where $\mathcal{X}_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(t) = 1$ if $t \in [\frac{k-1}{2^n}, \frac{k}{2^n})$ and 0 if $t \notin [\frac{k-1}{2^n}, \frac{k}{2^n})$

For instance, $r_1(t) = (+1)\mathcal{X}_{[0,\frac{1}{2})}(t) + (-1)\mathcal{X}_{[\frac{1}{2},1)}(t)$.

In what follows we stick with the basic graphical rerpresentation employed by Kac.

It is clear that the integral $\int_0^1 r_n(t)\, dt = 0$ for all n. There is an even number of $+1$ steps and and even number of -1 steps with equal bases hence the integral is zero. This generalises as follows for $i_1 < i_2 < \cdots < i_n$:

$$\int_0^1 r_{i_1}(t)\, r_{i_2}(t) \dots r_{i_n}(t)\, dt = 0 \tag{7}$$

This is a problem in Kac's book ([2], page 11, Problem 3).

2 The dyadic case

An obvious starting point to establish is the dyadic case and then use induction or some other technique to deal with the more general case. Zygmund in his famous work on trigonometric series does not prove the general case but he deals with the dyadic case as follows ([3], page 6). Because the $r_n(t)$ alternate between ± 1 inside the intervals $(0, 2^{-n-1})$, $(2^{-n-1}, 2 \cdot 2^{-n-1})$,$(2 \cdot 2^{-n-1}, 3 \cdot 2^{-n-1})$, ... so that if $m > n$ then the integral of $r_m(t)\, r_n(t)$ over any of these intervals is zero. Zygmund uses the periodicity rule:

$$r_n(t) = r_0(2^n t) \tag{8}$$

The proof of Zygmund's claim is as follows. Note that Zygmund takes $r_n(t)$ to be defined with $r_0(t)$ as the starting point ie:

$$\begin{aligned} r_0(t) &= 1 \quad (0 \le t < \frac{1}{2}) \\ r_0(t) &= -1 \quad (\frac{1}{2} \le t < 1) \\ r_0(t+1) &= r_0(t) \\ r_n(t) &= r_0(2^n\, t) \end{aligned} \tag{9}$$

Choose $m > n$. Then:

$$\begin{aligned}
\int_0^1 r_m(t)\, r_n(t)\, dt =& \sum_{k=1}^{2^m} \int_{\frac{k-1}{2^m}}^{\frac{k}{2^m}} r_m(t)\, r_n(t)\, dt \\
=& \sum_{k=1}^{2^m} \int_{\frac{k-1}{2^m}}^{\frac{k}{2^m}} (-1)^{k+1} r_n(t)\, dt \\
=& \sum_{k=1}^{2^m} \int_{\frac{k-1}{2^m}}^{\frac{k}{2^m}} (-1)^{k+1} r_0(2^n t)\, dt \\
=& \frac{1}{2^n} \sum_{k=1}^{2^m} \int_{\frac{k-1}{2^{m-n}}}^{\frac{k}{2^{m-n}}} (-1)^{k+1} r_0(u)\, du \text{ since } u = 2^n t \\
=& \frac{1}{2^n} \Big\{ \sum_{k=1}^{2^{m-n-1}} \int_{\frac{k-1}{2^{m-n}}}^{\frac{k}{2^{m-n}}} (-1)^{k+1} r_0(u)\, du + \sum_{k=2^{m-n-1}+1}^{2^m} \int_{\frac{k-1}{2^{m-n}}}^{\frac{k}{2^{m-n}}} (-1)^{k+1} r_0(u)\, du \Big\} \\
=& \frac{1}{2^n} \Big\{ \sum_{k=1}^{2^{m-n-1}} (-1)^{k+1} \frac{1}{2^{m-n}} + \sum_{k=2^{m-n-1}+1}^{2^m} (-1)^{k+1} \frac{-1}{2^{m-n}} \Big\} \\
=& \frac{1}{2^m} \Big\{ \sum_{k=1}^{2^{m-n-1}} (-1)^{k+1} - \sum_{k=2^{m-n-1}+1}^{2^m} (-1)^{k+1} \Big\} \\
=& \frac{1}{2^m} \Big\{ 0 - 0 \Big\} \\
=& 0
\end{aligned} \tag{10}$$

The integrations are merely sums of rectangles of width $\frac{1}{2^{m-n}}$ and because the domain of integration is broken up into $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$ in each sum there is an equal number of $+1$s and $-1s$ so the overall result is zero. The interchange of summation and integration works because we are dealing with step functions.

A more succinct and less mechanical proof is as follows (see [6], pages 42 -42). We suppose $m > n$ and construct intervals $I_k = [\frac{k-1}{2^n}, \frac{k}{2^n})$ for $k = 1, 2, \ldots, 2^n$. Then:

$$\int_0^1 r_m(t) r_n(t)\, dt = \sum_{k=1}^{2^n} \int_{I_k} r_m(t) r_n(t)\, dt \tag{11}$$

Because for each I_k, $r_n(t) = (-1)^{k+1}$ we have the following relationship:

$$\int_{I_k} r_m(t) r_n(t)\, dt = (-1)^{k+1} \int_{I_k} r_m(t)\, dt \tag{12}$$

Now in each I_k there are 2^{m-n} subintervals of width $\frac{1}{2^m}$ in which $r_m(t)$ alternates in sign from $+1$ to -1. Hence $\int_{I_k} r_m(t)\, dt = 0$ for all k. Thus the RHS of (11) is zero.

One can prove the dyadic case using a combinatorial style of approach. The function $u_n(t)$ has 2^{n-1} "-1" subintervals and 2^{n-1} "+1" subintervals (ie the heights are "-1" and "+1").
Case 1: Take the "-1" subintervals of $u_n(t)$ first. For each of these subintervals there are 2^{m-n} subintervals of $u_m(t)$ each of which has width $\frac{1}{2^m}$. Half of these subintervals are "-1". Hence the 2^{n-1} "-1" subintervals of $u_n(t)$ correspond to $2^{n-1} \times 2^{m-n-1} = 2^{m-2}$ "-1" subintervals of $u_m(t)$ which have width $\frac{1}{2^m}$. Similarly for our choice of "-1" subintervals of $u_n(t)$ there are symmetrically 2^{m-2} "+1" subintervals of $u_m(t)$ which have width $\frac{1}{2^m}$.
Case 2: Symmetrically, the "+1" subintervals of $u_n(t)$ also give rise to the same logic and for each of these subintervals we get the same number of $+1$ and -1 subintervals of $u_m(t)$ of width $\frac{1}{2^m}$.

Thus in the integral of the product we simply have equal numbers of rectangles with common width $\frac{1}{2^m}$ having $+1$ and -1 heights so the integral is zero . This result generalises to an integral of an arbitrary product. However, for reasons which will shortly become clearer it is better to work with the behaviour of the "epsilon" functions which, from a notational point of view, makes the combinatorial angle quite clear.

Implicit in the above approach is that each $r_k(t)$ is viewed as a finite sum of indicator functions of the form $\mathcal{X}_{[a,b)}(t)$ so that the resultant product of terms $r_{i_1}(t)\, r_{i_2}(t) \dots r_{i_n}(t)$ gives rise to an expression which depends essentially on the rule for the component terms:

$$\mathcal{X}_{A\cap B} = \mathcal{X}_A\, \mathcal{X}_B \tag{13}$$

Because of the intersection property, the smallest sub-interval (ie that with the highest index) will govern the result. Intersections of disjoint subintervals are "annihilated". This can be seen tediously if we write out the functions in detail, for example, as follows:

$$\begin{aligned}
r_1(t) \times r_2(t) &= \Big\{ \sum_{k=1}^{2} (-1)^{k-1} \mathcal{X}_{[\frac{k-1}{2},\frac{k}{2})}(t) \Big\} \times \Big\{ \sum_{k=1}^{4} (-1)^{k-1} \mathcal{X}_{[\frac{k-1}{4},\frac{k}{4})}(t) \Big\} \\
&= \Big\{ \mathcal{X}_{[0,\frac{1}{2})}(t) - \mathcal{X}_{[\frac{1}{2},1)}(t) \Big\} \times \Big\{ \mathcal{X}_{[0,\frac{1}{4})}(t) - \mathcal{X}_{[\frac{1}{4},\frac{1}{2})}(t) + \mathcal{X}_{[\frac{1}{2},\frac{3}{4})}(t) - \mathcal{X}_{[\frac{3}{4},1)}(t) \Big\} \\
&= \mathcal{X}_{[0,\frac{1}{2})}(t)\, \mathcal{X}_{[0,\frac{1}{4})}(t) - \mathcal{X}_{[0,\frac{1}{2})}(t)\, \mathcal{X}_{[\frac{1}{4},\frac{1}{2})}(t) + \mathcal{X}_{[0,\frac{1}{2})}(t)\, \mathcal{X}_{[\frac{1}{2},\frac{3}{4})}(t) - \mathcal{X}_{[0,\frac{1}{2})}(t)\, \mathcal{X}_{[\frac{3}{4},1)}(t) - \mathcal{X}_{[\frac{1}{2},1)}(t)\, \mathcal{X}_{[0,\frac{1}{4})}(t) + \\
&\quad \mathcal{X}_{[\frac{1}{2},1)}(t)\, \mathcal{X}_{[\frac{1}{4},\frac{1}{2})}(t) - \mathcal{X}_{[\frac{1}{2},1)}(t)\, \mathcal{X}_{[\frac{1}{2},\frac{3}{4})}(t) + \mathcal{X}_{[\frac{1}{2},1)}(t)\, \mathcal{X}_{[\frac{3}{4},1)}(t) \\
&= \mathcal{X}_{[0,\frac{1}{4})}(t) - \mathcal{X}_{[\frac{1}{4},\frac{1}{2})}(t) - \mathcal{X}_{[\frac{1}{2},\frac{3}{4})}(t) + \mathcal{X}_{[\frac{3}{4},1)}(t)
\end{aligned} \tag{14}$$

It is easier to see this result if we use as the common "base" for each step function that of $r_2(t)$ ie an interval width of $\frac{1}{4}$ and we write a " $+$ " for $+1$ and a " $-$ " for -1 on these intervals. The product $r_1(t) \times r_2(t)$ then looks like this:

r_1	+	+	-	-
r_2	+	-	+	-
$r_1\, r_2$	+	-	-	+

Signs are multiplied in each column (because each column represents the common base interval for which the function is either $+1$ or -1) and this result is the same as (14).

Rademacher does not prove (7) in his orginal paper [8]. One proof relies upon the periodicty of Rademacher functions (see [3] , p.185) while others employ a basic integration argument. For instance Kaczmarz and Steinhaus (Kac's PhD supervisor) prove the result this way (see [6], pages 236-237). There is a typo in the original paper with the right hand interval point being stated as $\frac{r+1}{2^{(l+1)}}$ and $r = 0, 1, 2, \dots 2^{l+1} - 1$ whereas the intervals should read $\Delta_r \equiv (\frac{r}{2^l}, \frac{r+1}{2^l})$ for $r = 0, 1, 2, \dots, 2^l$. With this the intergal $\int_0^1 r_j(x) r_k(x) r_l(x) r_m(x)\, dx$ where $j < k < l < m$ becomes:

$$\sum_r \int_{\Delta_r} r_j(x) r_k(x) r_l(x) r_m(x)\, dx = \sum_r \delta_r \int_{\Delta_r} r_m(x)\, dx \tag{15}$$

because $r_j(x), r_k(x)$ and $r_l(x)$ are constant on each interval Δ_r and thus

$\int_{\Delta_r} r_m(x)\, dx = 0$ and, because $m > l$, the result follows.

Kaczmarz and Steinhaus produce a similar proof in ([7], pages 125-126).

Interestingly, Kac demonstrates [5, p.9], *without* recourse to combinatorial techniques, that the probability that in n independent tosses of a fair coin, exactly l heads will appear is ($\mu[.]$ represents the measure of the set under consideration):

$$\mu[r_(t) + \cdots + r_n(t) = 2l - n] = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n x\, dx \tag{16}$$

but leaves it a straightforward integration exercise (see the Appendix) to show that:

$$\mu[r_1(t) + \cdots + r_n(t) = 2l - n] = \frac{1}{2^n} \binom{n}{l} \tag{17}$$

Of course, the RHS of (17) is immediate using a combinatorial approach but Kac wanted to show how Rademacher functions underpinned the integral on the RHS of (16). Billingsley notes in his probability textbook ([1], page 7) that the orthogonality relationships "stand on their own, in no way depend on any probabilistic interpretation". Gerald S. Goodman in his paper ([4], page 114) notes that because the binary coefficients $\epsilon_k(t)$ satisfy $\int_0^1 \epsilon_k(t)\, dt = \frac{1}{2}$ for all $k = 1, 2, \dots$ it follows that $\int_0^1 r_k(t)\, dt = 0$ for all $k = 1, 2, \dots$. He then notes that the "statistical independence of the $\epsilon_k(t)$ implies that the $r_k(t)$ are also independent , and , since they have mean zero, they satisfy the Multiplicativity Formula" (ie (7)).Thus he is taking the statistical independence as a previously established fact. However, as already noted above, there is a combinatorial angle one can employ to get the independence result. Indeed, with a bit of a notational shift one can get a result which is so clearly combinatorial that it screams the very statistical independence that orginally caught the eye of Borel and which formed the basis of Kac's book on statistical indepedence.

3 The combinatorial approach

To prove (7) using the combinatorial approach it is easiest to work with the $\epsilon_k(t)$ since they are 0 or 1. Invoking the dot product notation we have the following:

$$\epsilon_k(t) = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} \mathcal{X}_{[0,\frac{1}{2^k})}(t) \\ \mathcal{X}_{[\frac{1}{2^k},\frac{2}{2^k})}(t) \\ \vdots \\ \mathcal{X}_{[\frac{2^k-2}{2^k},\frac{2^k-1}{2^k})}(t) \\ \mathcal{X}_{[\frac{2^k-1}{2^k},\frac{2^k}{2^k})}(t) \end{pmatrix} \tag{18}$$

where $\mathcal{X}_A(t) = 1$ if $t \in A$ and 0 if $t \notin A$. In effect the $\epsilon_k(t)$ are sums of indicator functions over the subintervals with coefficients 0 and 1 as appropriate.

Using this notation we can compactly express the $\epsilon_k(t)$ as follows:

$$\epsilon_k(t) = \boldsymbol{\epsilon}_k \bullet \boldsymbol{\mathcal{X}}_k(t) \tag{19}$$

Where $\boldsymbol{\epsilon}_k = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ and $\boldsymbol{\mathcal{X}}_k(t)$ each have 2^k rows. Thus equation (2) can be written as:

$$r_k(t) = 1 - 2\,\boldsymbol{\epsilon}_k \bullet \boldsymbol{\mathcal{X}}_k(t) = 1 - \mu_k(t) \tag{20}$$

where $\mu_k(t) = 2\,\boldsymbol{\epsilon}_k \bullet \boldsymbol{\mathcal{X}}_k(t)$
Let:

$$I_n = \int_0^1 r_{i_1}(t)\, r_{i_2}(t) \dots r_{i_n}(t)\, dt = \int_0^1 \Pi_{k=1}^n (1 - \mu_{i_k}(t)\,)\, dt \tag{21}$$

Clearly because we are working with step functions when we work out the integral of a term in (21) which involves the integral of one of more terms $\mu_{k_i}(t)$ all we are doing is working out the area of some rectangles of height 1 and a certain base. Multiples of step functions are just step functions.The value of the integral is not affected by the jumps at the interval ends. Thus for instance:

$$\int_0^1 \mu_{k_i}(t)\, dt = 2\int_0^1 \boldsymbol{\epsilon}_{k_i} \bullet \boldsymbol{\mathcal{X}}_{k_i}(t)\, dt = 2 \times \underbrace{\frac{2^{k_i}}{2}}_{\text{number of rectangles of height 1}} \times \underbrace{\frac{1}{2^{k_i}}}_{\text{width of each rectangle}} = 1 \tag{22}$$

Clearly (22) holds generally.

From (21) we have:

$$I_n = \int_0^1 \left\{ 1 - \sum_{i=1}^n u_{k_i}(t) + \sum_{i<j}^n u_{k_i}(t) u_{k_j}(t) + \dots + (-1)^n u_{k_1}(t) u_{k_2}(t) \dots u_{k_n}(t) \right\} dt \tag{23}$$

This expression essentially encodes the combinatorial character of I_n.

In fact the following classical binomial expansion holds:

$$\boxed{I_n = \int_0^1 r_{i_1}(t)\, r_{i_2}(t) \ldots r_{i_n}(t)\, dt = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \cdots + (-1)^n \binom{n}{n} = (1-1)^n = 0} \tag{24}$$

To prove (24) one can proceed inductively. Note that:

$\int_0^1 \sum_{i=1}^n u_{k_i}(t)\, dt = \sum_{i=1}^n \int_0^1 u_{k_i}(t)\, dt = \sum_{i=1}^n 1 = \binom{n}{1}$ and $\int_0^1 dt = 1 = \binom{n}{0}$. Of course the integration and summations can be interchanged because we are dealing with step functions. Now consider the more interesting case:

$$\sum_{i<j}^n \int_0^1 u_{k_i}(t) u_{k_j}(t)\, dt \tag{25}$$

There are 2^{k_i-1} null (ie height zero) subintervals of $u_{k_i}(t)$ and to each of these there correspond $2^{k_j-k_i}$ subintervals of $u_{k_j}(t)$. Therefore there are $2^{k_i-1} \times 2^{k_j-k_i} = 2^{k_j-1}$ subintervals of $u_{k_j}(t)$ which are "annihilated". There are 2^{k_i-1} "+1" subintervals of $u_{k_i}(t)$ and to each of these there correspond $2^{k_j-k_i}$ subintervals of $u_{k_j}(t)$ and of these subintervals there are $2^{k_j-k_i-1}$ which are null subintervals. So in the product $2^{k_i-1} \times 2^{k_j-k_i-1} = 2^{k_j-2}$ subintervals of $u_{k_j}(t)$ are annihilated. Therefore the number of $u_{k_j}(t)$ "+1" subintervals which remain is:

$$2^{k_j} - (2^{k_j-1} + 2^{k_j-2}) = 2^{k_j-2} \tag{26}$$

Hence (bearing in mind the factor 2 in the definition of $u_k(t)$ in (20)):

$$\sum_{i<j}^n \int_0^1 u_{k_i}(t) u_{k_j}(t)\, dt = \sum_{i<j}^n 2^2 \times 2^{k_j-2} \times \frac{1}{2^{k_j}} = \sum_{i<j}^n 1 = \binom{n}{2} \tag{27}$$

More generally,

$$\int_0^1 u_{k_{i_1}}(t) u_{k_{i_2}}(t) \ldots u_{k_{i_n}}(t)\, dt = 2^n \frac{2^{i_n-n}}{2^{i_n}} = 1 \tag{28}$$

The factor 2^{i_n-n} is the number of +1 subintervals that remain in $u_{k_{i_n}}(t)$. This can be established inductively as follows. We need to show that:

$$\int_0^1 u_{k_{i_1}}(t) u_{k_{i_2}}(t) \ldots u_{k_{i_n}}(t) u_{k_{i_{n+1}}}(t)\, dt = 2^{n+1} \frac{2^{i_{n+1}-(n+1)}}{2^{i_{n+1}}} = 1 \tag{29}$$

We suppose that for all $j \leq n$ in the indices $\{k_{i_j}\}$ the number of "+1" terms in $v_n(t) = u_{k_{i_1}}(t) u_{k_{i_2}}(t) \ldots u_{k_{i_n}}$ which contribute to the product $u_{k_{i_1}}(t) u_{k_{i_2}}(t) \ldots u_{k_{i_n}}(t) u_{k_{i_{n+1}}}(t)$ is 2^{i_n-n}. By construction these are subintervals of $u_{i_n}(t)$. These 2^{i_n-n} "+1" subintervals correspond to $2^{i_{n+1}-i_n}$ subintervals of $u_{i_{n+1}}(t)$ half of which are null subintervals ie $2^{i_n-n} \times 2^{i_{n+1}-i_n-1} = 2^{i_{n+1}-n-1}$ are null.

The remaining $2^{i_n} - 2^{i_n - n} = 2^{i_n - n}(2^n - 1)$ null subintervals of $u_{i_n}(t)$ annihilate:

$(2^n - 1)2^{i_n - n} \times 2^{i_{n+1} - i_n} = (2^n - 1)2^{i_{n+1} - n}$ subintervals of $u_{i_{n+1}}(t)$.

Therefore the number of "+1" subintervals of $u_{i_{n+1}}(t)$ remaining in the product is:

$$\begin{aligned} 2^{i_{n+1}} - [(2^n - 1)2^{i_{n+1} - n} + 2^{i_{n+1} - n - 1}] =& 2^{i_{n+1}} - [2^{i_{n+1}} - 2^{i_{n+1} - n} + 2^{i_{n+1} - n - 1}] \\ =& 2^{i_{n+1} - n} - 2^{i_{n+1} - n - 1} \\ =& 2^{i_{n+1} - n - 1}[2 - 1] \\ =& 2^{i_{n+1} - (n+1)} \end{aligned} \tag{30}$$

Thus (29) holds generally. When we interchange integration and summation in (23) and apply the rule in (28) we get:

$$\begin{aligned} I_n =& \int_0^1 dt - \sum_{i=1}^{n} \int_0^1 u_{k_i}(t)\, dt + \int_0^1 \sum_{i<j}^{n} u_{k_i}(t) u_{k_j}(t)\, dt + \cdots + (-1)^n \int_0^1 u_{k_1}(t) u_{k_2}(t) \ldots u_{k_n}(t)\, dt \\ =& \binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \cdots + (-1)^n \binom{n}{n} \\ =& (1 - 1)^n \\ =& 0 \end{aligned} \tag{31}$$

4 Appendix

4.1 Proof of(16)

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n x\, dx =& \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \left(\frac{e^{ix} + e^{-ix}}{2} \right)^n dx \\ =& \frac{1}{2^n} \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \sum_{k=0}^{n} \binom{n}{k} e^{ikx}\, e^{-i(n-k)x}\, dx \\ =& \frac{1}{2^n} \frac{1}{2\pi} \int_0^{2\pi} \sum_{k=0}^{n} \binom{n}{k} e^{2i(k-l)x}\, dx \\ =& \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} \frac{1}{2\pi} \int_0^{2\pi} e^{2i(k-l)x}\, dx \\ =& \frac{1}{2^n} \binom{n}{l} \end{aligned} \tag{32}$$

since $\frac{1}{2\pi} \int_0^{2\pi} e^{2i(k-l)x}\, dx = 1$ if $k = l$ and 0 if $k \neq l$.

Extract from [7] pages 125-126.

§ 5. **Das Rademacher'sche System.**

Dieses (bereits in [225] angegebene) System ist unvollständig (vgl. [236]). Es gehört zur wichtigen Klasse der sog. lakunären Systeme, welche in Kapitel VII besprochen werden. Es besitzt eine interessante Eigenschaft:

Sind $j \leqslant k \leqslant l \leqslant m \leqslant \ldots \leqslant p \leqslant q$ *natürliche Zahlen und* $\{r_n(t)\}$ *die Rademacher'schen Funktionen, so ist* [451]

$$\int_0^1 r_j(t)\, r_k(t) \ldots r_p(t)\, r_q(t)\, dt = 0,$$

mit Ausnahme des Falles, wo der Integrand aus einer geraden Anzahl von Faktoren mit

$$j = k, \quad l = m, \quad \ldots, \quad p = q$$

besteht; dann ist das Integral gleich Eins.

Der Ausnahmsfall ist leicht zu erledigen, denn es sind dann $r_j^2(t), r_l^2(t), \ldots, r_p^2(t)$ sämtlich f. ü. gleich 1. Sonst kann man alle Paare mit gleichen Indizes streichen, denn sie liefern den Faktor 1 f. ü., und es bleiben unter dem Integralzeichen Funktionen mit verschiedenen Indizes ($j' < k' < \ldots < p' < q'$) übrig.

Das Produkt $r_{j'}(t) . r_{k'}(t) \ldots r_{p'}(t)$ hat $2^{p'}$ Konstanzintervalle (von der Länge $1/2^{p'}$), in welche $<0, 1>$ zerfällt. Ein jedes dieser Konstanzintervalle J zerfällt in $2^{q'-p'}$ gleiche Teile, innerhalb deren $r_{q'}(t)$ abwechselnd $+1$, -1 ist. Infolgedessen zerfällt

$$\int_0^1 r_{j'}(t)\, r_{k'}(t) \ldots r_{p'}(t)\, r_{q'}(t)\, dt$$

in $2^{p'}$ Summanden: const. $\int_J r_{q'}(t)\, dt$, und jeder Summand ist Null wegen $\int_J r_{q'}(t)\, dt = 0$.

Extract from [6], pages 42-43

on les intègre sur un intervalle de constance de $\varphi_{n_0+k}(x)$; or, nos intervalles partiels peuvent être décomposés en des tels intervalles. Il s'ensuit

$$\int_{Z_k}[s_{n_0+q}(x)-s_{n_0}(x)]^2\,dx \geqslant \int_{Z_k}[s_{n_0+k}(x)-s_{n_0}(x)]^2\,dx \geqslant \varepsilon^2\,|Z_k|,$$

$$\sum_{j=1}^{q} c_{n_0+j}^2 = \int_0^1 [s_{n_0+q}(x)-s_{n_0}(x)]^2\,dx \geqslant \sum_{k=1}^{q}\int_{Z_k}[s_{n_0+q}-s_{n_0}]^2\,dx$$

$$\geqslant \varepsilon^2 \sum_{k=1}^{q}|Z_k| \geqslant \frac{1}{2}\varepsilon^2\,|D^*| > \frac{1}{2}\varepsilon^2\,|D_\varepsilon| > 0;$$

ε et D_ε étant indépendants de n_0, la série (6) est divergente, c. q. f. d.

b) Si (1) est convergente dans un ensemble de mesure positive, alors cette série est uniformement convergente dans un ensemble E de mesure positive et l'on a

$$|s_n(x)| < M,$$

pour tout $x \in E$ et tout n, donc

$$|s_{n+p}(x)-s_n(x)| < 2M, \tag{8}$$

pour tout $x \in E$ et tous les n et p. Remarquons maintenant que l'intégrale

$$\int_0^1 \varphi_j(x)\,\varphi_k(x)\,\varphi_l(x)\,\varphi_m(x)\,dx \qquad (j \leqslant k \leqslant l \leqslant m) \tag{9}$$

n'est différente de zéro que dans le cas

$$j = k,\ l = m$$

(et il est évident qu'elle est égale à l'unité dans ce cas). En effet, pour $j \neq k$, $l = m$, l'intégrale devient

$$\int_0^1 \varphi_j(x)\,\varphi_k(x)\,dx = 0$$

cas $j < k < l < m$, on introduit les intervalles $\Delta_r = \langle \frac{r}{2^l}, \frac{r+1}{2^{l+1}} \rangle$ $(r = 0, 1 \ldots 2^{l+1}-1)$ et on écrit l'intégrale (9) comme

$$\sum_r \int_{\Delta_r} \varphi_j(x)\,\varphi_k(x)\,\varphi_l(x)\,\varphi_m(x)\,dx = \sum_r \delta_r \int_{\Delta_r}\varphi_m(x)\,dx,$$

car φ_j, φ_k et φ_l sont constantes dans chaque Δ_r; comme $\int_{\Delta_r}\varphi_m(x)\,dx = 0$, à cause de $m > l$, la remarque est justifiée.

(8) donne

$$\int_E [s_{n+p}(x)-s_n(x)]^2\,dx \leqslant 4M^2\,|E|,$$

$$4M^2|E| \geqslant \int_E \left(\sum_{n+1}^{n+p} c_i\varphi_i(x)\right)^2 dx = |E|\sum_{n+1}^{n+p} c_i^2 + 2\sum_{n+1 \leq i < k \leq n+p} c_i c_k \int_E \varphi_i(x)\,\varphi_k(x)\,dx. \tag{10}$$

Le système de tous les produits

$$\varphi_i(x)\,\varphi_k(x) \qquad (i < k)$$

est, d'après notre remarque, un système orthogonal et il est évident qu'il est normé; il s'ensuit que, pour $n > N(E)$,

$$\sum_{n+1 \leq i < k \leq n+p}\left[\int_E \varphi_i(x)\,\varphi_k(x)\,dx\right]^2 < |E|^2/9,$$

ce qui implique, pour la somme double $\sum$ de (10),

$$\left|\sum\right| \leqslant \frac{|E|}{3}\cdot\sqrt{\sum c_i^2 c_k^2} \leqslant \frac{|E|}{3}\cdot\sum_{n+1}^{n+p} c_i^2$$

et, à cause de (10),

$$4M^2|E| \geqslant \frac{|E|}{3}\sum_{n+1}^{n+p} c_i^2,$$

$$\sum_{n+1}^{n+p} c_i^2 \leqslant 12M^2,$$

pour $n > N(E)$ et pour tous les p; ceci est équivalent à la con-

5 References

[1] Patrick Billingsley, *"Probability and Measure"*, Third Edition, Wiley, 1995

[2] Émile Borel, (1909). Les probabilités dénombrables et leurs applications arithmétiques, Supplemento di rend. circ. Mat. Palermo 27, 247–271

[3] Gerald S. Goodman, *Statistical Independence and Normal Numbers: An Aftermath to Mark Kac's Carus Monograph*, The American Mathematical Monthly, Vol.106, No. 2 (Feb 1999) pp. 112-126 `http://www.jstor.org/stable/2589048`

[4] Masayoshi Hata, *Problems and Solutions in Real Analysis* ,World Scientific, 2007

[5] Mark Kac, *"Statistical Independence in Probability, Analysis and Number Theory"*, The Mathematical Association of America, 1964. This book can be accessed via this link: : `http://gibbs.if.usp.br/~marchett/estocastica/MarkKac-Statistical\-Independence.pdf`

[6] S Kaczmarz and H Steinhaus, *Le système orthogonal de M Rademacher*, Studia Mathematica 2.1 (1930): 231-247. This paper can be accessed via this link `http://eudml.org/doc/217255`

[7] S Kaczmarz and H Steinhaus, *Theorie der Orthogonalreihen*, Instytut Matematyczny Polskiej Akademi Nauk(Warszawa-Lwów), 1936. The book can be accessed via this link:`https://eudml.org/doc/219301`

[8] Hans Rademacher, *Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen*, Math. Ann. 87 (1922), no. 1-2, 112–138 (German)

[9] A Zygmund, *Trigonometric Series*, Third Edition, Vols 1 and 2 combined, Cambridge University Press, 2002)

6 History

Created 07/12/2016 11/12/2016 - corrected typographical error on page 2 formula (4)