A Slow Introduction to Slow Points

1 Some general background

I spent five happy academic years at Orsay, the first as a graduate student and the other four spread out over the next 20 years. I remember the splendid library with the books in alphabetical order (rather than by subject) so that one was constantly reminded of the broad spread of mathematics. I remember just sitting in beautiful grounds. (It was said that after the events of 1968, when the government wanted to have heavily muscled men around to keep order, the President of Orsay argued that this would look out place in such rural surroundings, but that gardeners were also heavily muscled and would do instead.) I remember the Orsay canteens and learning to drink strong black coffee and how to eat artichokes. (It was said that the first words of students from England when they arrived in the morning would be 'Quand est-ce que qu'on mange?') But most of all I remember the kindness shown to a novice by everyone whatever their academic standing.

I attended one lecture course by Kahane. It was on approximation of functions and beautifully put together, but not what this tyro expected from someone famous for his work on Hilbert's 13th problem. My mathematical education had prepared me for a course that took the fastest route to the most powerful result for the most general version of a particular problem. Instead Kahane would go back through the history of a topic and present other approaches which although less powerful from a general point of view shed a different light on the matter considered.

As an example which impressed me greatly, before giving the general version of the Stone–Weierstrass theorem, he gave versions of the Weierstrass, Bernstein and Lebesgue proofs of the polynomial approximation theorem pointing out the links between the Lebesgue proof and Stone's proof of the full theorem. The reader may recall that Lebesgue's proof that a continuous function $f : [0,1] \to \mathbb{R}$ can be uniformly approximated by polynomials consists of the observation that if

$$u_{m,r}(t) = m \max\{1 - m|r - mt|, 0\}$$

(that is to say, $u_{m,r}$ is the triangular function with base $[(r-1)/m, (r+1)/m]$ and height 1) then

$$\sum_{r=0}^{m} f(r/m)u_{m,r}(t) \to f(t)$$

uniformly on $[0,1]$. We now show (as is done in the standard proof of Stone–Weierstrass) that an individual triangle function $u_{r,m}$ can be uniformly approximated by polynomials on $[0,1]$ and this completes the proof.

In the sketch given above, we approximate f on $[0,1]$ by (the simplest) piecewise linear continuous function f_m with $f_m(r/m) = f(r/m)$ and allow $m \to \infty$, but, if we consider actual computations, it makes sense to take $m = 2^j$ only, so we can reuse values. There are various ways of writing f_{2^j} as the sum of triangle functions, but there is natural 'inductive' one

Exercise 1. *We work on* $[0,1]$. *Set* $v_0(t) = t$

$$v_{k,r}(t) = 2^k \max\{1 - 2^k|(r+1/2)2^{-k} - t|, 0\}$$

for $0 \leq r \leq 2^k - 1$. *If* $g : [0,1] \to \mathbb{R}$ *show that if we take*

$$a_{k,r} = g((r+1/2)2^{-k}) - \frac{g(r2^{-k}) + g((r+1)2^{-k})}{2}$$

then the function

$$g_m(t) = g(0) + \sum_{k=0}^{m} \sum_{r=0}^{2^k-1} a_{k,r} v_{k,r}(t)$$

is the simplest continuous piece-wise linear function with

$$g_m(2^{-m-1}r) = g(2^{-m-1}r)$$

for $0 \leq r \leq 2^{m+1}$. *(There is no need to this formally, but you should at least sketch a few graphs.)*

If g is continuous, explain why $g_m(t) \to g(t)$ uniformly as $m \to \infty$.

Seminar talks are different to lecture courses (as indeed are articles like this). The audience is less committed and different individual in that audience may have very different levels of knowledge. Kahane would often start of with a very simple result and then gradually build up to a much deeper finish. I will imitate him, in one point at least, by starting with a relatively simple idea, that of a Galton–Watson process.

Galton believed in a rather simple minded version of the 'survival of the fittest' and was bothered by the fact that although Shakespeare was the greatest English poet (and was therefore in Galton's view 'the fittest') he seemed to have no direct descendants. The same was true of several other 'great men'. In an effort to solve this 'paradox' he formulated it as a mathematical problem which was then attacked by Watson.

Here is Galton's problem. We start with an initial Adam and suppose that

$$\Pr(\text{Adam has } j \text{ sons}) = p_j$$

(where $0 \leq p_j$, $p_0 + p_1 + \ldots + p_m = 1$). Each son in turn has j sons with probability p_j (independently of each other) and the same goes for grandsons and so on. We ask 'what is the probability q that that the line of Adam will die out'.

Exercise 2. *What happens if $p_0 = 1$? What happens if $p_0 = 0$?*

From now on, we assume that $0 < p_0 < 1$. As the reader probably knows, the standard method of dealing with the problem is to look the probability q_n that Adam's line is extinct after n generations and to use the function

$$\phi(t) = \mathbb{E}t^X = \sum_{j=0}^{m} p_j t^j.$$

Exercise 3. *(i) By considering the probability that Adam's line goes extinct after at most $n+1$ generations in the particular case that Adam has exactly r sons show that*

$$q_{n+1} = \phi(q_n).$$

(ii) Observe that $1 \geq q_{n+1} \geq q_n$ and deduce that q_n converges to a limit q with $0 < q \leq 1$.

(iii) Show that ϕ is a convex function with $\phi'(1) = \mathbb{E}(X)$.

(iv) If $\mathbb{E}(X) > 1$, show that the equation $\phi(t) = t$ has exactly two roots α and 1 in $[0, 1]$. If $0 < t \leq \alpha$, show that $\phi(t) \leq \alpha$. Deduce that $q = \alpha$ and the probability of extinction is less than 1.

(v) If $\mathbb{E}(X) \leq 1$, show that the equation $\phi(t) = t$ has a unique root 1 in $[0, 1]$. Deduce that $q = 1$ and the probability of extinction is 1.

(vi) Remark that, if $\mathbb{E}(X) \leq 1$, then even if we start with a tribe of many Adams, extinction is certain (with probability 1) but, if $\mathbb{E}(X) > 1$, then, by starting with sufficiently many Adams, we can make the probability of extinction as small as we want.

[*If you cannot see what is happening draw a graph of ϕ for appropriate choices of p_j and look at the first few terms of the sequence (q_n).*]

If the matter appears obvious to the reader, she should note that Watson came to the mistaken conclusion that extinction was certain whatever the values of the p_j. She can also note the following result.

Exercise 4. *Let Y_n be the number of descendants in the n-th generation. If $\phi'(1) = 1$ show, by induction or otherwise, that $\mathbb{E}(Y_n) = 1$ and deduce that (although the probability of extinction is 1), the expected number of descendants is infinite!*

The Galton–Watson process is clearly relevant such things as the early stages of epidemics and queues (the sons of B are the members of the queue who arrive while B is being served).

2 A informal look at Brownian paths

We come now to the central topic of this essay:- Brownian paths. In 1827 the Scottish botanist Brown observes that very small particles of pollen suspended in water appeared to be in constant random motion. He then checked that the same was true for tiny particles of whatever description. One of the papers produced by Einstein in his'marvellous year' gave the correct reason that the random motion was due to constant jostling by the much tinier molecules of water. Einstein then predicted that the jostled particles would move in a 'random walk' of tiny random displacements and therefore *on average* move a distance $At^{1/2}$ in time t. The experimental verification of Einstein's prediction by Jean Perrin won Perrin a Nobel prize and provided strong evidence for the molecular structure of matter.

Let us try to imitate Brownian motion in 1-dimension. A simple model involves tossing a fair coin n times in one minute with rth toss a time r/n and a payment τ_n if the rth toss is heads and $-\tau_n$ if the rth toss is tails. If n is large our winnings after 1 minutes will be roughly normally distributed with mean 0 and variance $\tau_n n^{1/2}$. This strongly suggests that we take $\tau_n = n^{-1/2}$ so we shall do this. Now let $Y_k = n^{-1/2}$ if our kth throw is heads, $Y_k = -n^{-1/2}$ if tails and set $X_j = \sum_{k=1}^{j} Y_k$ so X_j is our total winnings at time j/n. If $0 \leq a < b < c < d \leq 1$ are fixed then provided n is large enough we will have $X_{[nb]} - X_{[na]}$ distributed more or less in the normal distribution $N(0, b-a)$ and the two random variables $X_{[nb]} - X_{[na]}$ and $X_{[nd]} - X_{[nc]}$ will be independent.

It is natural to make the following conjecture.

Conjecture 5. *There is a random variable X taking values in the space of functions from $[0,1]$ to $\mathbb{R}$ with the following properties.*

(1) $X(0) = 0$

(2) If $0 \leq s \leq t \leq 1$, then $X(t) - X(s)$ has the $N(0, t-s)$ distribution.

(3) If $0 = t_1 < t_2 < \ldots < t(n) = 1$ then the $X_{t(r+1)} - X_{t(r)}$ are independent.

If asked why we choose the normal distribution in (2) we mutter (reasonably enough) something about the central limit theorem.

In fact, Lévy proved the following theorem, which we give without proof.

Theorem 6. *Suppose that we have a random variable X taking values in the space of functions from $[0,1]$ to $\mathbb{R}$ with the following properties.*

(1) $X(0) = 0$

(2) If $0 \leq s \leq t \leq 1$, then $X(t) - X(s)$ has the same probability distribution as $X(t-s)$.

(3) If $0 = t_1 < t_2 < \ldots < t(n) = 1$, then the $X(t_{r+1}) - X(t_r)$ are independent.

(4) X is continuous with probability 1.

Then $X(t-s)$ has the normal distribution $N(0, \sigma^2(t-s))$ (where σ is a constant).

Lévy also showed that, if we drop the condition (4), the theorem is no longer true. (Since the X are then no longer continuous we need to adopt a convention about points of discontinuity in (3).)

The reader may worry that we have switched from considering '3-dimensional Brownian motion' to '1-dimensional Brownian motion'. However speaking in vague, but essentially correct, terms if $X(t)$, $Y(t)$, $Z(t)$ are 'independent representations of 1-dimensional Brownian random variables' then $(X(t), Y(t), Z(t))$ will represent a '3-dimensional Brownian random variables'. 1-dimensional Brownian random variables are linked directly to 'red noise' and indirectly to 'white noise' the hiss in electrical circuits.

In 1900 (so before Einstein's work) Bachelier introduced the idea of 1-dimensional Brownian motion as a model for stock market prices. The the price of a stock would be given by

$$Y(t) = a + bt + X(t)$$

where $X(t)$ was 1-dimensional Brownian random variables corresponding to many small decisions by individual traders. Since the model only depended on 3 variables a, b and σ^2 (with σ^2 the constant in Theorem 6) observation over time would enable a trader to estimate these constants and thus estimate the probability that the price of a stock would exceed a certain value after a particular time (say 1 month later). Bachelier's work was largely ignored by mathematicians and completely unread by economists during his lifetime. It is now considered as an essential tool for stock market trading. (I doubt that Kahane was particularly thrilled by this application.)

The paragraph above is the last reference I shall make to the 'real world' and from now on we stick to mathematics.

3 Construction of Brownian paths

The credit for turning Conjecture 5 into a theorem goes to Norbert Wiener using Fourier series with coefficients which are themselves random variables. The work involved in this procedure is not trivial and it is quite hard to use the resulting definition. Lévy introduced an alternative construction based on looking at the kind of piece-wise linear approximation we discussed in the first section.

Exercise 7. *Suppose, if possible, that conjecture 5 is correct. Let $X_m(t)$ be the simplest continuous piece-wise function such that $X_m(r2^{-m-1}) = X(r2^{-m-1})$.*

(i) Follow Exercise1 to show that

$$X_m(t) = \sum_{k=0}^{m} \sum_{r=0}^{2^k-1} Z_{k,r} \triangle_{k,r}(t)$$

where

$$\triangle_{k,r}(t) = \max\{0, 2^{-(k+1)/2}(1 - 2^{k+1}|t - (r + 1/2)2^{k+1}|)$$

(we retain this definition from now on) and

$$Z_{k,r} = 2^{(k+1)/2}\left(X((r + 1/2) - \frac{X((r + 1)2^{-k} + X(r2^{-k})}{2}\right)$$

(ii) Recall that a linear combination of independent normal random variables is normal and show that $Z_{k,r}$ is an $N(0, 1)$ random variable.

(iii) Recall that the two linear combinations U and V of independent normal random variables are independent if their covariance vanishes. Hence show that the $Z_{k,r}$ are independent for fixed k. Show that $Z_{k,r}$ is independent of $X((r+1)2^{-k} - X(r2^{-k})$.

We now 'reverse engineer' Exercise 7.

Lemma 8. *Suppose that the we have a collection of independent $N(0, 1)$ random variables $Z_{k,r}$ with $k, r \geq 0$. We define a sequence of random functions $X_m : [0, 1] \to \mathbb{R}$ inductively by the equations $X_0(t) = tZ_{0,0}$ and*

$$X_m(t) = \sum_{k=0}^{m}\sum_{r=0}^{2^k-1} Z_{k,r}\triangle_{k,r}(t)$$

Then the $X_m((r+1)2^{-m-1}) - X_m(r2^{-m-1})$ are independent $N(0, 2^{-m-1})$ random variables $[0 \leq r \leq 2^m - 1]$.

Proof. The conclusion is true for $m = 0$ by inspection.

Suppose the result is true for m. Write

$$V_I = X_{m+1}((r + 1)2^{-m-1}) - X_{m+1}(r2^{-m-1})$$

where $I = [r2^{-m}, (r + 1)2^m]$ and observe that

$$V_I = \begin{cases} 2^{-(m+2)/2}Z_{r+1,m} + \frac{1}{2}X_m((r + 1)2^{-m}) - \frac{1}{2}X_m(r2^{-m}) & \text{if } I = [r2^{-m}, (r + 1)2^{-m-1}] \\ -2^{-(m+2)/2}Z_{r+1,m} - \frac{1}{2}X_m((r + 1)2^{-m}) - \frac{1}{2}X_m(r2^{-m}) & \text{if } I = [(r + 1)2^{-m-1}, (r + 1)2^{-m}] \end{cases}$$

By the standard results on independent normal random variables referred to earlier, V_I is normal with mean 0 and variance 2^{-m-1}

The the V_I will be independent if $\operatorname{cov}(V_I, V_J) = 0$ for all $I \neq J$. By inspection V_I and V_J will be sums of disjoint collections of independent random variables and so independent with $\operatorname{cov}(V_I, V_J) = 0$ unless $I, J \subseteq [u2^{-m}, (u+1)2^{-m}]$ for some u. In this case we may assume without loss of generality that $J = [u2^{-m}, (u+1/2)2^{-m}]$, $I = [(u+1/2)2^{-m}, (u+1)2^{-m}]$ But now

$$\operatorname{cov}(V_{I,J}) = \mathbb{E}\left(\frac{1}{4}(X_m(2^{-m}(u+1) - X_m(2^{-m}(u)))^2 - (2^{-(m+2)/2}Z_{u,m})^2\right) = 0$$

and we are done. ∎

We still have to show that X_n converges in some sense to an appropriate X. The key to this and much more lies in two standard inequalities.

Lemma 9. *There exist constants $C_1 > C_2 > 0$ such that if Y is a $N(0,1)$ random variable and $\lambda > 1$.*

$$C_1 \frac{\exp(-\lambda^2/2}{\lambda} \geq \Pr(Y \geq \lambda) \geq C_2 \frac{\exp(-\lambda^2/2)}{\lambda}$$

Proof. Observe that

$$\Pr(Y \geq \lambda) \leq \frac{1}{\sqrt{2\pi}} \int_\lambda^\infty \frac{t}{\lambda} \exp(-t^2/2)\, dt = \frac{1}{\sqrt{2\pi}} \frac{\exp(-\lambda^2/2)}{\lambda}$$

and

$$\Pr(Y \geq \lambda) \geq \frac{1}{\sqrt{2\pi}} \int_\lambda^{\lambda+\lambda^{-1}} \exp(-t^2/2)\, dt \geq \frac{1}{\sqrt{2\pi}} \times \frac{\exp(-(\lambda+\lambda^{-1})^2/2)}{\lambda}.$$

∎

We shall use Lemma 9 over and over again. If the reader mutters 'but everybody knows these inequalities' let me confess that, although I may have known these inequalities, I most certainly did not understand their power. Then one day I gave a talk full of Chebychev inequalities. Afterwards Kahane quietly drew me aside and gave me a tutorial on sub-Gaussian random variables which showed me the power of related inequalities and has remained a source of inspiration to me through my mathematical life.

One of the main themes running through the subject may be summarised as follows. The only continuous homogeneous random paths are given by scaled normal increments. Because the normal distribution drops away very quickly this means that $X(t+h) - X(t)$ is unlikely to be very much larger than we expect.

Lemma 10. *(i) Let X_n be defined as in Lemma 8. If we write*

$$U_n = X_{n+1} - X_n,$$

then

$$U_n(t) = \sum_{r=0}^{2^n-1} Z_{n,r}\triangle_{n,r}(t).$$

(ii) There is a constant A such that

$$\Pr(\|U_n\|_\infty \geq A2^{-n/2}n^{-1/2} \leq 2^{-n}$$

(iii) The sequence X_n converges uniformly to a continuous function X

Proof. (i) See Exercise1.

(ii) We have

$$\begin{aligned}\Pr(\|U_n\|_\infty \leq A2^{-n/2}n^{1/2}) &= \Pr\{|Z_{n,r}| \leq An^{1/2})\text{for } 0 \leq r \leq 2^n - 1\}\\ &\leq 2^n \Pr\{|Z_{n,0}| \geq A_1 n^{1/2}\}\\ &\leq 2^{n+1}A_2\frac{\exp(-A_1^2 n/2}{A_1 n^{1/2}} \leq 2^{-n}\end{aligned}$$

for appropriate A_1 and A.

(iii) We have

$$\sum_{n=1}^{\infty} \Pr(\|U_n\|_\infty > A(2^{-n/2}n^{1/2}) < \infty)$$

so by the first Borel–Cantelli lemma there is probability 1 that

$$\|U_n\|_\infty < A2^{n/2}n^{1/2}$$

for sufficiently large n and so $\sum_{n=1}^{\infty} U_n$ converges uniformly. Since the U_n are continuous so is the sum. ■

Theorem 11. *The X obtained in Lemma 10 satisfy the conditions of Conjecture 5.*

Proof. The probability that X is not continuous is 0 so ignore this possibility. We observe that $X(r/2^n) = X_m(r/2^n)$ for all $m \geq n$. Thus writing D for the collection of intervals of the form $[r/2^n, r'/2^n]$ with $0 \leq r < r' \leq 2^[n, \geq 1$ we see that

(1) $X(0) = 0$

(2) If $[s, t] \in D$, then $X(t) - X(s)$ has the $N(0, t - s)$. distribution.

(3) If $0 \leq s_1 < t_1 \leq s_2 < t_2 < \ldots s_k < t_k \leq 1$ and the $s_j, t_j \in D$ then $X(t_j) - X(s_j)$ are independent.

We can now use monotone convergence and the fact that given $[a, b] \subseteq [0, 1]$ we can find $[a_n, b_n] \in D$ with $[a_n, b_n] \uparrow [a, b]$ to obtain the results required. ■

The estimates in Lemma 10 (ii) are much more powerful than we need to obtain (iii). However with a little work they show that there a constant C such that (with probability 1)

$$\sup_{t\in[0,1]} \limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{(|h|\log|h|^{-1})^{1/2}} \leq C.$$

With more work along similar lines we could obtain the following theorem of Lévy

Theorem 12. *If X is our standard Brownian random path then*

$$\sup_{t\in[0,1]} \limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{(2|h|\log|h|^{-1})^{1/2}} = 1.$$

Theorem 12 deals with the worst behaviour at any point. What about general behaviour?

Lemma 13. *With probability* 1 *we have*

$$\limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{(2|h|\log\log|h|^{-1})^{1/2}} \geq 1$$

for almost all (in Lebesgue measure) t.

Proof. Suppose $0 < a < 1$. Let $t \in [0,1]$ and $h \neq 0$ be such that $t+h \in [0,1]$. Observe that if Z is an $N(0,1)$ random variable and $n \geq 20$, then using Lemma 9

$$Pr\left(\frac{|X(t+2^{-n}h)-X(t+2^{-n-1})|}{(2|2^{-n-1}h|\log\log|2^{-n-1}h|^{-1})^{-1}} \geq a\right) = \Pr(|Z| \geq a(\log\log|2^{-n-1}h|^{-1})^{1/2}\}.$$

Using Lemma 9 we know there is a $C_3(h) > 0$ depending only on h such that

$$\Pr(|Z| \geq a(\log\log|2^{-n-1}h|^{-1})^{1/2} \geq C_3(h)n^{-a^2}.$$

Since $\sum_{n=1}^{\infty} n^{-a^2}$ diverges and the increments $|X(t+2^{-n}h)-X(t+2^{-n-1})|$ are independent we know by the second Borel–Cantelli lemma that

$$\frac{|X(t+2^{-n}h)-X(t+2^{-n-1})|}{(2|2^{-n-1}h|\log\log|2^{-n-1}h|^{-1})^{1/2}} \geq a$$

for infinitely many values of n with probability 1. Thus

$$\Pr\limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{(2|h|\log\log|h|^{-1})^{1/2}} \geq a = 1$$

with probability 1.

Using Fubini's theorem this gives

$$\limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{(2|h|\log\log|h|^{-1})^{1/2}} \geq a) = 1$$

except on a set of t of Lebesgue measure zero. The lemma follows on choosing an increasing sequence $a_n \to 1$. ■

The reader who knows the law of the iterated logarithm will recognise that we have essentially followed the easier half of that proof and will be able to extend the previous Lemma 13 to give the stronger result.

Theorem 14. *With probability* 1*, we have*

$$\limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{(2|h|\log\log|h|^{-1})^{1/2}} = 1$$

for almost all (in Lebesgue measure) t.

Lemma 13 shows that with probability 1, the Brownian curve is non-differentiable almost everywhere but does not show that it is nowhere differentiable.

However, Paley, Wiener and Zygmund proved the following result which does give nowhere differentiability.

Theorem 15. *If* $\epsilon > 0$ *then with probability* 1

$$\limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{|h|^{1/2+\epsilon}} = 1$$

for all t.

This was later improved by Dvoretsky to give the following result.

Theorem 16. *There exists a beta* > 0 *such that with probability* 1

$$\limsup_{h\to 0} \frac{|X(t+h)-X(t)|}{|h|^{1/2}} \geq \beta$$

for all t.

This completes our introduction. Readers now has all the background required to produce a proof that (with probability 1) a Brownian curve in nowhere differentiable. They may wish to try their hand at this before looking at the next section.

4 A short section

I believe that I was in the next office to Kahane when he came up with the following argument and so I one of the first to hear it explained. If my memory deceives me, please do not set me right.

Second proof of Theorem 16. Choose $\alpha > 0$ such that if Z is an $N(0,1)$ function

$$\Pr(|Z| \geq \alpha) \leq 1/2.$$

Let X be our standard Brownian curve $X : [0,1] \to \mathbb{R}$. and let N be a positive integer which we fix for the time being.

We start with the a population of 2^N Adams, specifically the dyadic intervals $[r2^{-N}, (r+1)2^{-N}]$ with $0 \leq r \leq 2^{-N}$ and proceed inductively as follows. Let $\mathcal{A}_m$ is the collection of male descendants $[r2^{-N-m}, (r+1)2^{-N-m}]$ in the mth generation. Every Adam $[u2^{-N-m}, (u+1)2^{-N-m}] \in \mathcal{A}_m$ has two sons $[2u2^{-N-m-1}, (2u+1)2^{-N-m-1}]$ and $[(2u+1)2^{-N-m-1}, (2u+2)2^{-N-m-1}]$ if and only if

$$|2X(2u+1)2^{-N-m-1} - X(u2^{-N-m}) - X((u+1)2^{-N-m})| \leq \alpha 2^{-(m+N)/2}$$

and otherwise has no sons. By the choice of α this means that the expected number of sons will be no greater than 1 and so extinction is certain (with probability 1).

Thus with probability 1 every $t \in [0,1]$ will belong to an interval $[u2^{-N-m}, (u+1)2^{-N-m}]$ with

$$|2X(2u+1)2^{-N-m-1} - X(u2^{-N-m}) - X((u+1)2^{-N-m})| \leq \alpha 2^{-m-N}.$$

and so with

$$\max\{|X(t) - X(2u+1)2^{-N-m-1}|, |X(t) - X(2u)2^{-N-m}|, |X(t) - X((u+1))2^{-N-m}| \geq \frac{\alpha}{4} 2^{-(m+N)/2}.$$

Since $t \in [u2^{-N-m}, (u+1)2^{-N-m}]$ we have

$$\frac{X(t+h) - X(t)}{|h^{1/2}|} \geq \frac{\alpha}{4}$$

for some h with $0 < |h| \leq 2^{-N-m}$. We have shown that with probability 1

$$\sup_{0<|h|<2-N} \frac{|X(t+h) - X(t)|}{|h|^{1/2}} \geq \frac{\alpha}{4}$$

for all $t \in [0,1]$.

If we now allow N to vary we see that, with probability 1

$$\limsup_{|h|\to 0} < \frac{|X(t+h) - X(t)|}{|h|^{1/2}|} \geq \frac{\alpha}{4}$$

for all $t \in [0,1]$. ∎

5 Slow points

The previous section shows Kahane's ability to come up with new and insightful methods approaches to old problems and I will be happy even if the reader goes no further. However major mathematicians not only exhibit ingenuity in finding new ideas but also the power to exploit those ideas to maximum effect.

In this case, Kahane was able to produce a beautiful complement to Theorem 16.

Theorem 17. *Consider Brownian paths X. There exists an* $\gamma > 0$ *such that with probability* 1 *there exists a* $v \in [0,1]$ *such that*

$$\limsup_{h\to 0} \frac{|X(v+h)-X(v)|}{|h|^{1/2}} \leq \gamma.$$

Points like v with

$$\limsup_{h\to 0} \frac{|X(v+h)-X(v)|}{|h|^{1/2}} < \infty$$

are called *slow points*.

It is natural to try to prove this by using the fact that the Galton–Watson process has positive non-extinction probability when the expected number of sons is strictly greater than 1. However, this naive approach soon runs into trouble. In Section 4 we relied on the fact that the information that

$$|2X((2u+1)2^{-k-1}) - X(u2^{-k}) - X((u+1)2^{-k})|$$

was large was sufficient information about $[u2^{-k},(u+1)2^{-k}]$ to tell us that that interval was 'suitable for our purposes'. However in our new situation inequalities are reversed the information that

$$|2X((2u+1)2^{-k-1}) - X(u2^{-k}) - X((u+1)2^{-k})|$$

is large not only tells us that $[(u2^{-k}),(u+1)2^{-k}]$ 'unsuitable for our purposes', but also that intervals $[r2^{-m}),(r+1)2^{-m}]$ close to $[(u2^{-k})-(u+1)2^{-k}]$ are 'unsuitable for our purposes'.

The way round this difficulty relies on the fact that very large values of a normal distribution are very rare and so very unsuitable intervals are very rare. In particular if Z has an $N(0,1)$ distribution

$$p_\chi = p_\chi(\lambda) = \Pr(|Z| \in [\lambda 2^\chi, \lambda 2^{\chi+1}]) \leq 2\Pr(|Z| \geq \lambda 2^\chi) \leq C_2 \frac{\exp(-\lambda^{-2}2^{2\chi}/2)}{\lambda 2^\chi}$$

for some constant C_2.

We observe that, if λ is small then p_0, may be large (so the arguments given below will not work), but if λ is reasonably large, not only will p_0 be very small, but the remaining terms p_j will converge to zero extremely rapidly.

Let X be a standard Brownian motion. We perform the following inductive construction. We take $\mathcal{A}_N$ to be the collection of all dyadic intervals $[2^{-N}u, 2^{-N}(u+1)]$. At the nth generation (with $n \geq N$) we have a collection $\mathcal{A}_n$ of descendant dyadic intervals $[2^{-n}u, 2^{-n}(u+1)]$. We look at *all* dyadic intervals $[2^{-n}v, 2^{-n}(v+1)]$ and declare that

$$[2^{-n}v, 2^{-n-1}(2v+1)]\,[2^{-n-1}(2v+1, 2^{-n}(v+1)] \in B_n(\chi)$$

if

$$Z_{v,n} = 2^{n/2}(X(2^{-n}v + \tfrac{1}{2}) - (X(2^{-n}(v+1)) + X(2^{-n}(v))$$

is such that $\lambda 2^{\chi} < |Z_{v,n}| \leq \lambda 2^{\chi+1}$. We write $\tilde{A}_{n+1}$ for the collection of $I = [2^{-(n+1)}u, 2^{-(n+1)}(u+1)$ with $I \subset J$ with $J \in \mathcal{A}_n$ and $\tilde{\mathcal{A}}_{n+1}(\chi)$ for the collection of $I \in \tilde{\mathcal{A}}_{n+1}$ at a distance between $2^{3\chi-2}2^{-n}$ and $2^{3\chi-1}2^{-n}$ from some interval of $B_n(\chi)$. We take

$$\mathcal{A}_{n+1} = \tilde{\mathcal{A}}_{n+1} \setminus \bigcup_{\chi=0}^{\infty} \tilde{A}_{n+1}(\chi)$$

(thus $\mathcal{E}_{n+1}$ consists of all possible sons with the exception of those too near an interval with bad behaviour).

We now do some elementary probability.

Exercise 18. *Suppose we toss a coin R times with probability q of heads. If S is the number of heads observed, use Chebychev's inequality to show that*

$$\Pr(S \geq R(1 + \theta(q(1-q))^{1/2}) \leq \frac{1}{\theta R}$$

and so there is an event Ω of probability at most $1/(Rq^2)$ such that if Ω does not occur

$$S < R(q + \theta(q(1-q))^{1/2}$$

Write $|A|$ for the number of elements in a set A. Then $|\tilde{B}_n(\chi)|$ is dominated by a binomial random variable $S_{n,\chi}$ corresponding to $R_{\chi,n} = 2^{3\chi}|\mathcal{A}_n|$ tosses of a coin with probability $p(\lambda 2^{\chi}) = \Pr(|Z| \geq \lambda 2^{\chi})$ of heads. By the previous exercise there is an event Ω_χ of probability at most $1/(|\mathcal{A}_n| p_\chi(\lambda))$ such that if Ω_χ does not occur

$$\tilde{B}_n(\chi) < |\mathcal{A}_n| 2^{3\chi}(p_\chi(\lambda) + \chi(p_\chi(\lambda)(1 - p_\chi(\lambda))^{1/2}).$$

Since the $p_\chi(\lambda)$ decrease extremely fast for large λ we can choose λ^* sufficiently large (and independent of $\mathcal{A}_n$) such that

$$\sum_{\chi=0}^{\infty} p_\chi(\lambda^*)) < 1$$

and

$$\sum_{\chi=0}^{\infty} 2^{3\chi}(p_\chi(\lambda^*) + \chi((p_\chi(\lambda^*)(1 - p_\chi(\lambda^*))^{1/2} < 1/8$$

With this choice of λ^*, there is an event Ω_n of probability at most $1/|\mathcal{A}_n|$ such that, if Ω_n does not occur,

$$\sum_{\chi=1}^{\infty} |\tilde{B}_n(\chi)| < |E_n|/8$$

and so

$$|\mathcal{A}_{n+1}| \geq 2^{3/2}|\mathcal{A}_n|$$

Exercise 19. *(i) Show by induction that, with this choice of* λ^**,*

$$\mathcal{A}_n \geq 2^{3n/2}$$

for all $n \geq N$ *with probability at least* $1 - \sum_{r=N}^{n} 2^{-3r/2}$.

(ii) Deduce that

$$\bigcap_{n=N}^{\infty} \bigcup_{I \in \mathcal{A}_n} I \neq \emptyset$$

with probability at least $1 - \sum_{r=N}^{\infty} 2^{-3r/2}$.

(iii) Conclude that, with probability 1*, a Brownian path X will have the property that there exists an N (depending on X) such that, if* $\mathcal{A}_N$ *is the collection of all dyadic intervals* $[u2^{-N}, (u+1)2^{-n}]$ *and we perform the process described above,* $\bigcap_{n=N}^{\infty} \bigcup_{I \in \mathcal{A}_n} I \neq \emptyset$*.*

Lemma 20. *Suppose that X and N have the properties described in Exercise19. Then there exists an* $M \geq N$ *(again dependent on X) and a constant C depending on* λ^* *but not on X, N or M such that, if* $t \in \bigcap_{n=N}^{\infty} \bigcup_{I \in \mathcal{A}_n} I$ *we have*

$$|X(t) - X(s)| \leq C|t-s|^{1/2} \textit{ for all } |t-s| \leq 2^{-M}$$

.

Proof. As usual, we will write

$$X(w) = \sum_{k=0}^{\infty} U_k(w)$$

with

$$U_k(w) = \sum_{r=0}^{2^k-1} Z_{k,r} \triangle_{k,r}(w),$$

but the reader should note that X is now a fixed function and therefore the $Z_{k,r}$ are fixed real numbers.

Observe first that $X_N(w) = \sum_{k=0}^N U_k(w)$ is a continuous piece-wise linear function and so we can find an $M \geq N$ such that

$$|X_N(t) - X_N(s)| \leq |t-s|^{1/2} \text{ for all } |t-s| \leq 2^{-M}.$$

The stated result will thus follow if we can show that there exists C' depending on λ^* but not on X, N or M such that, if $t \in \bigcap_{n=N}^{\infty} E_n$ we have

$$\left| \sum_{k=N+1}^{\infty} U_k(s) - U_k(t) \right| \leq C'|s-t|$$

whenever $|s-t| \leq 2^{-M}$.

Suppose therefore that $2^{-m} \leq |t-s| \leq 2^{-m+1}$ with $m \geq M+1$. If $k \leq m$ and $t \in I = [u2^{-k}(u+1), 2^{-k}]$, $s \in J = [v2^{-k}, (v+1)2^{-k}]$ then either $J = I$ or J and I are neighbouring intervals. In either case, $J \notin \bigcup_{\chi=1}^{\infty} |B_m(\chi)|$ so

$$|U_k(w)(t) - U_k(s)(t)| \leq \lambda^* 2^{k/2}|t-s|$$

and

$$\begin{aligned}\sum_{k=N+1}^{m} |U_k(t) - U_k(s)| &\leq \sum_{k=N+1}^{m} \lambda^* 2^{k/2}|t-s| \\ &\leq \lambda^* \sum_{k=N+1}^{m} 2^{k/2} 2^{-(m-1)/2}|t-s|^{1/2} \leq A_1\lambda^*|t-s|^{1/2}\end{aligned}$$

where $A_1, A_2, \ldots$ denote constants.

If $k > m$ we know that if $s \in |B_m(\chi)|$ then $|s-t| \geq 2^{3\chi}2^{-k}$ so, since $|t-s| \leq 2^{-m+1}$, we have

$$\chi \leq (k+1-m)/3.$$

Thus

$$|U_k(s)| \leq \lambda^* 2^{-(k+1-m)/3} 2^{-k/2} \leq A_2\lambda^* 2^{-(k+1-m)/3}|s-t|^{1/2}.$$

We also know that

$$|U_k(t)| \leq \lambda^* 2^{-k/2-1} \leq A_3\lambda^* 2^{-(k-m)/2}|s-t|^{1/2}$$

so

$$\sum_{k=m+1}^{\infty} |U_k(t) - U_k(s)| \leq \sum_{k=m+1}^{\infty} (|U_k(t)| + |U_k(s)|) \leq A_4\lambda^*|s-t|^{1/2}$$

and

$$\sum_{k=N+1}^{\infty} |U_k(t) - U_k(s)| \leq A_5 \lambda^* |s-t|^{1/2}$$

giving the desired result. ■

Putting Exercise 19 (iii) and Lemma 20 together gives Theorem 17. Just as there are 'slow points' which show that Theorem 16 is essentially best possible, so Orey and Taylor had shown that there are 'fast points' which show that Theorem 12 is essentially best possible.

Theorem 21. *Consider the collection of Brownian paths X. There exists an $\gamma > 0$ such that with probability 1 there exists a $v \in [0,1]$ such that*

$$\limsup_{h \to 0} \frac{|X(v+h) - X(v)|}{|h|^{1/2} \log(1/|h|)} \geq \gamma.$$

We outline a proof in the next exercise.

Exercise 22. *(i) Reread Lemma 10.*

(ii) Prove that there is a constant $K > 0$ such that

$$\Pr(\|U_n\|_\infty \leq K2^{-n/2}n) \leq 2^{-n}.$$

(iii) Deduce that given N, then, with probability 1, there is an $n \geq N$ and dyadic interval $J = [r2^{-n}, (r+1)2^{-n}]$ with

$$|U_n(r+1/2)2^{-n}| \geq K2^{-n/2}n$$

(iv) By using (iii) or otherwise, show that given N and a dyadic interval $I = [u2^{-M}, (u+1)2^{-M}]$, then, with probability 1, there is an $n \geq N, M$ and a dyadic interval $J = [r2^{-n}, (r+1)2^{-n}] \subseteq I$ with

$$|U_n((r+1/2)2^{-n})| \geq K2^{-n/2}n.$$

(v) Show that, with probability 1, we can find $n(q) \to \infty$ and nested dyadic intervals $J(q) = [r(q)2^{-n(q)}, (r(q)+1)2^{-n(q)}] \subseteq I$ with

$$|U_{n(q)}((r(q)+1/2)2^{-n(q)}| \geq K2^{-n(q)/2}n(q).$$

Take $\{v\} = \bigcap_{q=1}^{\infty} J(q)$ and show that there exist $\gamma > 0$ such that

$$\sup_{s \in J(q)} |X(s) - X(v)| \geq \gamma |s-v|^{1/2} \log|s-v|^{-1}$$

for each q

There is now an extensive literature on fast and slow points but the present author has read none of it and can only recommend Kahane's beautiful book *Some Random Series of Functions* issued by CUP as a superb introduction to application of probability theory to analysis.

6 Final remarks

Just as there are non-mathematicians who are proud of their inability to understand mathematics, so there are many mathematicians who are proud of their administrative incompetence. However mathematics at Orsay and beyond benefited not only from the glow shed by Kahane's mathematics, but from the quiet efficiency with which he and his splendid secretary Madame Dumas ran things.

To take one example, many mathematical conferences run from dawn to dusk. Some have 'invited speakers' in the morning and what must therefore be described as 'uninvited speakers' in the afternoons. When Kahane organised things the mornings were devoted to mathematics and the afternoons to walks through fine scenery.

Kahane loved and excelled in all aspects of mathematics:- teaching, researching and learning from others. I am fortunate to have known him.