Statistics 212: Lecture 22 (April 23, 2025)

Picard's Little Theorem and Cameron Martin

Instructor: Mark Sellke

Scribe: Swati Goel

1 Main Results

1.1 Picard's Little Theorem

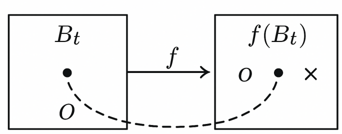
Recall from last time that $f : \mathbb{C} \to \mathbb{C}$ is an *entire, non-constant* function, then $f(B_t)$ is again a planar Brownian motion up to a time–change. Using it, we will outline a proof of the following celebrated complex analysis result.

Picard's Little Theorem. An entire function that omits *two* distinct complex values (without loss of generality -1 and 1) must be constant. (Omitting a single value can be done in many ways: $f(z) = e^{g(z)}$ never equals 0 if g is entire.)

For more details, see Picard's Theorem and Brownian Motion by Burgess Davis.

1.1.1 Proof Outline

The following diagram illustrates a Brownian Motion under f. The new Brownian Motion, $f(B_t)$ should get very tangled:



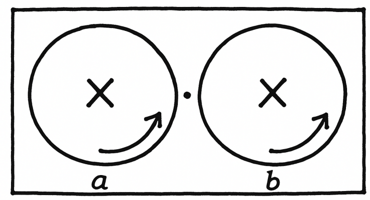
Note: There are no holes in the Brownian Motion on the left, but there can be holes (denoted by X) in the Brownian Motion on the right.

Observation 1. Recall that planar Brownian motion is *neighbourhood recurrent*—there are arbitrarily large times with $|B_t - B_0| \le \varepsilon$. At each such time, $B_{[0,t]}$ forms a loop (modulo the tiny distance between the endpoints) that can be continuously deformed to a single point in $\mathbb{C}$, for example via the deformation:

$$X_t^{(s)} = sB_t, \qquad s \in [0,1].$$

Such a continuous deformation is called a homotopy. Similarly (again up to a tiny distance between the endpoints), the pushforward under f of this loop can be deformed continuously to a point. Namely $Y_t^{(s)} := f\big(X_t^{(s)}\big)$ contracts $f(B_{[0,T]})$. However the latter deformation also avoids $\{-1,1\}$, which is a nontrivial property.

Topological fact: loops in $\mathbb{C} \setminus \{-1,1\}$ (up to homotopy) are in bijection with words in the free group $\{a,b,a^{-1},b^{-1}\}$.



Ex: consider the loop corresponding to the word: $a\,a\,b\,b\,a\,b^{-1}\,b^{-1}\,a\,b^{-1}\,a^{-1}\,a^{-1}$.

Because the group is *not commutative*, the loop represented by $aba^{-1}b^{-1}$ is non-trivial (but aa^{-1}, for example, is trivial).

Observation 2. Consider excursions $w_1,\dots,w_k$ of $f(B_t)$ from a small neighborhood of 0 of radius 0.001, which reach at least 0.1 away from 0. These excursions are approximately IID (e.g. recall the Poisson kernel formula from last time). Label the k-th excursion by a non-trivial word w_k. (Some excursions will come back quickly and give an empty word, but that's won't really matter.) We will argue that after k excursions, the composite word $w_1\cdots w_k$ has length $\mathrm{len}(w_1\cdots w_k)$ growing to ∞ with k.

Specifically, assume without loss of generality, that the last letter of the partial word $w_1\cdots w_{k-1}$ is a. Then

$$\mathrm{len}(w_1\cdots w_k) - \mathrm{len}(w_1\cdots w_{k-1}) \geq \begin{cases} \mathrm{len}(w_k) & \text{if } w_k \text{ does } \textit{not} \text{ start with } a^{-1}, \\ -\mathrm{len}(w_k) & \text{if } w_k \textit{ does} \text{ start with } a^{-1}. \end{cases}$$

Further, the excursion distributions are approximately symmetric, so these cases have respective probabilities approximately $3/4$ and $1/4$, and $|w_k|$ is approximately independent of its first letter. Thus, a law of large numbers style argument should imply that the total length grows to infinity, and in particular is never 0 after some finite time. However this contradicts Observation 1: we saw that every time B_t returns to near 0, the loop $f(B_{[0,t]})$ is homotopic to a point, hence corresponds to the trivial word. This proves the theorem.

A slightly delicate technical point is that the LLN is not actually applicable because $\mathrm{len}(w_k)$ will not be finite. This is because loops can spend a very long time far away from the origin, and make a large number of windings while they are far away. To get around this, Davis considers a modified definition of length which only counts adjacent occurences of $ab^{-1}, a^{-1}b, ba^{-1}, b^{-1}a$. This fixes the issue because the windings correspond to words like $abababab\ldots$, and a law of large numbers style argument can be applied to the modified definition of length.

1.2 Cameron–Martin Theorem

1.2.1 Radon–Nikodym Warm-Up

Let $Z \sim N(0,1)$ under $\mathbb{P}_0$. For any $\mu \in \mathbb{R}$ define

$$\frac{d\mathbb{P}_\mu}{d\mathbb{P}_0} = \exp\big(\mu Z - \tfrac{\mu^2}{2}\big).$$

Under $\mathbb{P}_\mu$ we have $Z \sim N(\mu, 1)$.

Why this warm up? This scalar tilt example is a 1D version of the change of measure idea we'll use to prove the Cameron-Martin Theorem. We'll now apply this idea to the entire Brownian Motion path (as opposed to a single random variable).

1.2.2 Theorem Statement

For Brownian motion B_t with respect to $\mathbb{P}_0$, define

$$Z_\mu(t) = \exp\big(\mu B_t - \tfrac{\mu^2}{2} t\big), \qquad 0 \le t \le T.$$

- Martingale property: $Z_\mu(t)$ is a martingale under $\mathbb{P}_0$.
- Change of measure: Introduce a new measure on the path space by

  $$d\mathbb{P}_\mu = Z_\mu(T)\, d\mathbb{P}_0.$$

- Resulting law: Under $\mathbb{P}_\mu$ the process

  $$\tilde{B}_t := B_t - \mu t$$

  is a standard Brownian motion; equivalently, B_t now has constant drift μ.

Intuition: apply the one–dimensional Radon–Nikodym tilt from the warm-up to every increment $B_{t+\varepsilon} - B_t$ and let $\varepsilon \to 0$.

Proof idea. By Itô's formula,

$$dZ_\mu(t) = \mu Z_\mu(t)\, dB_t + \Big(\tfrac{\mu^2}{2} - \tfrac{\mu^2}{2}\Big) Z_\mu(t)\, dt = \mu Z_\mu(t)\, dB_t,$$

so Z_μ is indeed a martingale. For any $\lambda \in \mathbb{R}$, we computed that

$$\mathbb{E}_{\mathbb{P}_\mu}\big[e^{\lambda B_t}\big] = \mathbb{E}_{\mathbb{P}_0}\big[e^{\lambda B_t} Z_\mu(t)\big] = \exp\big(\tfrac{\lambda^2}{2} t + \lambda \mu t\big),$$

which is the moment generating function (mgf) of $N(\mu t, t)$ and confirms the drift.

Note: moments do not uniquely determine a distribution in general, but if the moments don't grow too fast such that the mgf exists in a neighborhood of 0, this *does* uniquely define a distribution thanks to uniqueness of analytic continuation.

2 Looking Ahead

Some ideas we'll cover in more depth in the next lecture (don't worry too much about this for now):

(a) For a given μ_s we can tilt by $\displaystyle\int_0^\tau \mu_s\, dB_s$ to obtain a Brownian motion with drift μ_s.

(b) In $\mathbb{R}^d$, if $Z \sim N(0, I_d)$ and we tilt by $\exp\langle \mu, Z\rangle$, we see a mean shift from 0 to μ.
This is the d dimensional version of (a).

(c) Brownian motion can be viewed as a Gaussian vector for the Hilbert space

$$H^1 = \Big\{ f : [0,\tau] \to \mathbb{R} : f(0) = 0,\ f' \in L^2 \Big\}, \qquad \langle f, g\rangle_{H^1} = \int_0^\tau f'(s) g'(s)\, ds.$$

(d) If $f(s) = \int_0^s \mu_r\, dr$, then $\langle f, B\rangle_{H^1} = \int_0^\tau \mu_r\, dB_r$.
Thus adding drift corresponds to translating Brownian motion by an element of H^1.