A Summary of Brownian Motion.[1]

Definition. A STANDARD BROWNIAN MOTION $W = W(t)$, $t \geq 0$, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a collection of random variables $W(\omega, t)$ such that

(1) $W(0) = 0$;
(2) For every $0 < t_1 < \cdots t_n$, the vector $\big(W(t_1), \ldots W(t_n)\big)$ is Gaussian;
(3) $\mathbb{E}W(t) = 0$, $\mathbb{E}\big(W(t)W(s)\big) = \min(t, s)$, $t, s \geq 0$;
(4) For every $\omega \in \Omega$, the function $t \mapsto W(\omega, t)$ is continuous.

A standard Brownian motion in $\mathbb{R}^{\mathrm{d}}$, $\mathrm{d} > 1$, is the vector-valued process $t \mapsto (W_1(t), \ldots, W_{\mathrm{d}}(t))$, with independent standard Brownian motions W_k.

Note. If $W(0) = 0$, then $\mathbb{E}\big(W(t)W(s)\big) = \min(t, s)$ is equivalent to $\mathbb{E}|W(t) - W(s)|^2 = |t - s|$. Because W is Gaussian, the Kolmogorov continuity criterion implies that W has a continuous modification (in fact, Hölder continuous of every order less that $1/2$). It also follows that W has independent and stationary increments.

Three constructions of W.

(1) RANDOM WALK APPROXIMATION [Donsker, 1951]: $W_h(t),\ t \geq 0$, is the linear interpolation of $(t_n, W_h(t_n)),\ n \geq 0$, with

$$W_h(t_n) = W_h(t_{n-1}) + \sqrt{h}\xi_n,\ W_h(0) = 0, \quad t_n = nh,\ \xi_n \text{ iid } \mathcal{N}(0, 1).$$

If $\xi_n = \frac{W(t_n) - W(t_{n-1})}{\sqrt{h}}$, then, for $t \in [t_{n-1}, t_n]$ we have $W_h(t) = W(t_{n-1}) + (t - t_{n-1})\big(W(t_n) - W(t_{n-1})\big)/h$ and so $W - W_h$ is a Brownian bridge on $[t_{n-1}, t_n]$ [which, for most practical purposes, is the same as $\sqrt{h}B_0$, where B_0 is the Brownian bridge on $[0, 1]$.] As a result, the process $t \mapsto W(t) - W_h(t), t \geq 0$, is, in distribution, a collection of independent Brownian bridges. Then, keeping in mind that $B_0(t)$ is Gaussian with mean zero and variance $t(1-t)$,

$$\mathbb{E}\int_0^1 |W_h(t) - W(t)|\, dt = \sqrt{h}\int_0^1 \mathbb{E}|B_0(t)|\, dt = \sqrt{\frac{\pi h}{2}}\int_0^1 \sqrt{t(1-t)}\, dt = \sqrt{\frac{\pi h}{2}} B\left(\frac{3}{2}, \frac{3}{2}\right) = \sqrt{\frac{\pi}{32}}\sqrt{h}.$$

Similarly,[2]

$$\lim_{h \to 0+} \frac{1}{\sqrt{h|\ln h|}}\, \mathbb{E}\max_{0 \leq t \leq 1} |W_h(t) - W(t)| = \frac{1}{\sqrt{2}}.$$

(2) CHAOS EXPANSION: If $\{m_k(t),\ k \geq 1\}$ is an orthonormal basis in $L_2((0, T))$, then, for $t \in [0, T]$,

$$W(t) = \sum_{k=1}^{\infty} \overline{m}_k(t)\xi_k,\ \overline{m}_k(t) = \int_0^t m_k(s)\, ds,\ \xi_k \text{ iid } \mathcal{N}(0, 1).$$

The series converges in $L_2(\Omega \times [0, T])$ and with probability one for every t. A particular choice of the basis, such as the Fourier cosine basis or the Haar basis, makes it possible to establish uniform convergence. The Haar basis also corresponds to the popular *bisection method* for constructing W.

[1] Sergey Lototsky, USC. Most recent update on July 26, 2022.
[2] S. Asmussen, P. W. Glynn, *Stochastic Simulation* (Springer, 2007), Proposition X.2.1

(3) KL (Karhunen-Loève) expansion [around 1945]:

$$W(t)=\sqrt{\frac{2}{T}}\sum_{n=1}^{\infty}\frac{\sin(\lambda_n t)}{\lambda_n}\,\xi_n,\ \lambda_n=\left(n-\frac{1}{2}\right)\frac{\pi}{T},\ \xi_n \text{ iid } \mathcal{N}(0,1).$$

This is a particular case of chaos expansion, with $m_k(t)=\sqrt{2/T}\cos(\lambda_k t)$.

"Easy" properties of the Brownian motion.

(1) Scaling: for every $c>0$, the process $t\mapsto \sqrt{c}\,W(t/c)$ is a standard Brownian motion;
(2) Time reversal: the process $X(t)=tW(1/t)$, with $X(0)=0$, is a standard Brownian motion;
(3) Reflection principles:
  (a) if $\tau_a=\inf\{t>0: W(t)=a\}$, $a\neq 0$, then $\mathbb{P}(\tau_a\leq t)=2\mathbb{P}\big(W(t)>a\big)=\mathbb{P}\big(M(t)>a\big)$;
  (b) if $M(t)=\max_{0\leq s\leq t}W(s)$, then $M-W\stackrel{\mathcal{L}}{=}|W|$;
  (c) if τ is a stopping time, then the process $t\mapsto W(t)1(t\leq\tau)+\Big(2W(\tau)-W(t)\Big)1(t\geq\tau)$ is a Brownian motion.

"Highly non-trivial" properties of the Brownian motion.

(1) Continuity of the Brownian filtration: if $\mathcal{F}_t^W=\sigma(W(s),\ s\leq t)$ and is $\mathbb{P}$-complete (contains all $\mathbb{P}$-null sets), then $\mathcal{F}_t^W=\sigma\Big(\bigcup_{s<t}\mathcal{F}_s^W\Big):=\mathcal{F}_{t-}^W$, $t>0$ (by continuity of W), and $\mathcal{F}_t^W=\bigcap_{s>t}\mathcal{F}_s^W:=\mathcal{F}_{t+}^W$, $t\geq 0$ (by Blumenthal's $0-1$ law).
(2) Given a real number C, the process $t\mapsto C\max_{0\leq s\leq t}W(s)-W(t)$ is NOT Markov except for three special values of C: $C=0$ (obvious), $C=1$ (reflection principle), and $C=2$ (in this case, we get the 3-Bessel process, that is, the Euclidean norm of the standard Brownian motion in $\mathbb{R}^3$: J. W. Pitman, One-dimensional Brownian motion and the three-dimensional Bessel process, Advances in Appl. Probability, Vol. 7, No. 3, pp. 511–526, 1975).
(3) Skorokhod embedding/representation: If X is a square-integrable random variable with zero mean, then there exists a stopping time τ relative to $\mathcal{F}_t^W$ such that $W(\tau)$ has the same distribution as X and $\mathbb{E}X^2=\mathbb{E}\tau$. In particular, if X only takes two values a and b, then $\tau=\min\{t>0: W(t)\notin[a,b]\}$.

Sample Path Properties.

(1) Finite quadratic variation: if $t_{k,n}=Tk/n$, $k=0,\ldots,n$, $n=1,2,\ldots$, then

$$\lim_{n\to\infty}\sum_{k=1}^{n}|W(t_{k,n})-W(t_{k-1,n})|^2=T,$$

both with probability one and in $L_2(\Omega;\mathbb{P})$. As a result, with probability one, $\lim_{n\to\infty}\sum_{k=1}^{n}|W(t_{k,n})-W(t_{k-1,n})|^p=0$ $[p>2]$ or $+\infty$ $[0<p<2]$.
(2) Hölder continuity: With probability one, sample paths of W are Hölder continuous of any order less that $1/2$; with probability zero, sample paths of W are Hölder continuous of any order bigger that $1/2$;
(3) The Law of Iterated Logarithm: with probability one, as $t\to 0+$ or $t\to+\infty$, the set of limit points, of $W(t)/\sqrt{2t|\ln|\ln t||}$ is $[-1,1]$;

(4) Modulus of continuity: with probability one,

$$\limsup_{h\to 0+} \sup_{0<|t-s|<h} \frac{|W(t)-W(s)|}{\sqrt{2h|\ln h|}} = 1.$$

(5) At individual points in time:
- almost all (with respect to the Lebesgue measure) points $t > 0$ are `regular (ordinary)`: with probability one,

$$\limsup_{h\to 0} \frac{|W(t+h)-W(t)|}{\sqrt{2h|\ln|\ln|h|||}} = 1;$$

- Some points $t > 0$ are `rapid`: with probability one,

$$\limsup_{h\to 0} \frac{|W(t+h)-W(t)|}{\sqrt{h|\ln|h||}} > 0;$$

- Some other points $t > 0$ are `slow`: with probability one,

$$\limsup_{h\to 0} \frac{|W(t+h)-W(t)|}{\sqrt{|h|}} < \infty;$$

(6) Square root laws: While existence of rapid points is suggested by the modulus of continuity, existence of slow points is somewhat less obvious and leads to further discoveries. For example, for $c > 0$, define the (random) set

$$T_c = \left\{ t \in [0,1] : \limsup_{h\to 0} \frac{|W(t+h)-W(t)|}{\sqrt{|h|}} \le c \right\}.$$

For "small" c, we would expect the set T_c to be empty. Indeed, Dvoretsky (1963) showed that $\mathbb{P}(T_c = \emptyset) = 1$ if $c < 1/4$; Burgers Davis (1983) improved it to $c < 1$.

Deeper into stochastic analysis.

(1) (Strong) Markov Property: W has it, relative to $\mathcal{F}_t^W$, because the increments are independent;
(2) Martingale Property: W is a square integrable martingale relative to $\mathcal{F}_t^W$, with $\langle W\rangle(t) = t$ [that is, $W^2(t) - t$ is a martingale], again because of independent increments;
(3) Lévy characterization of the Brownian motion (1948): A `Wiener process`, that is, a continuous, square-integrable martingale $V = V(t)$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ satisfying the usual assumptions, with $\langle V\rangle(t) = t$, is a Brownian motion.
(4) Random time change/Dambis-Dubinis-Schwarz theorem (1965): If $N = N(t)$ is a continuous (local) martingale, with $\lim_{t\to\infty}\langle N\rangle(t) = +\infty$, and, for $u \ge 0$, $\tau(u) = \inf\{t > 0 : \langle N\rangle(t) = u\}$, then $V(u) = N(\tau(u))$ is a standard Brownian motion, adapted to $\mathcal{F}_{\tau(u)}$, and $N(t) = V\big(\langle N\rangle(t)\big)$.
(5) The filtration CAN make a difference: if W is a standard Brownian motion and

$$V(t) = W(t) - \int_0^t \frac{W(u)}{u}\,du,$$

then V is a standard Brownian motion [V is (obviously) Gaussian, and, by direct computation, $\mathbb{E}V(t) = 0$ and $\mathbb{E}|V(t) - V(s)|^2 = |t - s|$], and is therefore V a martingale relative to its own filtration $\mathcal{F}_t^V$, but V is not a martingale relative to $\mathcal{F}_t^W$. Indeed, if $t > s > 0$, then

$$\mathbb{E}\Big(V(t)|\mathcal{F}_s^W\Big) = W(s) - \int_0^s \frac{W(u)}{u}\,du - \int_s^t \frac{W(s)}{u}\,du = V(s) - W(s)\ln(t/s).$$

In other words, $\mathcal{F}_t^V \subsetneqq \mathcal{F}_t^W$ (strict inclusion).

(6) Markov, not Strong Markov: this is only possible in continuous time. Define the function $f : \mathbb{R} \to \mathbb{R}^2$ by

$$f(x) = \begin{cases} (x, 0), & x < 0; \\ (\sin x, 1 - \cos x), & 0 \le x \le 2\pi; \\ (x - 2\pi, 0), & x > 2\pi. \end{cases}$$

The curve $x \mapsto f(x)$ is the x-axis together with the unit circle centered at $(0, 1)$. If $W = W(t)$ is a standard Brownian motion, then the process $X(t) = f\big(W(t)+\pi\big)$ is Markov [the inverse f^{-1} exists everywhere except for $(0, 0)$, so, with probability one, $W(t) = f^{-1}(X(t)) - \pi$.] On the other hand, we cannot "restart" X at the stopping time $\tau = \inf\{t > 0 : |W(t)| > \pi\}$ because the behavior of $X(t)$, $t > \tau$, will depend on whether $W(\tau) = \pi$ or $W(\tau) = -\pi$.

A time line.

(1) **1827**: the random motion of pollen in water observed, through a microscope, by the English botanist Robert Brown (1773–1858), who started his career by dropping out of medical school and enlisting in the Royal Navy.
(2) **1905**: the physical theory based on random walk and the heat equation is developed by Albert Einstein (1879–1955), who, at the same time, also developed special relativity.
(3) **1906**: the same physical theory is developed, independently, by the Polish physicist Marian Smoluchowski (1872–1917), who, in his spare time, was doing skiing, mountain climbing, watercolor painting, and piano playing.
(4) **1923**: the mathematical construction, as a measure on the space of continuous functions, was presented by Norbert Wiener (1894–1964), whose father was related to Maimonides.
(5) **1933**: axiomatic (measure-theoretic) approach to probability is developed by the Soviet mathematician Andrey Nikolaevich Kolmogorov (1903–1987), who started as a history major, but quickly switched to mathematics.
(6) **1945**: the stochastic calculus is developed by the Japanese mathematician Kiyosi Itô, who was writing his papers not only in Japanese, but also in Chinese, English, French, and German.
(7) **1948**: the book *Processus Stochastiques et Mouvement Brownien* is published by the French mathematician Paul Lévy (1886–1971), who earlier (1934) introduced the concept of martingale as a stochastic process, and whose son-in-law was Laurent Schwartz,
(8) **1945–1950**: extensive computations with the Wiener process/Bronwian motion are carried out by American mathematicians William Ted Martin (1911–2004), who chaired the math department at MIT from 1947 to 1968, and Robert Horton Cameron (1908–1989), who supervised 35 Ph.D. students during his 30+ years at the University of Minnesota, including M. Donsker.
(9) **1951**: the mathematical justification of the Einstein-Smoluchowski construction, in the form of the functional central limit theorem/invariance principle, is provided by the American mathematician Monroe Donsker (1924–1991), who also served as Chair of the Board of Foreign Scholarships in the Ford and Carter administrations.