Random Walks and Brownian Motion Instructor: Ron Peled
Tel Aviv University Spring 2011

Lecture 9

Lecture date: May 02, 2011 *Scribe: Jonathan Hermon*

In today's lecture we present the Brownian motion (BM). We start with an intuitive discussion, describing the BM as a limit of SRW. We present some of the BM properties and try to explain why we can expect these properties from a limit of SRW. We then give a formal definition for BM and prove such a process exists (Wiener 1923). Before the proof we give some important facts about the normal distribution and Gaussians. We end the lecture with some theorems about the continuity of BM.

1 Brownain motion as a limit of SRW

Let S_n be a SRW on $\mathbb{Z}$. By the CLT we have $\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$.
Now we fix n and think of a "SRW" on [0,1]. We take the k^{th} step at time $\frac{k}{n}$ and our increments are $\pm\frac{1}{\sqrt{n}}$ with probability $\frac{1}{2}$ each. At time 1 our RW will have value $\frac{S_n}{\sqrt{n}}$ so it is roughly distributed like $N(0,1)$.
Now think of this RW as a random function that is defined at $\frac{k}{n}$ as $\frac{S_k}{\sqrt{n}}$ and in between two such points it is linear. Denote this random function by f_n. It is easy to see that for any fixed point x, $f_n(x) \xrightarrow{d} N(0,x)$.
In fact, $f_n \xrightarrow{d} w$ for some continuous function w, where the converges is in distribution over the space of continuous functions $C[0,1]$. This w is a random function that satisfies $w(t) \sim N(0,t)$ and is a BM.

Remark 1 *As we will see in a few lectures, this is a special case of Donsker invariance principle. The principle describe the universality of BM and states that BM can be achieved as a limit of every RW with mean 0 that has a finite variance. Furthermore, we can do the same for RW with mean zero that does not have finite variance, and the limit is called a Lévy process, and will be discontinuous at all points, with probability 1.*

Before defining the BM formally, we ask ourselves what properties do we expect the BM (denoted as $\{B(t)\}_{t\geq 0}$) will have as a limit of SRW.

1. Independent increments: $\forall\ 0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, $B(t_n) - B(t_{n-1}), ..., B(t_2) - B(t_1)$

are independent.

2. Stationary increments: fix $h > 0$, then for any $t \geq 0$, $B(t+h) - B(t)$ has the same distribution, independent of $t \geq 0$. By the previous discussion and the CLT we expect this distribution to be $N(0, h)$.

3. The random function $B(t)$ is a.s a continuous function.

2 Conformal Invariance

We can define Brownian motion in $\mathbb{R}^d$ as a limit of SRW in $\mathbb{R}^d$. It seems reasonable that even if we rotate or stretch the SRW, the limit of the process (again the convergence is in distribution over the space of continuous functions) will be a BM. For this reason we can expect the BM have some invariance properties under rotation and scaling. It turns out that this is indeed the case. Similarly, we have seen in the first lecture that if we reverse the order of the increments of a SRW, we still get a SRW. Thus we expect the BM to have a time reversal property. Furthermore, the BM has a much stronger property, it is invariant under conformal maps. More precisely, if f is a conformal map, then up to time changes $f(B)$ has the same distribution as $B(f)$.
Most of these statements would be stated in a more precise manner in the next lecture and would be proven. For more details, see chapter 1 section 1.13 and chapter 7 section 2 of (1).

3 Formal Definition

Definition 2 *A real-valued stochastic process $\{B(t)\}_{t \geq 0}$ is a (linear, i.e 1 dimensional) Brownain motion starting at point x if*

1. $B(0) = x$
2. *Independent increments:* $\forall\ 0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, $B(t_n) - B(t_{n-1}), ..., B(t_2) - B(t_1)$ *are independent.*
3. $\forall\ t > s$, $B(t) - B(s) \sim N(0, t-s)$.
4. *The random function $B(t)$ is a.s a continuous function.*

Remark 3 *Note that for any fixed* $w \in \Omega$, $B(w,t)$ *is a real-valued function. for a set of probability one we get a continuous function. Similarly, for any fixed t, B(t) is a N(0,t) random variable.*

Definition 4 *For a stochastic process B, the finite dimensional distributions are all the joint distribution of the form* $(B(t_1), B(t_2), ..., B(t_n))$ *for some* $n \in \mathbb{N}$ *and* $t_1, ..., t_n \in \mathbb{R}$.

It is easy to see that they generate (span the σ-field of) all events which depends only on countable many co-ordinates.
Notice that properties 1,2,3 in the definition of BM characterize the finite dimensional distribution of the BM. The following example demonstrate that property 4 is not determined by the finite dimensional distributions. However, given that property 4 does hold, the finite dimensional distributions determine not only the events that depends only on countable many co-ordinates, but by continuity they also determine some events depending on continuum of co-ordinates.

Example 5 *Suppose that* $\{B(t)\}_{t\geq 0}$ *is a Brownian motion and U is an independent random variable, which is uniformly distributed on [0, 1]. Then the process* $\{\tilde{B}(t)\}_{t\geq 0}$ *defined by*

$$\tilde{B}(t) = \begin{cases} B(t) & if\, t \neq U \\ B(t)+1 & if\, t = U \end{cases}$$

B(t) has the same finite-dimensional distributions as a Brownian motion, but is discontinuous at U if B is continuous at U. Hence this process is not a Brownian motion.

Theorem 6 ***(Wiener 1923)*** *Standard Brownian motion exists.*

It is a substantial issue whether the conditions imposed on the finite-dimensional distributions in the definition of Brownian motion allow the process to have continuous sample paths, or whether there is a contradiction (thus there is no point studying BM). Given that the name of our course is RW and Brownian motion, it is not so surprising that indeed Brownian motion exists. We follow Paul Lévy's construction of Brownian motion, constructing it as an **a.s** uniform limit of continuous functions, to ensure that it automatically **a.s** has continuous paths. Note that we need only to construct a standard Brownian motion $B(t)$, as $X(t) = x + B(t)$ is a Brownian motion with starting point x.
The proof exploits properties of Gaussian random vectors, which are the higher dimensional analogue of the normal distribution. Therefore, before starting the construction we will state some facts about Gaussians that we will need. Some proofs can be found in appendix 12 of (1)

4 Facts about Gaussians

Definition 7 *A random vector $X = (X_1, ..., X_n)$ is called a Gaussian random vector if there exists an $n \times m$ matrix A, and an n-dimensional vector b such that*
$X^T = AY + b$, where Y is an m-dimensional vector with independent standard normal entries.

Proposition 8 *The distribution of a Gaussian vector $X = (X_1, ..., X_n)$ is uniquely determined by its mean vector $(E(X_1), ..., E(X_n))$ and its covariance matrix $C_{ij} = Cov(X_i, X_j)$.*

This could be proved using characteristic functions

Corollary 9 *If $X = (X_1, ..., X_n)$ is a Gaussian, and the X_i-s are pairwise independent, then they are independent.*

Corollary 10 *If X_1 and X_2 are independent $N(0, \sigma^2)$ distributed, then $X_1 + X_2$ and $X_1 - X_2$ are independent and both are $N(0, 2\sigma^2)$ distributed. Also, if $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ are independent, then $X_1 + X_2 \sim \mathbb{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ and $X_1 - X_2 \sim \mathbb{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$*

Proposition 11 *Let $\{X_n\}_{n \geq 0}$ be a sequence of Gaussian random vectors in R^d which converges in distribution to a random vector X, then X is a Gaussian with mean vector and covariance matrix that are resp. the component-wise limits (which exist) of $E(X_n)$ and $Cov(X_n)$*

Proposition 12 *If $X \sim N(0, 1)$, then for $x \geq 1$:*

$$\frac{x}{1 + x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \leq \mathbb{P}(X \geq x) \leq \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

5 BM exists

Definition 13 *Let $\mathcal{D}_n = \{\frac{k}{2^n} : 0 \leqslant k \leqslant 2^n\}$, then the set $\mathcal{D} = \bigcup_{n \geq 0} \mathcal{D}_n$ is the set of dyadic points in $[0, 1]$.*

$\mathcal{D}$ is of course dense, as if x is any point in the interval [0,1], then for each n there is a k such that, $\frac{k}{2^n} \leq x \leq \frac{k+1}{2^n}$, so there is a point in $\mathcal{D}_n$ such that its distance from x is at most $\frac{1}{2^{n+1}}$. Also, a very convenient property of $\mathcal{D}$ for our construction is that $\forall n \, \mathcal{D}_n \subset \mathcal{D}_{n+1}$.

Proof of Wiener theorem: As mentioned before it is sufficient to construct a BM starting at 0. It is also sufficient to construct a BM on the interval $[0, 1]$. Assuming we have such a construction, then we can take a sequence $B_0, B_1, ...$ of independent $C[0, 1]$-valued random functions with the distribution of this process, and define $\{B(t)\}$ by gluing together the parts, more precisely by:

$$B(t) = B_{\lfloor t \rfloor}(t - \lfloor t \rfloor) + \sum_{i=0}^{\lfloor t \rfloor - 1} B_i(1) \quad \forall t \geq 0.$$

One can easily check that this random function satisfy properties 1-4 in the definition of BM.

The idea is to construct the right joint distribution of Brownian motion step by step on the finite sets $\mathcal{D}_n$. We then interpolate the values on $\mathcal{D}_n$ linearly and check that the uniform limit of these continuous functions exists a.s and is a Brownian motion. To do this, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a collection $\{Z_t : t \in \mathcal{D}\}$ of independent, $N(0, 1)$ r.v-s, can be defined. Let $B(0) := 0$ and $B(1) := Z_1$. For each $n \in \mathbb{N}$ we define the random variables $B(d)$, $d \in \mathcal{D}_n$ by induction such that:

1. $\forall\ r < s < t$ in $\mathcal{D}_n$, $B(t) - B(s) \sim N(0, t - s)$, $B(s) - B(r) \sim N(0, s - r)$ and they are independent.
2. The vectors $(B(d) : d \in \mathcal{D}_n)$ and $(Z_t : t \in \mathcal{D} \setminus D_n)$ are independent.
3. Let $I_1 = [a, b]$, $I_2 = [c, d]$ be 2 disjoint intervals contained in [0,1], then for any $x_1, x_2, x_3, x_4 \in \mathcal{D}_n$ such that $x_1, x_2 \in I_1$, $x_3, x_4 \in I_2$ we have that $B(x_1) - B(x_2)$ and $B(x_3) - B(x_4)$ depend on disjoint sets of $(Z_t : t \in \mathcal{D}_n)$, and in particular are independent.

Note that we have already done this for $\mathcal{D}_0 = \{0, 1\}$. Now, assume that we have succeeded in doing so for $\mathcal{D}_{n-1}$. In order to do so also for $\mathcal{D}_n$ we only need to define $B(d)$ for $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ in a way that the three properties remains valid. Set (for $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$)

$$B(d) = \frac{B(d - 2^{-n}) + B(d + 2^{-n})}{2} + \frac{Z_d}{2^{\frac{n+1}{2}}}.$$

Note that the first summand is the linear interpolation of the values of B at the neighbouring points of d (i.e the 2 nearest points in $\mathcal{D}_n$) which lay in $\mathcal{D}_{n-1}$ (thus B is already defined on them and satisfies our induction hypothesis). Therefore, the second property is satisfied by the induction hypothesis and the fact that in the n^{th} stage $(B(d) : d \in \mathcal{D}_n)$ is defined by the previous stage only by adding Z_t-s for $t \in \mathcal{D}_n$. It is also clear that property 3 remains valid after the n^{th} stage of our construction. This can by seen by using property three of the induction hypothesis on the previous stage and the fact that in the n^{th} stage of our construction we added to different intervals different Z_t-s. We first note that in order to prove that property 1 remains valid after the n^{th} stage, we only need to show that the increments $B(d) - B(d-2^{-n})$, $d \in \mathcal{D}_n \setminus \{0\}$, are pairwise independent and have $N(0, 2^{-n})$ distribution. This is sufficient by corollary 9, as the vector of these increments is a Gaussian. We already know by property three (that has already been established) that the pairwise independence is true for increments of disjoint intervals, so it remains only to verify this for increments of intervals with a common edge and while doing so verify that the distribution of the increments is indeed $N(0, 2^{-n})$. Note that by the induction hypothesis $X_1 := \frac{B(d+2^{-n})+B(d-2^{-n})}{2}$ depends only on $(Z_t : t \in \mathcal{D}_{n-1})$ and thus is independent of $X_2 := \frac{Z_d}{2^{\frac{n+1}{2}}}$. By corollary 10 we have that the difference (and sum) of independent normal r.v-s is a normal r.v whose mean is the difference (resp. sum) of their means, and its variance is the sum of their variances. Moreover, if X_1, X_2 are independent with $N(0, \sigma^2)$ distribution, then $X_1 + X_2$ and $X_1 - X_2$ are independent normal r.v-s with mean 0 and $2\sigma^2$ variance. Thus if $X_1 = \frac{B(d+2^{-n})-B(d-2^{-n})}{2}$, $X_2 = \frac{Z_d}{2^{\frac{n+1}{2}}}$, then by the first part of the lemma X_1 is distributed $N(0, 2^{-(n+1)})$ which is the same as X_2. We already saw that X_1 and X_2 are independent, thus since both have $N(0, 2^{-(n+1)})$ distribution, we can apply the second part of the corollary, and obtain that $X_1 - X_2 = B(d + 2^{-n}) - B(d)$ and $X_1 + X_2 = B(d) - B(d - 2^{-n})$ are independent and both have $N(0, 2^{-n})$ distribution. Thus property one is established and the induction step is completed.

We have defined the values of B on $\mathcal{D}$. We now interpolate these values linearly as follows:

$$F_0(t) = \begin{cases} Z_1 & t = 1 \\ 0 & t = 0 \\ \text{linear in between} \end{cases}$$

Similarly for $n \geq 1$,

$$F_n(t) = \begin{cases} \frac{Z_t}{2^{\frac{n+1}{2}}} & t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1} \\ 0 & t \in \mathcal{D}_{n-1} \\ \text{piecewise linear between consecutive points of } \mathcal{D}_n \end{cases}$$

These functions are continuous on $[0, 1]$ and by construction satisfy for all n and $d \in \mathcal{D}_n$,

$$B(d) = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{\infty} F_i(d).$$

To verify that the previous equality indeed follows from our construction we can use induction. It holds for $n = 0$. Suppose it holds for $n - 1$. Let $d \in D_n \setminus \mathcal{D}_{n-1}$. First of all note that for any $m \geq n + 1$ we have $F_m(d) = 0$ (actually this holds $\forall$d$\in D_n$, and this explains why we can take either an infinite sum or the sum until $i = n$). Using the fact the for $0 \leq i \leq n - 1$, F_i is linear on $[d - 2^{-n}, d + 2^{-n}]$, we get:

$$\sum_{i=0}^{n-1} F_i(d) = \sum_{i=1}^{n-1} \frac{F_i(d - 2^{-n}) + F_i(d + 2^{-n})}{2} = \frac{B(d - 2^{-n}) + B(d + 2^{-n})}{2}$$

The first equality follows by the linearity argument mentioned before and the second one by the induction hypothesis. Combine this with $F_n(d) = \frac{Z_d}{2^{-\frac{n+1}{2}}}$ to get the desired equality.

Define for all $t \in [0,1]$, $B(t) = \sum_{i=0}^{\infty} F_i(t)$. We first use an estimate we have on normal r.v-s and the Borel-Cantelli lemma to show that with probability 1 this sum is uniformly convergent (B is a sum of normal r.v-s). This will imply B is continuous (a.s), as the finite sums are continuous by construction, and the infinite sum is B (if g_n are continuous and uniformly converge to g, then g is continuous). We have already shown that the other properties of the BM (concerning finite dimensional distributions of increments) are satisfied on $\mathcal{D}$, which is a dense set in $[0,1]$. Therefore, by continuity we will have the desired finite dimensional distributions of increments for all points.
By proposition 12, we have for $c > 1$ and large enough n,

$$\mathbb{P}\{|Z_d| \geq c\sqrt{n}\} \leq exp(\frac{-c^2 n}{2}).$$

So for $c > \sqrt{2log2}$ we get that the r.h.s is of the form a^n for $0 < a < \frac{1}{2}$, thus by union bound:

$$\sum_{n=0}^{\infty} \mathbb{P}\{\exists d \in \mathcal{D}_n \; s.t \; |Z_d| \geq c\sqrt{n}\} \leq \sum_{n=0}^{\infty} \sum_{d \in \mathcal{D}_n} \mathbb{P}\{|Z_d| \geq c\sqrt{n}\} \leq \sum_{n=0}^{\infty} (2^n+1) exp(\frac{-c^2 n}{2}) \leq \sum_{n=0}^{\infty} 2(2a)^n \leq \infty.$$

So for such c, by Borel-Cantelli lemma, there exists a.s a (random) N such that for all $n \geq N$ and $d \in \mathcal{D}_n :\ |Z_d| < c\sqrt{n}$. Hence for all $n \geq N$,

$$\|F_n\|_\infty < c\sqrt{n} 2^{-\frac{n}{2}}.$$

This upper bound implies that indeed the sum $B(t) = \sum_{i=0}^{\infty} F_i(t)$ is uniformly convergent on [0,1] and thus B is continuous.
It remains only to verify that the increments of this process have the correct finite-dimensional distribution. As mentioned earlier this follows directly from the properties of B on the dense set $\mathcal{D} \subset [0,1]$ and the continuity of B. Indeed let $t_1 < t_2 < ... < t_n$ be in [0,1]. there exist $t_{1,k} \leq t_{2,k} \leq ... \leq t_{n,k}$ in $\mathcal{D}$ with $t_{i,k} \to \ t_i$. Thus by continuity of B for all $0 \leq i \leq n-1$ we have:

$$B(t_{i+1}) - B(t_i) = \lim_{k\to\infty} B(t_{i+1,k}) - B(t_{i,k}).$$

Since $\lim_{k\uparrow\infty} \mathbb{E}[B(t_{i+1,k}) - B(t_{i,k})] = 0$ and
$\lim_{k\uparrow\infty} Cov(B(t_{i+1,k}) - B(t_{i,k}), B(t_{j+1,k}) - B(t_{j,k})) = 1_{\{i=j\}}(t_{i+1,k} - t_{i,k}) = \lim_{k\uparrow\infty} 1_{\{i=j\}}(t_{i+1} - t_i)$, we get that the increments $B(t_{i+1}) - B(t_i)$ are independent normal variables with mean 0 and variance $t_{i+1} - t_i$ as required. □

6 Continuity properties of BM

Definition 14 *For a continuous function f on [0,1], we say that f has modulus of continuity φ, if*

$$\limsup_{h\downarrow 0} \sup_{0\leq t\leq 1-h} \frac{f(t+h)-f(t)}{\varphi(h)} \leq 1$$

Theorem 15 *$\exists C>0$ such that with probability 1, the function $\varphi(h) := C\sqrt{-hlog(h)}$ is a modulus of continuity for the BM B. That is, a.s exist (a random) h_0 such that $|B(t+h)-B(t)| \leq C\sqrt{-hlog(h)}$ for every $h<h_0$, for all $0\leq t\leq 1-h$.*

Proof This will follow from our constructions. Recall that for $t\in[0,1]$ $\quad B(t)=\sum_{i=0}^{\infty}F_i(t)$, and for any $c>\sqrt{2log2}$ with probability 1 exists (a random) N such that for any $n\geq N$ we have
$\|F_n(t)\|_\infty < c\sqrt{n}2^{-\frac{n}{2}}$. Also, by definition F_n is piece-wise linear, and in particular $\exists F_n'(t)$ and for any $n\geq N$ $\|F_n'(t)\|_\infty \leq \frac{2\|F_n(t)\|_\infty}{2^{-n}} \leq 2c\sqrt{n}2^{-\frac{n}{2}}$. Now for each $t,t+h\in[0,1]$, using the mean-value theorem,

$$|B(t+h)-B(t)| \leq \sum_{n=0}^{\infty}|F_n(t+h)-F_n(t)| \leq \sum_{n=0}^{l} h\|F_n'(t)\|_\infty + \sum_{n=l+1}^{\infty} 2\|F_n(t)\|_\infty.$$

Now, if we take $l\geq N$, we get that the last expression is bounded by

$$h\sum_{n=0}^{N}\|F_n'(t)\|_\infty + 2ch\sum_{n=N+1}^{l}\sqrt{n}2^{-\frac{n}{2}} + 2c\sum_{n=l+1}^{\infty}\sqrt{n}2^{-\frac{n}{2}}.$$

We now suppose that h is (again random and) small enough that the first summand is smaller than $\sqrt{-hlog(h)}$ and that l, defined by $2^{-l}<h\leq 2^{-l+1}$, is bigger than N. For this choice of l the second and third summands are also bounded by constant multiples of $\sqrt{-hlog(h)}$ as both sums are dominated by their largest element. This completes the proof
$\square$

Remark 16 *The estimate in the previous theorem is in fact sharp up to the constant, as the next theorem shows*

Theorem 17 *For any constant $c<\sqrt{2}$, a.s, for any $\epsilon>0$, there exist $0<h<\epsilon$ and $t\in[0,1-h]$ such that*

$$|B(t+h)-B(t)| \geq c\sqrt{-hlog(h)}$$

Proof Fix $c < \sqrt{2}$ and define, for integers k, $n \geq 0$, the events

$$A_{k,n} := \{B((k+1)e^{-n} - B(ke^{-n}) > c\sqrt{n}e^{-\frac{n}{2}}\}.$$

Then using proposition 12, for any $k \geq 0$,

$$\mathbb{P}(A_{k,n}) = \mathbb{P}\{B(e^{-n}) > c\sqrt{n}e^{-\frac{n}{2}}\} = \mathbb{P}\{B(1) > c\sqrt{n}\} \geq \frac{c\sqrt{n}}{1+c^2n}\frac{1}{\sqrt{2\pi}}e^{-\frac{c^2n}{2}}$$

By our assumption on c, we have $e^n\mathbb{P}(A_{k,n}) \to \infty$ as $n \uparrow \infty$. Therefore, using $1 - x \leq e^{-x}$, for all real x, we get:

$$\mathbb{P}(\bigcap_{k=0}^{\lfloor e^n - 1\rfloor} A^c_{k,n}) = (1 - \mathbb{P}(A_{0,n}))^{e^n} \leq exp(-e^n\mathbb{P}(A_{0,n})) \to 0.$$

By considering $h = e^{-n}$ one can now see that, for any $\epsilon > 0$,

$$\mathbb{P}\{|B(t+h) - B(t)| \leq c\sqrt{-hlog(h)}\ \forall h \in (0,\epsilon), t \in [0, 1-h]\} = 0.$$

□

Remark 18 *We will not prove this (see (1) page 16 for a proof), but Lévy's modulus of continuity theorem says that:*

$$\limsup_{h\downarrow 0} \sup_{0\leq t\leq 1-h} \frac{B(t+h) - B(t)}{\sqrt{-2hlog(h)}} = 1.$$

We proved one of the inequalities.

References

[1] Yuval Peres and Peter Mörters book: Brownian Motion. Chapter one covers all of the theorems that was proved in today's lecture. See appendix 12 for material on normal distribution and Gaussians.