Proof. Since the martingale property of ζ ensures that the family of measures induced on $\mathcal{F}_t^0$, $t \geq 0$, are consistent, the existence of the measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F}^0)$ follows from the Daniell-Kolmogorov theorem. The rest of the theorem follows by replacing γ by $\gamma \mathbb{I}_{[0,t]}$ in Theorem 2.6. □

Remark 2.8. Theorem 2.7 also holds with $\mathcal{F}_t^0$ replaced everywhere by $\mathcal{F}_{t+}^0$. Note that W remains a Brownian motion relative to $\mathcal{F}_{t+}$ and so this allows a little more flexibility.

Remark 2.9. Under the conditions of Theorem 2.7, observe that for every $T < \infty$ and functional $f : \mathcal{C}[0,T] \mapsto \mathbb{R}$, we have

$$\mathbb{E}_{\tilde{\mathbb{P}}}\left[f(W_s, s \leq t)\right] = \mathbb{E}_{\mathbb{P}}\left[f(W_s, s \leq t)\zeta_T\right] \tag{2.17}$$

2.3. **Conditions for exponential martingales to be true martingales.** The statement of Theorem 2.6 motivates the need to identify general conditions under which, given a local martingale M, the corresponding exponential martingale $\mathcal{E}(M)$ is actually a (uniformly integrable) martingale. Due to the fact that $\mathcal{E}(M)$ is a supermartingale (see (2.11)) and the fact that a supermartingale with constant expectation is a martingale, this is equivalent to asking when

$$\mathbb{E}\left[\mathcal{E}(M)_t\right] = 1 \qquad \forall t. \tag{2.18}$$

The main criteria here are Kazamaki's and Novikov's conditions. But first we show that (2.18) holds in the bounded case.

A. The Bounded Case

Theorem 2.10. *Suppose that for each t there exists a constant K_t such that*

$$\langle M \rangle_t < K_t \text{ a.s.} \tag{2.19}$$

Then for every t, and every $y > 0$,

$$\mathbb{P}\left(\max_{s \leq t} M_s > y\right) \leq \exp\left(-y^2/2K_t\right) \tag{2.20}$$

and $\mathcal{E}(\theta M)$ is a martingale for any $\theta \in \mathbb{R}$.

Proof. Since $\mathcal{E}(\theta M)$ is a non-negative supermartingale such that $\mathcal{E}(\theta M)_0 = 1$, for $\alpha > 0, y > 0$, we have

$$\begin{aligned}
\mathbb{P}\left(\max_{s \leq t} M_s > y\right) &= \mathbb{P}\left(\exp\left(\alpha \max_{s \leq t} M_s\right) > \exp(\alpha y)\right) \\
&\leq \mathbb{P}\left(\max_{s \leq t} \mathcal{E}(\alpha M)_s > \exp\left(\alpha y - \frac{1}{2}\alpha^2 K_t\right)\right) \\
&\leq \exp\left(-\alpha y + \frac{1}{2}\alpha^2 K_t\right),
\end{aligned}$$

where the second-last inequality uses (2.19) and the last line uses Doob's inequality for the supermartingale $\mathcal{E}(\alpha M)$ and (2.11). Substituting $\alpha = y/K_t$ yields (2.20).

Now define $S_t \doteq \max_{s\in[0,t]} M_s$. Then, for any $\theta > 0$, using an integration by parts and the inequality (2.20), we obtain

$$\begin{aligned}\mathbb{E}\left[e^{\theta S_t}\right] &= \int_{\mathbb{R}} e^{\theta y}\mathbb{P}\left(S_t \in dy\right) \\ &= 1+\theta\int_{\mathbb{R}} e^{\theta y}\mathbb{P}\left(S_t > y\right)dy \\ &\leq 1+\theta\int_{\mathbb{R}} e^{\theta y}\exp\left(-y^2/2K_t\right)dy < \infty.\end{aligned}$$

Let $T^{(n)}$ be a localizing sequence for $\mathcal{E}(\theta M)$. Then

$$\mathbb{E}\left[\mathcal{E}(\theta M)_{t\wedge T^{(n)}} \mid \mathcal{F}_s\right] = \mathcal{E}(\theta M)_{s\wedge T^{(n)}}.$$

However, $\mathcal{E}(\theta M)_{t\wedge T^{(n)}} \leq e^{\theta S_t}$ for all $n \in \mathbb{N}$. So, letting $n \to \infty$ and using the dominated convergence theorem, we have

$$\mathbb{E}\left[\mathcal{E}(\theta M)_t \mid \mathcal{F}_s\right] = \mathcal{E}(\theta M)_s.$$

For $\theta < 0$, simply consider θM as $(-\theta)(-M)$. ☐

Remark 2.11. As a direct corollary of the above result, we see that Theorem 2.7 holds whenever γ is a.s. bounded. Indeed, apply the last theorem with $\theta = 1$ and $M = \int \gamma dW$ and note that $\langle M\rangle_t = \int_0^t |\gamma|^2\, dt$.

B. Kazamaki's and Novikov's conditions

Before we state the main conditions, we need to establish some preliminary results.

Lemma 2.12. *Suppose $M \in M^{c,\mathrm{loc}}$. Then*

$$\mathbb{E}\left[e^{\frac{1}{2}M_t}\right] \leq \mathbb{E}\left[e^{\frac{1}{2}\langle M\rangle_t}\right]^{\frac{1}{2}}.$$

Proof. Straightforward algebra shows that

$$(\mathcal{E}(M)_t)^{\frac{1}{2}} = \left[e^{M_t-\frac{1}{2}\langle M\rangle_t}\right]^{\frac{1}{2}} = e^{\frac{1}{2}M_t}\left(e^{-\frac{1}{2}\langle M\rangle_t}\right)^{\frac{1}{2}},$$

which implies that

$$e^{\frac{1}{2}M_t} = (\mathcal{E}(M)_t)^{\frac{1}{2}}\left(e^{\frac{1}{2}\langle M\rangle_t}\right)^{\frac{1}{2}}.$$

Using the Cauchy-Schwarz inequality and the fact that $\mathbb{E}\left[\mathcal{E}(M)_t\right] \leq 1$, the lemma follows. ☐

Lemma 2.13. *Suppose $M \in M^{c,\mathrm{loc}}$ and let $1 < p < \infty$. If*

$$\sup_{\substack{\tau:\\ \tau \text{ bounded}\\ \text{stopping time}}} \mathbb{E}\left[e^{\frac{\sqrt{p}}{2\sqrt{p}-1}M_\tau}\right] < \infty$$

then $\mathcal{E}(M)$ is an $\mathbb{L}^q$ bounded martingale, where q satisfies $1/p + 1/q = 1$.

Proof. Let $1 < p < \infty$ and $r \doteq (\sqrt{p}+1)/(\sqrt{p}-1)$. Then $s = (\sqrt{p}+1)/2$ and $1/r + 1/s = 1$. Also, note that

$$\left(q - \sqrt{\frac{q}{r}}\right) s = \frac{\sqrt{p}}{2(\sqrt{p}-1)}.$$

Then $\mathcal{E}(M)^q = e^{qM - \frac{q}{2}\langle M\rangle} e^{(q-\sqrt{\frac{q}{r}})M}$. Applying Hölder's inequality for a bounded stopping time τ, we obtain

$$\begin{aligned}\mathbb{E}\left[\mathcal{E}(M)^q_\tau\right] &= \mathbb{E}\left[e^{\sqrt{qr}M_\tau - \frac{qr}{2}\langle M,M\rangle_\tau}\right]^{\frac{1}{r}} \mathbb{E}\left[e^{(q-\sqrt{\frac{q}{r}})M_\tau}\right]^{\frac{1}{s}} \\ &= \left(\mathbb{E}\left[\mathcal{E}\left(\sqrt{qr}M\right)_\tau\right]\right)^{\frac{1}{r}} \left(\mathbb{E}\left[e^{(\frac{\sqrt{p}}{2\sqrt{p}-1})}M_\tau\right]\right)^{\frac{1}{s}}\end{aligned}$$

Since the first term on the right-hand side is bounded by 1 due to (2.11), it follows that $\mathcal{E}(M)_\tau$ is L^q bounded. Since $q > 1$, this implies that the family $\mathcal{E}(M)_\tau$, τ a bounded stopping time, is uniformly integrable and thus $\mathcal{E}(M)$ is of class DL. Since a local martingale of class DL is a martingale, it follows that $\mathcal{E}(M)$ is an L^q bounded martingale. □

Theorem 2.14. (Kazamaki's criterion) *Suppose $M \in M^{c,\text{loc}}$. If*

$$\sup_{\substack{\tau:\\ \tau \text{ bounded} \\ \text{stopping time}}} \mathbb{E}\left[e^{\frac{1}{2}M_\tau}\right] < \infty \tag{2.21}$$

then $\mathcal{E}(M)$ is a uniformly integrable martingale.

Proof. Let $0 < a < 1$ and $p > 1$ be such that $\sqrt{p}/(\sqrt{p}-1) < 1/a$. Lemma 2.13, along with (2.21), shows that $\mathcal{E}(aM)$ is an $\mathbb{L}^q$-bounded martingale, where $1/p + 1/q = 1$, which in turn implies it is a uniformly integrable martingale. However,

$$\begin{aligned}\mathcal{E}(aM) &= e^{aM - \frac{a^2}{2}\langle M\rangle} = e^{a^2 M - \frac{a^2}{2}\langle M\rangle} e^{a(1-a)M} \\ &= \mathcal{E}(M)^{a^2} e^{a(1-a)M}.\end{aligned}$$

Hölder's inequality with a^{-2} and $(1-a^2)^{-1}$ yields

$$\begin{aligned}1 \overset{\text{u.i.}}{=} \mathbb{E}\left[\mathcal{E}(aM)_\infty\right] &= \mathbb{E}\left[\mathcal{E}(M)^{a^2}_\infty e^{a(1-a)M_\infty}\right] \\ &\leq \left(\mathbb{E}\left[\mathcal{E}(M)_\infty\right]\right)^{a^2} \left(\mathbb{E}\left[e^{\frac{a}{1+a}M_\infty}\right]\right)^{1-a^2} \\ &\leq \left(\mathbb{E}\left[\mathcal{E}(M)_\infty\right]\right)^{a^2} \left(\mathbb{E}\left[e^{\frac{1}{2}M_\infty}\right]\right)^{2a(1-a)},\end{aligned}$$

where the last equality holds because, since $a < 1$ ensures that $(1+a)/2a > 1$, the $\mathbb{L}^1$ norm is less than the $\mathbb{L}^{(1+a)/2a}$ norm. Let a increase to 1 and note that then the second term on the right side of the last inequality above converges to 1 since $2a(1-a) \to 0$. Thus, $1 \leq \mathbb{E}\left[\mathcal{E}(M)_\infty\right]$, and since we know that the opposite inequality $\mathbb{E}\left[\mathcal{E}(M)_\infty\right] \leq 1$ always holds, we are done. □

As a corollary, we obtain Novikov's criterion.

Theorem 2.15. (Novikov's Criterion) *Let* $M \in M^{c,\mathrm{loc}}$ *and suppose that*

$$\mathbb{E}\left[e^{\frac{1}{2}\langle M\rangle_\infty}\right] < \infty.$$

Then $\mathcal{E}(M)$ *is a uniformly integrable martingale.*

Proof. By Lemma 1, we have, for any bounded stopping time τ,

$$\mathbb{E}\left[e^{\frac{1}{2}M_\tau}\right] \leq \left(\mathbb{E}\left[e^{\frac{1}{2}\langle M\rangle_\tau}\right]\right)^{\frac{1}{2}}.$$

The theorem then follows via an application of Kazamaki's criterion. □

Remark 2.16.

1. Even though Kazamaki's condition is slightly stronger than Novikov's condition, it can be shown that the constant $1/2$ in Novikov's condition is the best possible. Indeed, the proof with a $1/2$ replaced by $1/2+\varepsilon$, for some $\varepsilon > 0$, is considerably simpler. Novikov's condition is often easier to verify.
2. When combined with Theorem 2.7, Novikov's condition shows that if
$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T |H_s|^2\, ds\right)\right] < \infty \quad \forall 0 \leq T \leq \infty$$
then
$$\zeta_t = \exp\left(\int_0^t H_s dB_s - \frac{1}{2}\int_0^t |H_s|^2\, ds\right)$$
is a martingale on $[0,T]$ and so, $\forall T < \infty$,
$$\tilde{W}_t = W_t - \int_0 H_s\, ds \quad 0 \leq t < T,$$
is a Brownian motion on $\left(\Omega, \mathcal{F}_T^0, \tilde{\mathbb{P}}_T\right)$, where $\frac{d\tilde{\mathbb{P}}_T}{d\mathbb{P}} = \zeta_T$ on $\mathcal{F}_T^0$.

2.4. **Applications of Girsanov's Theorem.** In this section we motivate the form of Girsanov's theorem by considering importance sampling for random variables, and then illustrate an application of Girsanov's theorem to compute the distribution of hitting times of Brownian motion with drift.

A. Importance Sampling. We will describe how changes of measure serve as a convenient tool for designing the simulation of rare events (i.e., events with very small probabilities). Suppose we want to calculate $\mathbb{E}[f(X)]$, where f is a known function and X is a random variable with a given distribution. For the sake of concreteness and to help motivate the form of Girsanov's theorem, in the discussion below, we will assume $X \sim \mathcal{N}(0,1)$, i.e., X is a standard normal random variable, although the general philosophy of importance sampling does not depend on the particular form of the normal distribution. Let $\mathbb{E}_0$ represent expectation under the standard normal distribution. A natural (but, as we will see, sometimes naive) way to calculate

$\mathbb{E}_0[f(X)]$ is to use direct simulation. In other words, one can try to approximate $\mathbb{E}_0[f(X)]$ as follows:

$$\mathbb{E}_0[f(X)] \approx \frac{1}{n}\sum_{i=1}^{n} f(X_i), \tag{2.22}$$

where X_i are i.i.d. samples of X (i.e., pseudorandom numbers generated by a computer that are advertised to be independent and $\mathcal{N}(0,1)$ distributed). Assuming $\mathbb{E}_0[f(X)] < \infty$, the strong law of large numbers tells us that (for almost every realization) as $n \to \infty$, the right-hand side of (2.22) converges to the left-hand side. The question then arises as to how large n has to be in order for the approximation to be "good", i.e., for the relative absolute error

$$\frac{\mathbb{E}[f(X)] - \frac{1}{n}\sum_{i=1}^{n} f(X_i)}{\mathbb{E}[f(X)]}$$

to be smaller than some specified value $\varepsilon > 0$. In order for the direct simulation to be computationally feasible, you need this value n to be of a reasonable magnitude.

However, we provide an illustrative example of when this can fail to happen, so that simulation by the naive approximation in (2.22) is not feasible. Suppose $f(x) = \mathbb{I}_{[30,\infty)}(x)$, so that we want to estimate

$$\mathbb{P}(X \geq 30) = \mathbb{E}_0[f(X)].$$

Let $\mathcal{N}$ be the number of samples that must be generated in order to obtain even one sample that is no less than 30 (note that for $n < \mathcal{N}$, the sum on the right-hand side of (2.22) is zero, which is clearly a very bad approximation to the desired probability, giving a relative error of 1). Explicit calculations, taking into account the rapid decay of the tail probability of the standard normal distribution show that $\mathbb{E}_0[\mathcal{N}] > 10^{100}$. Moreover, we would require 100 or so non-zero summands before one is confident that the first significant digit of $\mathbb{E}_0[f(X)]$ has been correctly calculated. This makes the calculation of $\mathbb{E}_0[f(X)]$ via direct simulation infeasible from a computational point of view.

On the other hand, if we had wanted to estimate $\mathbb{E}_0[h(X)]$, where $h(x) = \mathbb{I}_{[0,\infty)}$, then both the standard normal samples and the most informative behavior of h would have been centered at zero and so a direct simulation would have been completely feasible (in the sense that we would have been able to achieve a high degree of accuracy by generating a moderate number of samples).

The idea behind importance sampling is to find another probability measure (in this case, other than the standard normal probability measure) under which the event whose probability we want to estimate is no longer a rare event. We then generate samples according to this other probability measure, and make appropriate modifications in our estimate in order to obtain the probability of this event under the original probability measure.

For example, in the example described above, we will shift the mean of the normal distribution to 30. (observing that f under this shifted measure looks very much like h does under the original measure). Under this new distribution the event f is much more likely to happen and so it should be much easier to generate relative samples.

More precisely, for $\mu \in \mathbb{R}$, let $\mathbb{E}_\mu$ denote expectation under the probability distribution induced by the random variable $X \sim N(\mu, 1)$ with mean μ. We can then rewrite $\mathbb{E}_0$ expectations in terms of $\mathbb{E}_\mu$ expectations as follows:

$$\begin{aligned}
\mathbb{E}_0[f(X)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-x^2/2}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-\mu x + \mu^2/2} e^{-(x-\mu)^2/2}\, dx \\
&= \mathbb{E}_\mu \left[f(X) e^{-\mu X + \mu^2/2} \right].
\end{aligned}$$

In other words, we have shown that

$$\mathbb{E}_0[f(X)] = \mathbb{E}_\mu[g(X)], \tag{2.23}$$

where

$$g(x) = f(x) e^{-\mu x + \mu^2/2}.$$

Thus, using direct simulation to approximate the quantity on the right-hand side of (2.23), we can try the approximation

$$\mathbb{E}_0[f(x)] \approx \frac{1}{n} \sum_{i=1}^{n} g(Y_i),$$

where Y_i are independent, $N(\mu, 1)$ samples. The quantity $e^{-\mu x + \mu^2/2}$ is referred to as the likelihood ratio between the two probability distributions associated with $\mathbb{E}_\mu$ and $\mathbb{E}_0$.

The question now arises as to how to choose μ wisely so as to improve the "efficiency" of the simulation. Optimally, one would like to choose the μ that minimizes the coefficient of variation

$$c(\mu) = \frac{\mathbb{E}[g(Y_1)]}{[\mathrm{Var}(g(Y_1))]^{\frac{1}{2}}}.$$

However, $c(\mu)$ is typically not known as a function of μ – indeed, $c(0)$ equals the probability we were trying to estimate in the first place, which is unknown. There are other theoretically justifiable ways to choose μ that we wont go into here (it is a separate topic of interest on its own), but it is not hard to see from the above discussion that the heuristic choice of $\mu = 30$ would lead to a more efficient simulation than the original direct simulation. In this context, $\mathbb{E}_\mu$ is sometimes referred to as the *tilted* distribution.

In a similar fashion, we can also perform a "tilting" when we want to evaluate probabilities associated with a process. Consider Brownian motion with drift

$$X_t = B_t + \mu t$$

and suppose that we want to calculate the expectation

$$\mathbb{E}[f(X_{t_1}, X_{t_2}, \dots X_{t_n})],$$

where $0 = t_0 < t_1 < t_2 < \dots < t_n \leq T$ for some bounded Borel function f. In order to exploit the independence of the increments of Brownian motion, we first find g such that

$$f(x_1, \dots, x_n) = g(x_1, x_2 - x_1, \dots, x_n - x_{n-1}).$$

We can write the density of $(X_{t_1}, X_{t_2} - X_{t_1}, \dots X_{t_n} - X_{t_{n-1}})$ as the product of the normalizing constant

$$\frac{1}{(2\pi)^{-n/2}} t_1^{-1/2} (t_2 - t_1)^{-1/2} \dots (t_n - t_{n-1})^{-1/2}$$

and an exponential function of the x_i's (where we set $x_0 \doteq 0, t_0 \doteq 0$):

$$\prod_{i=1}^{n} \exp\left(-\{(x_i - x_{i-1}) - \mu(t_i - t_{i-1})\}^2 / 2(t_i - t_{i-1})\right).$$

Expanding the quadratic components and collecting terms, we obtain an alternating representation, containing a product where terms containing μ are aggregated, we can rewrite the last expression as

$$\prod_{i=1}^{n} \exp\left(-\left(x_i - x_{i-1}\right)^2 / 2(t_i - t_{i-1})\right) \cdot \prod_{i=1}^{n} \exp\left(\mu\left(x_i - x_{i-1}\right) - \frac{1}{2}\mu^2\left(t_i - t_{i-1}\right)\right)$$

$$= \prod_{i=1}^{n} \exp\left(-\left(x_i - x_{i-1}\right)^2 / 2\left(t_i - t_{i-1}\right)\right) \cdot \exp\left(\mu x_n - \frac{1}{2}\mu^2 t_n\right)$$

Comparing the last product to the density of the vector $\left(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}\right)$, our original expectation of the function f of the process $X_t = B_t + \mu t$ can be expressed as a modified expectation of another function of Brownian motion, as follows:

$$\mathbb{E}\left[f\left(X_{t_1}, X_{t_2}, \dots, X_{t_n}\right)\right] = \mathbb{E}\left[f\left(B_{t_1}, B_{t_2}, \dots, B_{t_n}\right) \underbrace{\exp\left(\mu B_{t_n} - \frac{1}{2}\mu^2 t_n\right)}_{Z_{t_n}}\right].$$

Since Z_{t_n} is a martingale and $f\left(X_{t_1}, \dots, X_{t_n}\right)$ is $\mathcal{F}_{t_n}$-measurable, we can rewrite this, when $0 \leq t_1 \leq \dots \leq t_n \leq T$, as

$$\mathbb{E}\left[f\left(X_{t_1}, \dots, X_{t_n}\right)\right] = \mathbb{E}\left[f\left(B_{t_1}, \dots, B_{t_n}\right) Z_T\right].$$

Compare this with Girsanov's theorem (see Remark 2.9 and equation (2.17)). In fact, due to the continuity of Brownian paths, using approximation arguments, note that these finite-dimensional calculations already imply that

$$\mathbb{E}\left[\mathbb{I}_{\left\{\max_{t\in[0,T]}(B_t+\mu t)<x\right\}}\right] = \mathbb{E}\left[\mathbb{I}_{\left\{\max\limits_{t\in[0,T]} B_t<x\right\}} \exp\left(\mu B_T - \frac{1}{2}\mu^2 T\right)\right].$$

B. Hitting time of Brownian motion with drift. Let $\Omega \doteq \mathcal{C}[0,\infty)$, $\mathcal{F}_t^0 \doteq \sigma(\omega_s, s \le t)$, let $\mathbb{P}$ be Wiener measure and let $\tilde{\mathbb{P}}$ be the distribution induced by Brownian motion with drift. For $a > 0$, let

$$\tau_a \doteq \inf\{t > 0 : w_t = a\}.$$

Under $\tilde{\mathbb{P}}$, τ_a is the hitting time of Brownian motion with drift μ. By Girsanov's theorem, we have for $t \in [0,\infty)$,

$$\tilde{\mathbb{P}}(\tau_a \le t) = \tilde{\mathbb{E}}\left[\mathbb{I}_{\{\tau_a \le t\}}\right] = \mathbb{E}\left[\mathbb{I}_{\{\tau_a\le t\}} Z_t\right].$$

Conditioning on the σ-field $\mathcal{F}^0_{t\wedge\tau_a}$, using the fact that $\{\tau_a \le t\}$ is $\mathcal{F}^0_{t\wedge\tau_a}$-measurable and Z is an $\mathcal{F}_t^0$-martingale, we find that

$$\begin{aligned}\mathbb{E}\left[\mathbb{I}_{\{\tau_a\le t\}} Z_t\right] = \mathbb{E}\left[\mathbb{I}_{\{\tau_a\le t\}} Z_{t\wedge\tau_a}\right] &= \mathbb{E}\left[\mathbb{I}_{\{\tau_a\le t\}} \exp\left(\mu a - \frac{1}{2}\mu^2\tau_a\right)\right] \\ &= \int_0^t \exp\left(\mu a - \frac{1}{2}\mu^2 s\right)\frac{a}{s^{3/2}}\varphi\left(\frac{a}{\sqrt{s}}\right) ds,\end{aligned}$$

where φ is the standard normal density and we have used the known explicit expression for the density of the first hitting time for Brownian motion (see, for example, relation (8.5) on page 96 of [1]). Differentiating the last integral and completing the square in the exponent, we see that the density f_{τ_a} of τ_a is given by

$$f_{\tau_a}(t) = \frac{a}{t^{3/2}}\varphi\left(\frac{a-\mu t}{\sqrt{t}}\right).$$

Another very important application of Girsanov's theorem is in establishing existence and uniqueness of so-called weak solutions to stochastic differential equations. This will be covered in the next section.