1. (Durrett 4.1.3) Let $x \in \mathbb{R}$, then on all of Ω, we get that

$$(E(X|G) + xE(Y|G))^2 = E(X|G)^2 + x^2E(Y|G)^2 + 2xE(XY|G) \geq 0$$

So this quadratic has discriminant ≤ 0 on all of Ω, writing down the discriminant this gives us

$$(2E(XY|G))^2 - 4E(X|G)^2E(Y|G)^2 \leq 0$$

and dividing by 4 gives us

$$E(XY|G)^2 \leq E(X|G)^2E(Y|G)^2$$

Finally Jensen's inequality for conditional expectation implies that $E(X|G)^2 \leq E(X^2|G)$ and $E(Y|G)^2 \leq E(Y^2|G)$, and since both are positive we get the following inequality

$$E(XY|G)^2 \leq E(X|G)^2E(Y|G)^2 \leq E(X^2|G)E(Y^2|G)$$

□

2. (Durrett 4.1.7) Since $\mathcal{F} = \{\Omega, \emptyset\}$, $E(Z|\mathcal{F})$ must be a constant function for any random variable Z, since any non-constant function is not $\mathcal{F}$ measurable. Because of this,

$$\text{Var}(X|\mathcal{F}) = P(\Omega)\text{Var}(X|\mathcal{F}) = \int \text{Var}(X|\mathcal{F})$$

Now we can just compute using the property of integrating conditional expectation.

$$\begin{aligned}
E(\text{Var}(X|\mathcal{F})) + \text{Var}(E(X|\mathcal{F})) &= \int \text{Var}(X|\mathcal{F}) + \int E(X|\mathcal{F})^2 - \left(\int E(X|\mathcal{F})\right)^2 \\
&= \int E(X^2|\mathcal{F}) - E(X|\mathcal{F})^2 + \int E(X|\mathcal{F})^2 - \left(\int X\right)^2 \\
&= \int E(X^2|\mathcal{F}) - (EX)^2 \\
&= \int X^2 - (EX)^2 = EX^2 - (EX)^2 = \text{Var}X
\end{aligned}$$

□

3. (Durrett 4.1.9) The main step is a computation, note since $X \in \mathcal{G}$, so is X^2 this gives us

$$E((X-Y)^2|\mathcal{G}) = E(X^2|\mathcal{G}) + E(Y^2|\mathcal{G}) - 2E(XY|\mathcal{G}) = X^2 + E(Y^2|\mathcal{G}) - 2XE(Y|\mathcal{G}) = E(Y^2|\mathcal{G}) - X^2$$

Now we can apply the assumption $EY^2 = EX^2$, by taking expectations to get

$$E[(X-Y)^2] = E[E((X-Y)^2|\mathcal{G})] = E[E(Y^2|\mathcal{G})] - EX^2 = EY^2 - EX^2 = 0$$

Assume for contradiction that $X \neq Y$ a.s. then $A = \left\{(X-Y)^2 > 0\right\}$ has $P(A) > 0$, since $A = \bigcup_1^\infty A_n$ where $A_n = \left\{(X-Y)^2 \geq \frac{1}{n}\right\}$, we get that $P(A_n) \to P(A)$ by continuity from below, hence for some n we have $P(A_n) > 0$, so that

$$0 = E[(X-Y)^2] \geq E[(X-Y)^2 1_{A_n}] \geq \int_{A_n} \frac{1}{n} = P(A_n)/n > 0$$

which is the desired contradiction. □

4. (Durrett 4.1.10) Let Z be a random variable on $L^1(\mathcal{F})$, and let $X = E(Z|\mathcal{G})$, further assume that $X \overset{\text{dist}}{=} Z$, this of course implies that $P(X > x) = P(Z > x)$ for all $x \in \mathbb{R}$. Now we can apply layer cake to get that

$$EZ^+ = \int_0^\infty P(Z > x)dx = \int_0^\infty P(X > x)dx = EX^+$$

Then the property of conditional probability further implies that $E[Z1_{X>0}] = E[X1_{X>0}] = EZ^+$, which implies that $EZ^+ = E[Z^+1_{X>0}]$, so that $Z > 0$ implies $X > 0$ a.s. Moreover, there does not exist a set B of positive measure with $Z < 0$ on B, and $B \subset X > 0$, else we would have

$$EZ^+ = E[Z1_{X>0}] = E[Z1_B] + E[Z1_{\{X>0\}\setminus B}] < E[Z1_{\{X>0\}\setminus B}] \leq E[Z^+1_{\{X>0\}\setminus B}] \leq EZ^+$$

which is a contradiction, hence $Z < 0$ implies that $X \leq 0$ a.s. Finally $P(Z = 0) = P(X = 0)$, and $\int_{X=0} X = \int_{X=0} Z$ alongside $Z > 0$ implying $X > 0$ a.s. gives us that $\{X = 0\} = \{Z = 0\}$ a.s.

Now for any $c \in \mathbb{R}$ we can choose $Z = Y - c$, and since $Y \overset{\text{dist}}{=} E[Y|\mathcal{G}]$, we have $Y - c \overset{\text{dist}}{=} E[Y - c|\mathcal{G}] = E[Y|G] - c$, applying the above gives us that $Y > c$ if and only if $X := E[Y|G]$ is. It follows that for $q \in \mathbb{Q}$ and $n \in \mathbb{Z}_{>0}$ any set of the form $\left\{q - \frac{1}{2n} < X \leq q + \frac{1}{2n} \text{ and } Y \notin (q - \frac{1}{2n}, q + \frac{1}{2n}]\right\}$ has probability zero, we can use this to get a contradiction. Suppose that $\{X \neq Y\}$ has positive probability, then $\{X \neq Y\} = \bigcup_1^\infty \left\{|X - Y| > \frac{1}{n}\right\}$, so that for some n, one of these sets must have positive probability. Then

$$\left\{|X - Y| > \frac{1}{n}\right\} \subset \bigcup_{q\in\mathbb{Q}} \left\{q - \frac{1}{2n} < X \leq q + \frac{1}{2n} \text{ and } Y \notin (q - \frac{1}{2n}, q + \frac{1}{2n}]\right\}$$

by countable subadditivity one of the sets in the union must have positive probability, but this is a contradiction. So we can conclude indeed $P(X \neq Y) = 0$. $\square$

5. (Durrett 4.2.1) Since $X_1, \ldots, X_n \in \mathcal{G}_n$, so are the σ-algebra they generate, i.e. $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) \subset \mathcal{G}_n$, the properties that $X_n \in \mathcal{F}_n$, and $E|X_n| < \infty$ are immediate by definition of $\mathcal{F}$ and $(X_n, \mathcal{G}_n)$ being a martingale. To check the Martingale property we use that $E(E(X_{n+1}|\mathcal{G}_n)|\mathcal{F}_n) = E(X_{n+1}|\mathcal{F}_n)$, this allows us to write

$$E(X_{n+1}|\mathcal{F}_n) = E(E(X_{n+1}|\mathcal{G}_n)|\mathcal{F}_n) = E(X_n|\mathcal{F}_n) = X_n$$

$\square$

6. (Durrett 4.2.3) First note that for measurable X, Y we can write

$$\max\{X, Y\} = \frac{1}{2}(X + Y) + \frac{1}{2}|X - Y| = \frac{1}{2}(X + Y) + \frac{1}{2}(X - Y)^+ + \frac{1}{2}(Y - X)^+$$

Applying this to $\max\{X_{n+1}, Y_{n+1}\}$ and using linearity of conditional expectation yields

$$2E(\max\{X_{n+1}, Y_{n+1}\}|\mathcal{F}_n) = E(X_{n+1}|\mathcal{F}_n) + E(Y_{n+1}|\mathcal{F}_n) + E((X_{n+1} - Y_{n+1})^+|\mathcal{F}_n) + E((Y_{n+1} - X_{n+1})^+|\mathcal{F}_n)$$

Since $X \mapsto X^+$ is convex, we can apply the submartingale property, along with Jensen's to get

$$2E(\max\{X_{n+1}, Y_{n+1}\}|\mathcal{F}_n) \geq X_n + Y_n + E((X_{n+1} - Y_{n+1})|\mathcal{F}_n)^+ + E((Y_{n+1} - X_{n+1})|\mathcal{F}_n)^+$$

Now using the fact that $X \mapsto X^+$ is increasing, alongside the submartingale property gives us

$$E((X_{n+1} - Y_{n+1})|\mathcal{F}_n)^+ \geq (X_n - Y_n)^+ \text{ and } E((Y_{n+1} - X_{n+1})|\mathcal{F}_n)^+ \geq (Y_n - X_n)^+$$

We can apply these inequalities to get

$$2E(\max\{X_{n+1}, Y_{n+1}\}|\mathcal{F}_n) \geq X_n + Y_n + (X_n - Y_n)^+ + (Y_n - X_n)^+ = 2\max\{X_n, Y_n\}$$

Dividing out by 2, we get the submartingale property. Now note that since $X_n, Y_n \in \mathcal{F}_n$, and max is borel measurable we have $\max\{X_n, Y_n\} \in \mathcal{F}_n$ and $\max X_n, Y_n \leq |X_n| + |Y_n| \in L^1(\mathcal{F}_n)$, since L^1 is a vectorspace. $\square$

7. (Durrett 4.2.6)

(a) A non-negative Martingale is in particular a nonnegative super-Martingale so by the super-martingale convergence theorem, there is some random variable X_∞, such that $X_n \to X_\infty$. Assume for contradiction that $P(A) > 0$, where $A = \{X_\infty > 0\}$. Then on A, we get that

$$\lim_{n\to\infty} Y_n 1_A = \lim_{n\to\infty} \frac{X_n}{X_{n-1}} 1_A = 1_A \frac{X_\infty}{X_\infty} = 1_A$$

Since $X_\infty \neq 0$ on A. Hence we get that $Y_n \to 1$ on A. Note that by continuity from above, we get that $\lim_{k\to\infty} P(|Y_1 - 1| < \frac{1}{k}) = P(Y_1 = 1) < 1$, the pointwise convergence on A tells us for any $k \in \mathbb{Z}_{>0}$, there exists some $N \in \mathbb{Z}_{>0}$ such that for $n \geq N$, we get $|Y_n - 1| < \frac{1}{k}$ on A. Letting k be arbitrary, and choosing such an N, we get for any $n \geq N$

$$P(A) \leq P\left(|Y_N - 1| < \frac{1}{k}, \dots, |Y_n - 1| < \frac{1}{k}\right) = P\left(|Y_1 - 1| < \frac{1}{k}\right)^{n-N}$$

So that for any $m \in \mathbb{Z}_{>0}$, we have $P(A) \leq P\left(|Y_1 - 1| < \frac{1}{k}\right)^m$, now since k was arbitrary, and A, m are independent of k, this holds for all k, so taking the limit

$$P(A) \leq \lim_{k\to\infty} P\left(|Y_1 - 1| < \frac{1}{k}\right)^m = P(Y_1 = 1)^m$$

And since this holds for any m, we get that

$$P(A) \leq \lim_{m\to\infty} P(Y_1 = 1)^m = 0$$

which is a contradiction. □

(b) $\frac{1}{n} \log X_n = \frac{1}{n} \sum_1^n \log Y_j$, by the strong law of large numbers this converges to $E \log Y_1$, so it suffices to check that $E \log Y_1 < 0$. We can rewrite $E \log Y_1 = E[1_{Y_1 > 1} \log Y_1] + E[1_{Y_1 < 1} \log Y_1]$, now since $Y_1 \neq 1$ we get $0 < \int_{Y_1 > 1} Y_1 + \int_{Y_1 < 1} Y_1 \leq EY_1 = 1$, and moreover since the expectation is 1, if one of the two integrals is nonzero, then so is the other so that

$$0 < \int_{Y_1 > 1} Y_1, \int_{Y_1 < 1} Y_1 < 1$$

This means that we can apply the inequality (strict because both values are strictly between $0, 1$)

$$\begin{aligned} 0 = \log(EY_1) &\geq \log\left(\int_{Y_1>1} Y_1 + \int_{Y_1<1} Y_1\right) > \log\left(\int_{Y_1>1} Y_1\right) + \log\left(\int_{Y_1<1} Y_1\right) \\ &\geq \int_{Y_1>1} \log Y_1 + \int_{Y_1<1} \log Y_1 = E[1_{Y_1>1} \log Y_1] + E[1_{Y_1<1} \log Y_1] = E \log Y_1 \end{aligned}$$

To justify the inequality interchanging the log with the integral, we use Jensen's inequality with exponentiation and the fact log is increasing (i.e. suppose X is a random variable positive on A), then

$$\exp\left(\int_A \log X\right) \leq \int_A \exp \log X = \int_A X$$

so since log is increasing, taking log of both sides preserves the inequality

$$\int_A \log X = \log \exp\left(\int_A \log X\right) \leq \log \int_A X$$

□

8. (Durrett 4.4.2) Since $N \geq M$, we have $1_{M<n} - 1_{N<n} \geq 0$, so we can take $H_n = 1_{M<n} - 1_{N<n}$ which is positive as above, obviously bounded and also predictable since $\{M < n\} = \{M \leq n-1\} \in \mathcal{F}_{n-1}$ and $\{N < n\} = \{N \leq n-1\} \in \mathcal{F}_{n-1}$, hence $1_{M<n} \in \mathcal{F}_{n-1}$ and $1_{N<n} \in \mathcal{F}_{n-1}$, which implies their difference is also in $\mathcal{F}_{n-1}$. Since H constitutes a positive martingale transform and X_n is a martingale, so is $(H \bullet X)_n$, so that (on the set $N \leq k$ which is a.s.)

$$\begin{aligned}
(H \bullet X)_k &= \sum_1^k (1_{M<n} - 1_{N<n})(X_n - X_{n-1}) = \sum_1^k 1_{M<n}(X_n - X_{n-1}) - \sum_1^k 1_{N<n}(X_n - X_{n-1}) \\
&= \sum_1^k (1 - 1_{M \geq n})(X_n - X_{n-1}) + \sum_1^k (1_{N \geq n} - 1)(X_n - X_{n-1}) \\
&= X_k - X_{k \wedge M} + -X_k + X_{k \wedge N} = X_{k \wedge N} - X_{k \wedge M} = X_N - X_M
\end{aligned}$$

Now since $(H \bullet X)_n$ is a submartingale, we get

$$EX_N - EX_M = E[(H \bullet X)_k] \geq E[(H \bullet X)_0] = 0$$

. □

9. (Durrett 4.4.9) Let $1 \leq m \leq n$, then

$$E[(X_m - X_{m-1})(Y_m - Y_{m-1})] = E[X_m Y_m] + E[X_{m-1} Y_{m-1}] - E[X_m Y_{m-1}] - E[Y_m X_{m-1}]$$

Notice that

$$\begin{aligned}
E[X_m Y_{m-1}] &= E[E(X_m Y_{m-1} | \mathcal{F}_{m-1})] = E[Y_{m-1} E(X_m | \mathcal{F}_{m-1})] = E[X_{m-1} Y_{m-1}] \\
E[Y_m X_{m-1}] &= E[E(Y_m X_{m-1} | \mathcal{F}_{m-1})] = E[X_{m-1} E(Y_m | \mathcal{F}_{m-1})] = E[X_{m-1} Y_{m-1}]
\end{aligned}$$

so the above expression simplifies to

$$E[(X_m - X_{m-1})(Y_m - Y_{m-1})] = E[X_m Y_m] - E[X_{m-1} Y_{m-1}]$$

of course that means we can telescope the following series

$$\sum_1^n E[(X_m - X_{m-1})(Y_m - Y_{m-1})] = \sum_1^n E[X_m Y_m] - E[X_{m-1} Y_{m-1}] = E[X_n Y_n] - E[X_0 Y_0]$$

□

10. (Durrett 4.6.1) Denote $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$, and $\mathcal{F} = \sigma\left(\bigcup_1^\infty \mathcal{F}_n\right)$, then since $\mathcal{F}_n \uparrow \mathcal{F}$, we get $E(\theta|\mathcal{F}_n) \to E(\theta|\mathcal{F})$ a.s. and in L^1 (Durrett theorem 4.6.8). From here the problem will be done if we can show that $\theta \in \mathcal{F}$, since denoting $\phi = E(\theta|\mathcal{F})$, then $A = \{\theta > \phi\} \in \mathcal{F}$ and $B = \{\phi > \theta\} \in \mathcal{F}$, so that we can use the property of conditional expectation to conclude

$$E|\theta - \phi| = \int_A \theta - \phi + \int_B \phi - \theta = \int_A \theta - \theta + \int_B \theta - \theta = 0$$

so indeed $\theta = \phi$ a.s. so that $E(\theta|\mathcal{F}_n) \to \theta$ a.s. and in L^1. It remains to show that $\theta \in \mathcal{F}$, to do so we can use that the pointwise limit of measurable functions is measurable, and realize θ as a pointwise limit of $\mathcal{F}$ measurable functions, since $E|Z_i| < \infty$, and iid, the strong law of large numbers tells us that $\frac{1}{n}\sum_1^n Z_j \to EZ_1$ almost surely, but then this of course implies that

$$S_n := -EZ_1 + \frac{1}{n}\sum_1^n Y_j = -EZ_1 + \frac{1}{n}\sum_1^n Z_j + \theta \to \theta$$

almost surely. Since each S_n is $\mathcal{F}$ measurable, so is its pointwise limit θ. So $\theta \in \mathcal{F}$ and we are done by the argument above. □