(4.4) Fix $n$ and define the bounded stopping time $N = n \wedge \min\{k : |S(k)| \geq x\}$. Since $|\xi_m| \leq K$ we have $|S_N| \leq x + K$, so $ES_N^2 \leq (x + K)^2$. Also
\[ ES_N^2 \geq E(s_n^21_{\{N=n\}}) = s_n^2P(N = n) \geq s_n^2P(\max_{m \leq n} |S_m| \leq x). \]

By Exercise 2.6, $\{S_n^2 - s_n^2\}$ is a martingale, so by Theorem 4.1,
\[ 0 = E(S_n^2 - s_n^2) = E(S_N^2 - s_n^2) \leq (x + K)^2 - s_n^2P(\max_{m \leq n} |S_m| \leq x). \]

Equivalently,
\[ P(\max_{m \leq n} |S_m| \leq x) \leq \frac{(x + K)^2}{s_n^2}. \]

(4.5) Since $(x + c)^2$ is a convex function of $x$ and $\{X_n\}$ is a martingale, $\{(X_n + c)^2\}$ is a submartingale for all $c$. Consider $c > 0$ and $\lambda > 0$. By Doob’s Inequality, since $EX_n = EX_0 = 0$ for all $n$,
\[ P(\max_{m \leq n} X_m \geq \lambda) \leq P(\max_{m \leq n} (X_m + c)^2 \geq (\lambda + c)^2) \]
\[ \leq \frac{E(X_n + c)^2}{(\lambda + c)^2} \]
\[ = \frac{EX_n^2 + c^2}{(\lambda + c)^2}. \]

Let $f(c)$ be the fraction on the right side of (1). Since (1) is valid for all $c > 0$, we can minimize $f(c)$ over $c$: $f'(c) = 0$ for $c = EX_n^2/\lambda$ by easy calculus, so
\[ P(\max_{m \leq n} X_m \geq \lambda) \leq f\left(\frac{EX_n^2}{\lambda}\right) = \frac{EX_n^2}{EX_n^2 + \lambda^2}. \]

(4.8) Let $W_m = X_m - X_{m-1}$ and $\Delta_m = Y_m - Y_{m-1}$. For $m < l$ we have $\Delta_m \in \mathcal{F}_{l-1}$, and since $\{W_l\}$ is a martingale we have $E(W_l \mid \mathcal{F}_{l-1}) = 0$. Therefore
\[ E(W_l \Delta_m \mid \mathcal{F}_{l-1}) = \Delta_m E(W_l \mid \mathcal{F}_{l-1}) = 0, \]
so \( E(W_l \Delta_m) = 0 \) and similarly \( E(W_m \Delta_l) = 0 \). Thus

\[
\sum_{m=1}^{n} E(W_m \Delta_m) = \sum_{m=1}^{n} \sum_{l=1}^{n} E(W_m \Delta_l)
\]

\[
= E\left( \sum_{m=1}^{n} W_m \right) \sum_{l=1}^{n} \Delta_l
\]

\[
= E(X_n - X_0)(Y_n - Y_0)
\]

\[
= E(X_n Y_n) - E(X_0 Y_n) - E(X_n Y_0) + E(X_0 Y_0).
\]

Now \( E(X_0 Y_n | \mathcal{F}_0) = X_0 E(Y_n | \mathcal{F}_0) = X_0 Y_0 \), so \( E(X_0 Y_n) = E(X_0 Y_0) \). Similarly \( E(X_n Y_0) = E(X_0 Y_0) \). Therefore the right side of (2) is \( E(X_n Y_n) - E(X_0 Y_0) \).

(5.1) Let \( \epsilon > 0 \). There exists \( M \) such that \( x \geq M \) implies \( x \leq \epsilon \varphi(x) \). Hence for all \( i \in I \),

\[
E(|X_i||\{\|X_i\| \geq M\}) \leq \epsilon E(\varphi(\|X_i\|)\{\|X_i\| \geq M\}) \leq \epsilon E\varphi(\|X_i\|) \leq C\epsilon.
\]

Since \( \epsilon \) is arbitrary this shows that \( \{X_i : i \in I\} \) is u.i.

(5.2) Let \( \mathcal{F}_n = \sigma(Y_1, ..., Y_n) \) and \( \mathcal{F}_\infty = \sigma(\cup_{n \geq 1} \mathcal{F}_n) \). We have

\[
\lim_{n} \frac{Y_1 + ... + Y_n}{n} = \theta + \lim_{n} \frac{Z_1 + ... + Z_n}{n} = \theta + 0 = \theta \text{ a.s.,}
\]

so \( \theta \in \mathcal{F}_\infty \). Hence \( E(\theta | \mathcal{F}_n) \to E(\theta | \mathcal{F}_\infty) = \theta \text{ a.s.} \)

(5.8) Suppose \( \mathcal{F}_n \not\subset \mathcal{F}_\infty \) and \( Y_n \to Y \) in \( L^1 \). \( \{E(Y | \mathcal{F}_n) : n \geq 1\} \) is u.i. and converges a.s. to \( E(Y | \mathcal{F}_\infty) \) (shown in lecture) so by Theorem 5.3, \( E(Y | \mathcal{F}_n) \to E(Y | \mathcal{F}_\infty) \) in \( L^1 \). Hence using Jensen’s inequality,

\[
|E(Y_n | \mathcal{F}_n) - E(Y | \mathcal{F}_\infty)| \leq |E(Y_n | \mathcal{F}_n) - E(Y | \mathcal{F}_n)| + |E(Y | \mathcal{F}_n) - E(Y | \mathcal{F}_\infty)|
\]

\[
\leq E(|Y_n - Y| | \mathcal{F}_n) + E(Y | \mathcal{F}_n) - E(Y | \mathcal{F}_\infty)|.
\]

The convergence in \( L^1 \) means the expected value of the last term here \( \to 0 \), so taking \( E(\cdot) \) of both sides we get

\[
E(|E(Y_n | \mathcal{F}_n) - E(Y | \mathcal{F}_\infty)|) \leq E|Y_n - Y| + o(1),
\]

which \( \to 0 \) as \( n \to \infty \).

(A) We want to show that

\[
P(A) \leq \delta \implies \int_{A} |X| \, dP \leq \int_{\{\|X\| > M\}} X \, dP.
\]
In fact,

\[
\int_A |X| \, dP = \int_{A \cap \{|X| > M\}} |X| \, dP + \int_{A \cap \{|X| \leq M\}} |X| \, dP \\
\leq \int_{A \cap \{|X| > M\}} |X| \, dP + M(P(A) - P(A \cap \{|X| > M\})) \\
\leq \int_{A \cap \{|X| > M\}} |X| \, dP + M(P(|X| > M) - P(A \cap \{|X| > M\})) \\
= \int_{A \cap \{|X| > M\}} |X| \, dP + MP(A^c \cap \{|X| > M\}) \\
\leq \int_{A \cap \{|X| > M\}} |X| \, dP + \int_{A \cap \{|X| > M\}} |X| \, dP \\
= \int_{\{|X| > M\}} |X| \, dP.
\]