1. **Exercise 4.1.6** Provide an example of a probability space \((\Omega, \mathcal{F}, P)\), a filtration \(\{\mathcal{F}_n\}\) and a stochastic process \(\{X_n\}\) adapted to \(\mathcal{F}_n\) such that:

(a) \(\{X_n\}\) is a martingale with respect to its canonical filtration but \((X_n, \mathcal{F}_n)\) is not a martingale.

**ANS:** Take \(\Omega = \{a, b\}\), \(\mathcal{F}_0 = \mathcal{F} = 2^\Omega\), \(X_0 = 0\), \(X_1 = \pm 1\) with probability \(1/2\) and \(X_n = X_1\) for all \(n \geq 2\). Then \(\{X_n\}\) is a martingale with respect to its canonical filtration since:

\[
X_0 = 0 = \mathbb{E}(X_1) = \mathbb{E}(X_1 | X_0)
\]

and

\[
X_n = \mathbb{E}(X_n | X_n) = \mathbb{E}(X_{n+1} | X_n) = \mathbb{E}(X_{n+1} | \sigma(X_0, \ldots, X_n))
\]

for all \(n \geq 1\). Now consider the filtration \(\{\mathcal{F}_n\}\) where \(\mathcal{F}_n = 2^\Omega\) for all \(n\). Then,

\[
X_0 = 0 \neq X_1 = \mathbb{E}(X_1 | \mathcal{F}_0),
\]

so that \((X_n, \mathcal{F}_n)\) is not a martingale.

(b) Provide a probability measure \(Q\) on \((\Omega, \mathcal{F})\) under which \(\{X_n\}\) is not a martingale even with respect to its canonical filtration.

**ANS:** Let \(Q\) be a probability measure on \((\Omega, \mathcal{F})\) such that \(X_1 = 1\) with probability \(p > 1/2\) and \(X_1 = -1\) with probability \(1 - p < 1/2\). Then

\[
\mathbb{E}X_1 = (2p - 1) > 0 \neq 0 = X_0
\]

so that \(\{X_n\}\) is not a martingale with respect to its canonical filtration.

2. **Exercise 4.1.23** Let \(\xi_1, \xi_2, \ldots\) be independent with \(\mathbb{E}\xi_i = 0\) and \(\mathbb{E}\xi_i^2 = \sigma_i^2\).

(a) Let \(S_n = \sum_{i=1}^n \xi_i\) and \(s_n^2 = \sum_{i=1}^n \sigma_i^2\). Show that \(\{S_n^2\}\) is a sub-martingale and \(\{S_n^2 - s_n^2\}\) is a martingale.

**ANS:** Using the same argument of Example 4.1.8 we know that \(\{S_n\}\) is a martingale with respect to its canonical filtration. Moreover, from the fact that \(S_n^2 = \sum_{i=1}^n \xi_i^2 + 2 \sum_{1 \leq i < j \leq n} \xi_i \xi_j\) it is clear that \(|\mathbb{E}|S_n^2| < \infty\) for all \(n\). Thus since \(x \mapsto x^2\) is a convex function it follows from the conditional Jensen inequality that \(S_n^2\) is a sub-martingale. Letting \(\mathcal{F}_n = \sigma(S_1, \ldots, S_n)\) and using that \(\xi_{n+1}\) is independent of \(\mathcal{F}_n\), we have

\[
\mathbb{E}|S_{n+1}^2|_{\mathcal{F}_n} = \mathbb{E}|(S_n + \xi_{n+1})^2|_{\mathcal{F}_n} = \mathbb{E}|S_n^2 + 2\xi_{n+1} S_n + \xi_{n+1}^2|_{\mathcal{F}_n} = S_n^2 + \sigma_{n+1}^2.
\]
Hence,

\[ \mathbb{E}[S_{n+1}^2 - s_{n+1}^2 | \mathcal{F}_n] = S_n^2 - s_{n+1}^2 + \sigma_{n+1}^2 - s_n^2. \]

Thus \( \{ S_n^2 - s_n^2 \} \) is a martingale as desired.

(b) Suppose also that \( m_n = \prod_{i=1}^n \mathbb{E}[e^{\xi_i}] < \infty \). Show that \( \{ e^{S_n} \} \) is a sub-martingale and \( M_n = e^{S_n}/m_n \) is a martingale.

**ANS:** By assumption \( m_n < \infty \) giving us that \( \{ e^{S_n} \} \) is an integrable SP. Since \( \{ S_n \} \) is a martingale and \( x \mapsto e^x \) is convex it follows from the conditional Jensen inequality that \( \{ e^{S_n} \} \) is a sub-martingale. Letting \( \mathcal{F}_n = \sigma(M_1, \ldots, M_n) \), the independence of \( \xi_{n+1} \) and \( \mathcal{F}_n \) gives us that

\[ \mathbb{E}[M_{n+1}|\mathcal{F}_n] = \frac{1}{m_{n+1}} \mathbb{E}[e^{S_n} e^{\xi_{n+1}}|\mathcal{F}_n] = \frac{e^{S_n}}{m_{n+1}} \mathbb{E}[e^{\xi_{n+1}}] = \frac{e^{S_n}}{m_n} = M_n. \]

Therefore \( \{ M_n \} \) is a martingale.

3. **Exercise 4.2.5**. Let \( \mathcal{G}_t \) denote the canonical filtration of a Brownian motion \( W_t \).

(a) Show that for any \( \lambda \in \mathbb{R} \), the S.P. \( M_t(\lambda) = \exp(\lambda W_t - \lambda^2 t/2) \), is a continuous time martingale with respect to \( \mathcal{G}_t \).

**ANS:** Note that \( \mathbb{E}[M_t(\lambda)] = e^{-\lambda^2(t/2)} \mathbb{E}[e^{\lambda W_t}] \) which since \( W_t \) is a Gaussian random variable, we know to be finite. Further, \( \mathbb{E}e^{\lambda(W_{t+h} - W_t)} = e^{\lambda^2 h/2} \) yielding the identity \( \mathbb{E}[M_{t+h}(\lambda)|\mathcal{G}_t] = e^{-\lambda^2(t/2) + \lambda W_t} = M_t(\lambda) \), so \( M_t(\lambda) \) is a martingale.

(b) Explain why \( \frac{\partial}{\partial \lambda} M_t(\lambda) \) are also martingales with respect to \( \mathcal{G}_t \).

**ANS:** Fixing \( \lambda \in \mathbb{R} \), let \( \lambda_m = \lambda + m^{-1} \) and recall that \( M_t(\lambda_m) \) is a MG with respect to \( \mathcal{G}_t \). The adapted process \( X_t(m, \lambda) := m(M_t(\lambda_m) - M_t(\lambda)) \) is then also a MG with respect to \( \mathcal{G}_t \). That is, \( \mathbb{E}[X_{t+h}(m, \lambda) - X_t(m, \lambda)|\mathcal{G}_t] = 0 \) for any non-random \( h, t \geq 0 \). Further, considering \( m \to \infty \) we get by definition of the derivative that \( X_t(m, \lambda) \) converges a.s. to the adapted S.P. \( Z_t(\lambda) := \frac{\partial}{\partial \lambda} M_t(\lambda) \).

Now, by the mean value theorem

\[ \sup_m \{|X_t(m, \lambda)|\} \leq Y_t := \sup\{|Z_t(\lambda + u)| : 0 \leq u \leq 1\}. \]

Computing explicitly \( Z_t(\lambda + u) \), it is not hard to check that \( Y_t \leq (|W_t| + (|\lambda| + 1)t)e^{(|\lambda|+1)|W_t|} \) is integrable (per fixed \( t \geq 0 \)). From the preceding we thus deduce by dominated convergence for C.E. that a.s. \( \mathbb{E}[Z_{t+h}(\lambda) - Z_t(\lambda)|\mathcal{G}_t] = 0 \). Consequently, given \( \lambda \in \mathbb{R} \) non-random, the process \( Z_t(\lambda) \) is a MG with respect to \( \mathcal{G}_t \). Applying the same reasoning with \( M_t(\lambda) \) replaced by \( Z_t(\lambda) \) extends our claim from \( k = 1 \) to \( k = 2 \), and arguing inductively in \( k \), the same applies for \( k = 3, 4, \ldots \).

(c) Compute the first three derivatives in \( \lambda \) of \( M_t(\lambda) \) at \( \lambda = 0 \) and deduce that the S.P. \( W_t^2 - t \) and \( W_t^4 - 3tW_t \) are also MGs.
ANS: Fixing \( x, t \in \mathbb{R} \), the derivative of \( M(\lambda) := e^{\lambda x - \lambda^2 t/2} \) is \( M'(\lambda) = (x - \lambda t)M(\lambda) \), yielding that \( M''(\lambda) = [(x - \lambda t)^2 - t]M(\lambda) \) and \( M'''(\lambda) = (x - \lambda t)[(x - \lambda t)^2 - 3t]M(\lambda) \). In case \( \lambda = 0 \) we have \( M(0) = 1 \) resulting with \( M'(0) = x \), \( M''(0) = x^2 - t \) and \( M'''(0) = x^3 - 3tx \). Setting \( x = W_t \) we deduce by the preceding that \( W_t^2 - t \) and \( W_t^3 - 3tW_t \) are also MGs.

4. Exercise 4.2.10 Given a positive MG \((Z_t, \mathcal{F}_t)\) with \( \mathbb{E}Z_0 = 1 \) consider for each \( t \geq 0 \) the probability measure \( \tilde{\mathbb{P}}_t : \mathcal{F}_t \to \mathbb{R} \) given by \( \tilde{\mathbb{P}}_t(A) = \mathbb{E}[Z_t I_A] \).

(a) Show that \( \tilde{\mathbb{P}}_t(A) = \tilde{\mathbb{P}}_s(A) \) for any \( A \in \mathcal{F}_s \) and \( 0 \leq s \leq t \).

ANS: Since \( Z_t \) is a martingale and \( I_A \) is \( \mathcal{F}_s \)-measurable, we have by the tower property and taking out what is known that
\[
\tilde{\mathbb{P}}_t(A) = \mathbb{E}[Z_t I_A] = \mathbb{E} [\mathbb{E}[Z_t I_A | \mathcal{F}_s]] = \mathbb{E}[\mathbb{E}[Z_t | \mathcal{F}_s] I_A] = \mathbb{E}[Z_t I_A] = \tilde{\mathbb{P}}_s(A).
\]

(b) Fixing \( 0 \leq u \leq s \leq t \) and \( Y \in L^1(\Omega, \mathcal{F}_s, \tilde{\mathbb{P}}_t) \), set \( X_{s,u} = \mathbb{E}(YZ_u | \mathcal{F}_u)/Z_u \). With \( \tilde{\mathbb{E}}_t \) denoting the expectation under \( \tilde{\mathbb{P}}_t \), deduce that \( \tilde{\mathbb{E}}_t(Y | \mathcal{F}_u) = X_{s,u} \) almost surely under \( \tilde{\mathbb{P}}_t \) (hence also under \( \mathbb{P} \), by Exercise 1.4.32).

ANS: First note that \( YZ_s \in L^1(\Omega, \mathcal{F}_s, \mathbb{P}) \) since
\[
\mathbb{E}(|Y| | Z_s) = \mathbb{E}(|Y| \mathbb{E}(Z_t | \mathcal{F}_s)) = \mathbb{E}(\mathbb{E}(|Y| Z_t | \mathcal{F}_s)) = \mathbb{E}(|Y| Z_t) = \tilde{\mathbb{E}}_t(|Y|) < \infty.
\]
Consequently, the \( \mathcal{F}_u \) measurable random variable \( X_{s,u} = \mathbb{E}(YZ_u | \mathcal{F}_u)/Z_u \) is well defined. Further, fixing \( A \in \mathcal{F}_u \), recall that \( YI_A \) is \( \mathcal{F}_s \) measurable and \((Z_t, \mathcal{F}_t)\) a martingale. Hence, using the tower property, taking out what is known and applying part (a) for the \( \mathcal{F}_u \) measurable \( X_{s,u} I_A \) we get that
\[
\tilde{\mathbb{E}}_t[YI_A] = \mathbb{E}[Z_t YI_A] = \mathbb{E}[\mathbb{E}(Z_t YI_A | \mathcal{F}_s)] = \mathbb{E}[YI_A \mathbb{E}(Z_t | \mathcal{F}_s)] = \mathbb{E}[YZ_s I_A]
\]
\[
= \mathbb{E}[\mathbb{E}(YZ_s I_A | \mathcal{F}_u)] = \mathbb{E}[\mathbb{E}(YZ_s | \mathcal{F}_u) I_A] = \mathbb{E}[Z_u X_{s,u} I_A] = \tilde{\mathbb{E}}_u[X_{s,u} I_A] = \tilde{\mathbb{E}}_t[X_{s,u} I_A].
\]

Since this applies for any \( A \in \mathcal{F}_u \), we have by definition of the conditional expectation in the probability space \((\Omega, \mathcal{F}, \tilde{\mathbb{P}}_t)\) that \( X_{s,u} = \tilde{\mathbb{E}}_t(Y | \mathcal{F}_u) \) up to a set \( N \in \mathcal{F} \) such that \( \tilde{\mathbb{P}}_t(N) = 0 \). Recall Exercise 1.4.32 that \( \tilde{\mathbb{P}}_t(N) = 0 \) if and only if \( \mathbb{P}(N) = 0 \), so the identity \( X_{s,u} = \tilde{\mathbb{E}}_t(Y | \mathcal{F}_u) \) holds for \( \mathbb{P} \) almost every \( \omega \), as claimed.

5. Exercise 5.1.8 Compute the mean and the auto-covariance functions of the processes \( B_t, Y_t, U_t, \) and \( X_t \).
ANS: We compute,

\[\mathbf{E}(B_t) = 0,\]
\[\mathbf{E}(B_tB_s) = s(1 - t) \text{ when } 0 \leq s \leq t \leq 1,\]
\[\mathbf{E}(B_tB_s) = s - 1 \text{ when } 0 \leq 1 \leq s \leq t \text{ and} \]
\[\mathbf{E}(B_tB_s) = 0 \text{ when } 0 \leq s \leq 1 \leq t;\]
\[\mathbf{E}(Y_t) = e^{t/2} \text{ and } \mathbf{E}[(Y_t - e^{t/2})(Y_s - e^{s/2})] = e^{(t+s)/2}(e^{\min(t,s)} - 1);\]
\[\mathbf{E}(U_t) = 0 \text{ and } \mathbf{E}(U_tU_s) = e^{-|t-s|/2}.\]

\[\mathbf{E}(X_t) = x + \mu t \text{ and},\]
\[\mathbf{E}[(X_t - x - \mu t)(X_s - x - \mu s)] = \sigma^2 \mathbf{E}(W_tW_s) = \sigma^2 \min(t,s).\]

Justify your answers to:

(a) Which of the processes \(W_t, B_t, Y_t, U_t, X_t\) is Gaussian?

**ANS:** We know that \(W_t\) is a Gaussian process. The f.d.d. of the S.P. \(B_t\) and \(U_t\) correspond to deterministic linear combinations of the joint Gaussian r.v. \(W_t\), hence both \(B_t\) and \(U_t\) are Gaussian processes. Since \(Y_t = e^{W_t}\) is strictly positive and not almost surely a constant, it can not be a Gaussian r.v, hence \(Y_t\) is not a Gaussian process. Finally, \(X_t\) is just an affine (time-dependent) translate of a Gaussian process and hence Gaussian.

(b) Which of these processes is stationary?

**ANS:** Stationarity implies the process has constant mean and its auto-covariance \(\rho(t, s)\) is a function only of \(|t - s|\). The S.P. \(W_t, B_t, Y_t\) and \(X_t\) fail to have this property so are non-stationary. The S.P. \(U_t\) satisfies these conditions and being also Gaussian, this suffices for \(U_t\) being a stationary process.

(c) Which of these processes has continuous sample paths?

**ANS:** \(W_t\) has continuous sample paths by the definition of Brownian motion so \(B_t, Y_t, U_t, X_t\) are finite compositions of functions continuous in \(t\). Therefore, all five processes have continuous sample paths.

(d) Which of these processes is adapted to the filtration \(\sigma(W_s, s \leq t)\) and which is also a sub-martingale for this filtration?

**ANS:** Recall that \(W_t\) is adapted and is a martingale for its canonical filtration. The processes \(B_t\) and \(U_t\) depend on values of \(W_s\) for \(s > t\) so they are not adapted to this filtration. The S.P. \(Y_t\) is the composition of the convex function \(e^x\) and a martingale and hence a submartingale. Finally, as \(X_t\) is an affine translate of \(W_t\), it is visibly adapted to the filtration and is a submartingale provided
that \( \mu \geq 0: \)
\[
E[X_t|\sigma(W_s : s \leq t)] = x + \mu t + \sigma W_s \geq x + \mu s + \sigma W_s = X_s.
\]

Note that if \( \mu < 0 \) we get the reverse inequality.

6. **Exercise 5.1.11.** Suppose \( W_t \) is a Brownian motion.

(a) Compute the probability density function of the random vector \((W_s, W_t)\). Then compute \(E(W_s|W_t)\) and \(\text{Var}(W_s|W_t)\), first for \( s > t \), then for \( s < t \).

*Hint:* Consider Example 2.4.5.

*ANS:* Suppose first that \( t < s \). Then, \( W_s - W_t \) is independent of \( W_t \), having a Gaussian distribution of zero mean and variance \( s - t \). Therefore, \( E(W_s|W_t) = W_t \) and \( \text{Var}(W_s|W_t) = E((W_s - W_t)^2|W_t) = s - t \). Moving to deal with \( t > s \), note that \((W_s, W_t)\) is a Gaussian random vector, of zero mean and covariance matrix \( \Sigma \) whose entries are \( \Sigma_{11} = \Sigma_{12} = \Sigma_{21} = s, \Sigma_{22} = t \). Upon finding that \( \Sigma \) is invertible and computing its inverse, we get that \((W_s, W_t)\) has the (joint) probability density function \( f_{W_s,W_t}(x,y) = \exp(-x^2/(2s) - (y-x)^2/(2(t-s)))/(2\pi\sqrt{s(t-s)}) \). With the density of \( W_t \) being \( g_{W_t}(y) = \exp(-y^2/2t)/\sqrt{2\pi t} \), we have by Example 2.4.5 that the conditional density of \( W_s \) given \( W_t \) is \( f_{W_s|W_t}(x|W_t) \) for
\[
f_{W_s|W_t}(x|y) = f_{W_s,W_t}(x,y)/g_{W_t}(y) = \exp(-(x - sy/t)^2/(2\sigma^2))/\sqrt{2\pi\sigma}
\]
where \( \sigma^2 = s(t-s)/t \). The latter is the density of a Gaussian random variable of mean \( sy/t \) and variance \( \sigma^2 \), so as explained in Example 2.4.5 we have that \( E(W_s|W_t) = (s/t)W_t \) and \( \text{Var}(W_s|W_t) = s - s^2/t \).

(b) Explain why the Brownian Bridge \( B_t, 0 \leq t \leq 1 \) has the same distribution as \( \{W_t, 0 \leq t \leq 1, \text{conditioned upon } W_1 = 0\} \) (which is the reason for naming \( B_t \) a Brownian bridge).

*Hint:* Both Exercise 2.4.6 and parts of Exercise 5.1.8 may help here.

*ANS:* For \( s \leq t \leq 1 \) we know that \( X = W_1 - W_t \) is independent of the random vector \((Y, Z) = (W_s, W_t)\). Consequently, combining part (a) with Exercise 2.4.6 we have that \( E(W_s|W_1 - W_t) = E(W_s|W_t) = (s/t)W_t \). Further, \( \sigma(W_t, W_1) = \sigma(W_t, W_1 - W_t) \), so also \( E(W_s|W_t, W_1) = (s/t)W_t \). Thus, applying the tower property for \( \sigma(W_1) \subseteq \sigma(W_t, W_1) \) and taking out what is known, we see that
\[
E[W_sW_t|W_1] = E[W_tE(W_s|W_t, W_1)|W_1] = (s/t)E(W_t^2|W_1).
\]

Recall that by part (a), \( E(W_t^2|W_1) = \text{Var}(W_t|W_1) + [E(W_t|W_1)]^2 = t - t^2 + t^2W_1^2 \),
implying that

$$E(W_sW_t|W_1) = s(1-t) + stW_1^2.$$ 

Though we shall not do so in detail, fixing $0 < s_1 < \ldots < s_n < 1$ one can compute the density of $(W_{s_1}, \ldots, W_{s_n})$ conditional on $\{W_1 = 0\}$, per Example 2.4.5, and verify that it is the density of a (zero-mean) non-degenerate Gaussian random vector. Consequently, $\{W_t, 0 \leq t \leq 1\}$ conditional on the event $\{W_1 = 0\}$ is a Gaussian S.P. Recall Exercise 5.1.8, that $E(B_t) = 0$ and $E(B_sB_t) = s(1-t)$ for all $0 \leq s \leq t \leq 1$. In conclusion, we have established that the Gaussian S.P. $\{W_t, 0 \leq t \leq 1\}$ conditional on the event $\{W_1 = 0\}$, has the same mean and auto-covariance functions as the Gaussian S.P. $B_t$. Therefore, these two S.P. have the same distribution (i.e. the same f.d.d.).