Lectures on 15.4. and 17.4.
We finish the construction of the stochastic integral with respect to a continuous martingale $M$. Let $H$ be a simple left-continuous $\mathbb{F}$-adapted process:

$$
H=\sum_{k=1}^{n} \alpha_{k} I_{\left(a_{k}, b_{k}\right]}(\cdot)
$$

where $\alpha_{k} \in F_{a_{k}}$ and $\alpha_{k}$ is a bounded random variable. Define the stochastic integral by

$$
\begin{equation*}
N_{t} \doteq(H \circ M)_{t} \doteq \int_{0}^{t} H_{s} d M_{s} \doteq \sum_{k=1}^{n} H_{k}\left(M_{t \wedge b_{k}}-M_{t \wedge a_{k}}\right) \tag{4.4}
\end{equation*}
$$

Theorem 4.1. Let $H$ be a simple left-continuous process and let $M \in \mathcal{M}^{2}$. If $N=H \circ M$ is given by (4.4), then $N \in \mathcal{M}^{2}$, we have the isometry

$$
\begin{equation*}
\mathbb{E} N_{\infty}^{2}=\mathbb{E} \int_{0}^{\infty} H_{s}^{2} d[M, M]_{s} \tag{4.5}
\end{equation*}
$$

and

$$
\begin{equation*}
[N, N]=H^{2} \cdot[M . M] \tag{4.6}
\end{equation*}
$$

Proof We can assume that $a_{k}, b_{k}$ satisfy

$$
a_{1}<b_{1} \leq a_{2}<b_{2} \leq \cdots \leq a_{n-1}<b_{n-1} \leq a_{n}<b_{n}
$$

Now its easy to check that $N$ is a continuous martingale and $N \in \mathcal{M}^{2}$. Next we check (4.5): Note first that
$\mathbb{E} N_{\infty}^{2}=\mathbb{E}\left(\sum H_{k}^{2}\left(M_{b_{k}}-M_{a_{k}}\right)^{2}\right)+2 \mathbb{E}\left(\sum_{i<j} H_{i} H_{j}\left(M_{b_{i}}-M_{a_{i}}\right)\left(M_{b_{j}}-M_{a_{j}}\right)\right)$.
We show that

$$
\mathbb{E}\left(H_{i} H_{j}\left(M_{b_{i}}-M_{a_{i}}\right)\left(M_{b_{j}}-M_{a_{j}}\right)\right)=0
$$

this follows from the fact that

$$
\mathbb{E}\left(H_{i} H_{j}\left(M_{b_{i}}-M_{a_{i}}\right)\left(M_{b_{j}}-M_{a_{j}}\right) \mid F_{a_{j}}\right)=0
$$

On the other hand

$$
\begin{aligned}
\mathbb{E}\left[H_{k}^{2}\left(M_{b_{k}}-M_{a_{k}}\right)^{2}\right] & =\mathbb{E}\left[H_{k}^{2} \mathbb{E}\left[\left(M_{b_{k}}-M_{a_{k}}\right)^{2} \mid F_{a_{k}}\right]\right] \\
& =\mathbb{E}\left[H_{k}^{2} \mathbb{E}\left[M_{b_{k}}^{2}-M_{a_{k}}^{2} \mid F_{a_{k}}\right]\right] \\
& =\mathbb{E}\left[H_{k}^{2} \mathbb{E}\left[[M, M]_{b_{k}}-[M, M]_{a_{k}} \mid F_{a_{k}}\right]\right] \\
& =\mathbb{E}\left[H_{k}^{2}\left([M, M]_{b_{k}}-[M, M]_{a_{k}}\right)\right]
\end{aligned}
$$

This proves the equality (4.5). To finish the proof one must show that the process $N^{2}-H^{2} \cdot[M, M]$ is a martingale. This is shown similarly.
Denote by $\mathcal{L}(\mathbb{F})$ left-continuous $\mathbb{F}$ - adapted processes with right-hand limits.
Then we can write

$$
\mathcal{P}(\mathbb{F})=\sigma(\mathcal{L}(\mathbb{F})
$$

Lemma 4.2. Let $H \in \mathcal{P}(\mathbb{F})$ with

$$
\mathbb{E} \int_{0}^{\infty} H_{s}^{2} d\langle M, M\rangle_{s}<\infty
$$

Then there exists a sequence $H^{n}$ of simple left-continuous processes such that

$$
\begin{equation*}
\mathbb{E} \int_{0}^{\infty}\left(H_{s}-H_{s}^{(n)}\right)^{2} d\langle M, M\rangle_{s} \rightarrow 0 \tag{4.7}
\end{equation*}
$$

as $n \rightarrow \infty$.
Proof Assume that we have proved the claim for bounded $H$. If $H$ satisfies $\mathbb{E} \int_{0}^{\infty} H_{s}^{2} d\langle M, M\rangle_{s}<\infty$, then for every $\epsilon>0$ we can find $K \geq K_{\epsilon}$ such that

$$
\mathbb{E} \int_{0}^{\infty}\left(H_{s}-H_{s}^{K}\right)^{2} d\langle M, M\rangle<\epsilon
$$

where $H^{K}=H \wedge K$.
Let $H$ be left-continuous $\mathbb{F}$ - adapted process with (4.1) and $H \geq 0$. We assume that $H \leq K$. Let $\frac{k-1}{2^{n}}<s \leq \frac{k}{2^{n}}$ and $H_{\frac{k-1}{2^{n}}} \leq n ; H_{s}^{n}=H_{\frac{k-1}{2^{n}}}$ and if $H_{\frac{k-1}{2^{n}}}>n$, then put $H_{s}^{n}=n$, when $k=1, \ldots, n 2^{n}$ and otherwise $H_{s}^{n}=0$. We have that by left-continuity $H=\lim _{n} H^{n}$ and the process $H^{(n)}$ satisfies (4.1). DCT theorem implies that $H^{n} \rightarrow H$ in the space $\mathbb{L}^{2}\left(\mathbb{P} \otimes[M, M]_{\infty}\right)$. Now if $H \in \mathcal{L}(\mathbb{F})$ is bounded and the process $H$ satisfies

$$
\mathbb{E} \int_{0}^{\infty} H_{s}^{2} d\langle M, M\rangle_{s}<\infty
$$

then by considering the decomposition $H=H^{+}-H^{-}$we get in the usual way that (4.7) holds.
Define the distance of two process $H, L$ by $\rho(H, L)$

$$
\rho(H, L)=\left\|\left(\int_{0}^{\infty}\left(H_{s}-L_{s}\right)^{2} d\langle M, M\rangle_{s}\right)^{\frac{1}{2}}\right\|_{2}
$$

Denote by $\mathcal{H}$ the following class of processes

$$
\mathcal{H}=\left\{\begin{array}{c}
H \in \mathcal{P}:|H| \leq K, \mathbb{E} \int_{0}^{\infty} H^{2} d\langle M, M\rangle_{s}<\infty \text { and } \\
\forall \epsilon>0 \exists L \in \mathcal{L} \text { such that } \rho(H, L)<\epsilon
\end{array}\right\}
$$

We have that the space of $K$ - bounded left-continuous processes $\mathcal{L}^{K}(\mathbb{F}) \subset$ $\mathcal{H}$. Moreover, if $H^{n} \in \mathcal{H}, H^{n} \uparrow H \geq 0$ and $H \in \mathbb{L}^{2}\left(\mathbb{P} \otimes[M, M]_{\infty}\right)$, then $H \in \mathcal{H}$. So $\mathcal{H}$ is a monotonic class. Hence $\sigma\left(\mathcal{L}^{K}\right)=\mathcal{H}$.
Corollary 4.1. Let $H \in \mathcal{P} \cap L^{2}\left(\mathbb{P} \otimes\langle M, M\rangle_{\infty}\right)$. Then there exists a sequence $H^{n}$ of simple left-continuous processes such that

$$
H^{n} L^{2}\left(\mathbb{P} \otimes\langle M, M\rangle_{\infty}\right) \quad H .
$$

Stochastic integral. Let us define the stochastic integral for $H \in \mathcal{P} \cap L^{2}(\mathbb{P} \otimes$ $\left.\langle M, M\rangle_{\infty}\right)$. Let $H^{n}$ be a sequence of simple left-continuous processes, which satisfy $H^{n} \xrightarrow{L^{2}\left(\mathbb{P} \otimes\langle M, M\rangle_{\infty}\right)} H$. By Doob's $\mathbb{L}^{2}$ maximal inequality

$$
\begin{aligned}
\left\|\left(H^{n}-H^{m}\right) \circ M\right\|_{\mathcal{M}^{2}} & =\left\|\left(\left(H^{n}-H^{m}\right) \circ M\right)_{\infty}\right\|_{2} \\
& \leq 2\left(\mathbb{E} \int_{0}^{\infty}\left(H_{s}^{n}-H_{S}^{m}\right)^{2} d\langle M, M\rangle_{s}\right)^{\frac{1}{2}} \rightarrow 0
\end{aligned}
$$

as $m, n \rightarrow \infty$. So the sequence $H^{n} \circ M$ is a c-sequence in the space $\mathcal{M}^{2}$, and so it has a limit $N \in \mathcal{M}^{2}$. Moreover $\sup _{t}\left(\left(H^{n} \circ M\right)_{t}-N_{t}\right) \rightarrow 0$ in the space $L^{2}(\mathbb{P})$. Let us check that the limit $N$ does not depend on the approximating sequence $H^{n}$ : let $\tilde{H}^{n}$ be another sequence, which satisfies $\rho\left(H, \tilde{H}^{n}\right) \rightarrow 0$, as $n \rightarrow \infty$, then

$$
\mathbb{E}\left(\int_{0}^{\infty}\left(H_{s}^{n}-\tilde{H}_{s}^{n}\right) d M_{s}\right)^{2}=\mathbb{E} \int_{0}^{\infty}\left(H_{s}^{n}-\tilde{H}_{s}^{n}\right)^{2} d\langle M, M\rangle_{s} \rightarrow 0,
$$

as $n \rightarrow \infty$; this means that the limit $N$ is independent of the sequence $H^{n}$ used for the approximation of $H$.
Process $N$ is the stochastic integral of $H$ with respect to the martingale $M \in \mathcal{M}^{2}$; we continue to use the notation $N=H \circ M$. We have the isometry

$$
\begin{equation*}
\mathbb{E}\left(\int_{0}^{\infty} H_{s} d M_{s}\right)^{2}=\mathbb{E} \int_{0}^{\infty} H_{s}^{2} d\langle M, M\rangle_{s} \tag{4.8}
\end{equation*}
$$

and

$$
\begin{equation*}
\langle H \circ M, H \circ M\rangle=H^{2} \cdot\langle M, M\rangle . \tag{4.9}
\end{equation*}
$$

The equality (4.8) follows directly the fact that it is true for the approximating sequence. The proof of the claim (4.9) is an exercise.
Let us discuss the following situation:
Let $H$ be a predictable process, which satisfies

$$
\begin{equation*}
\int_{0}^{\infty} H_{s}^{2} d\langle M, M\rangle_{s}<\infty, \tag{4.10}
\end{equation*}
$$

but the condition (4.1) is not fulfilled. Then we can define the stochastic integrals locally as follows: let $K>0$ and put

$$
\tau_{K}=\inf \left\{t: \int_{0}^{t} H_{s}^{2} d\langle M, M\rangle_{s}=K\right\}
$$

Note that $H I_{\left(0, \tau_{K}\right]}$ is a predictable process, and we have

$$
\int_{0}^{\infty}\left(H I_{\left[0, \tau_{K}\right]}\right]_{s}^{2} d\langle M, M\rangle_{s} \leq K
$$

and the condition (4.1) is in force. On the other hand

$$
\left.\int_{0}^{\infty}\left(H I_{\left(0, \tau_{K}\right]}\right)_{s} d M_{s}=(H \circ M)^{\tau_{K}}\right)_{\infty}=\left(H \circ M^{\tau_{K}}\right)_{\infty}
$$

[for more details see the next weeks exercise]. Hence we can define the stochastic integral as a local martingale $H \circ M$, where the pair of processes $H,\langle M, M\rangle$ satisfies the condition (4.10) with the localizing sequence $\tau_{K}$, $K \geq 1$. Here $\tau_{K} \rightarrow \infty$, and so $\int_{0}^{t} H_{s} d M_{s}:=\lim _{K \rightarrow \infty} \int_{0}^{t} H_{s} d M_{s}^{\tau_{K}}$, where the limit is almost surely. The same argument applies when $M$ is only a local martingale.

Remark 4.2. Note that the condition $\mathbb{E} \int_{0}^{T} H_{s}^{2} d\langle M, M\rangle_{s}<\infty$ gives that the stochastic integral $(H \circ M)$ is a square integrable martingale on the interval $[0, T]$ But if we have only that $\int_{0}^{T} H_{s}^{2} d\langle M, M\rangle_{s}<\infty$ the stochastic integral is only a local martingale on $[0, T]$.

Remark 4.3. Consider the case when $H=f(M), f$ is continuous, bounded and we have

$$
\mathbb{E} \int_{0}^{\infty} f^{2}\left(M_{s}\right) d\langle M, M\rangle_{s}<\infty
$$

Then $H=f(M)$ is continuous, and hence predictable. We can approximate $H$ by

$$
H_{s}^{n}=f\left(M_{t_{k-1}}\right) 1_{\left(t_{k-1}, t_{k}\right]}(s)
$$

with $t_{k} \leq n$, and $H_{s}^{n}=f\left(M_{n}\right) 1_{(n, \infty)}(s)$. Then, if $T \leq n$,

$$
\left(H^{n} \circ M\right)_{T}=\sum_{t_{k} \leq T} f\left(M_{t_{k-1}}\right)\left(M_{t_{k}}-M_{t_{k-1}}\right) \rightarrow \int_{0}^{T} f\left(M_{s}\right) d M_{s}
$$

in $L^{2}(\mathbb{P})$ and hence in probability, too. In this sense stochastic integral is a Riemann-Stieltjes integral.

### 4.3. Itô formula.

Theorem 4.2 (Integration by parts). Let $M, N \in \mathcal{M}^{2}$. Then we have the following integration by parts formula:

$$
\begin{equation*}
M N=M \circ N+N \circ M+\langle M, N\rangle \tag{4.11}
\end{equation*}
$$

Proof Note that the processes $M, N$ are continuous, and hence they are predictable, and so the stochastic integrals in the formula (4.11) well defined, and at least local martingales.
From the polarization formula $M N=\frac{1}{4}\left((M+N)^{2}-(M-N)^{2}\right)$ it follows that it is enough to prove the claim in the case of $N=M$ :

$$
M^{2}=2 M \circ M+\langle M, M\rangle
$$

From the theorem 3.10 we obtain that the process $M^{2}-\langle M, M\rangle$ is a local martingale. Further, the process $C^{n}$ constructed in the proof of theorem 3.10 , where

$$
C_{t}^{n}=\sum_{k} M_{\tau_{k}^{n}} I_{\left\{t \in\left(\tau_{k}^{n}, \tau_{k+1}^{n}\right]\right\}}
$$

is predictable, $C^{n} \rightarrow M$, and hence we must have the equality $M^{2}-$ $[M, M]=2 M \circ M$. This proves the claim.
Assume that $M$ is a bounded continuous martingale, $\pi$ is a partition of the interval $[0, T]$, with $|\pi| \rightarrow 0$. From the Abel summation formula we obtain for $t \in[0, T]$, as $|\pi| \rightarrow 0$ :

$$
\begin{aligned}
M_{t}^{2} & =2 \sum_{k} M_{t \wedge t_{k-1}}\left(M_{t \wedge t_{k}}-M_{t \wedge t_{k-1}}\right)+\sum_{k}\left(M_{t \wedge t_{k}}-M_{t \wedge t_{k-1}}\right)^{2} \\
& \doteq 2\left(M^{\pi} \circ M\right)_{t}+Q_{t}^{\pi}
\end{aligned}
$$

We shall show that $Q_{t}^{\pi} \xrightarrow{\mathbb{P}}\langle M, M\rangle_{t}$, as $\pi \rightarrow 0$. Note first that if $\eta$ is another partition, then $\left(M^{\pi}-M^{\eta}\right)_{T}^{*} \xrightarrow{\mathbb{P}} 0$, as $|\pi| \vee|\eta| \rightarrow 0$. From this we obtain $M \pi^{n} \circ M$ is a c-sequence, if $\left|\pi^{n}\right| \rightarrow 0$, as $n \rightarrow \infty$. Hence the martingales $M^{\pi} \circ M$ converge towards a continuous martingale $N$. From this we obtain that also the sequence of random variables $Q^{\pi}$ converges to the limit $\tilde{Q}$. After passing to the limit we will get that the process $M^{2}-\tilde{Q}$ is a martingale [recall that we have assumed that $M$ is a continuous and bounded martingale] and the process $\tilde{Q}$ is continuous and increasing. From this we
obtain that $\tilde{Q}=\langle M, M\rangle$, because the angle bracket process $\langle M, M\rangle$ is the unique increasing continuous process with the property that $M^{2}-\langle M, M\rangle$ is a (local) martingale..

## Continuous semimartingales.

Definition 4.1. Let $X$ be a continuous process adapted to $\mathbb{F}$. The process $X$ is a (continuous) semimartingale, if it satisfies $X=X_{0}+M+A$, where $M$ is a continuous local martingale and the process $A$ has locally bounded variation, and we assume that $M_{0}=A_{0}=0$.

The representation of $X$ as $X=X_{0}+M+A$ is unique; this follows from theorem 3.7.

Integration with respect to a continuous semimartingale. Let $X$ be a continuous semimartingale with the decomposition $X=X_{0}+M+A$ and let $H$ be a predictable process such that the Riemann-Stieltjes-integrals of $H$ with respect to $A$ and $\langle M, M\rangle$ satisfy $\int_{0}^{t}\left|H_{s}\right| d A_{s}<\infty$ and $\int_{0}^{t} H_{s}^{2} d\langle M, M\rangle_{s}<\infty$. Then the integral $H \circ X:=H \circ M+H \cdot A$ is well-defined: the first integral is a stochastic integral and the second is a Riemann-Stieltjes integral.

Integration by parts formula for continuous semimartingales. Let $X=X_{0}+$ $M+A$. Define $\langle X, X\rangle=\langle M, M\rangle$. Then we have that

$$
\begin{equation*}
\langle X, X\rangle_{t}=\mathbb{P}-\lim _{|\pi| \rightarrow 0} \sum_{t_{k} \in \pi}\left(X_{t_{k}}-X_{t_{k-1}}\right)^{2} \tag{4.12}
\end{equation*}
$$

We shall prove this. We can assume that $M$ and $A$ are bounded and that $X_{0}=0$. Then also $X$ is bounded. Now

$$
\begin{aligned}
\sum_{k}\left(X_{t_{k}}-X_{t_{k-1}}\right)^{2}= & \sum_{k}\left(M_{t_{k}}-M_{t_{k-1}}\right)^{2}+\sum_{k}\left(A_{t_{k}}-A_{t_{k-1}}\right)^{2} \\
& +2 \sum_{k}\left(A_{t_{k}}-A_{t_{k-1}}\right)\left(M_{t_{k}}-M_{t_{k-1}}\right)
\end{aligned}
$$

Because the process $A$ has bounded variation, then

$$
\sum_{k}\left(A_{t_{k}}-A_{t_{k-1}}\right)^{2} \leq \max _{t_{k}}\left|A_{t_{k}}-A_{t_{k-1}}\right| \mathcal{V}_{t}(A) \xrightarrow{\mathbb{P}} 0
$$

Further, since the process $M$ is continuous, we obtain using the bounded variation property of $A$ that

$$
\left|\sum_{k}\left(A_{t_{k}}-A_{t_{k-1}}\right)\left(M_{t_{k}}-M_{t_{k-1}}\right)\right| \leq \max _{t_{k}}\left|M_{t_{k}}-M_{t_{k-1}}\right| \mathcal{V}_{t}(A) \xrightarrow{\mathbb{P}} 0
$$

After these observations it is clear that in (4.12): the only non-zero limit is the angle bracket process of the martingale $M$.
The following is an exercise:
Theorem 4.3. Let $X, Y$ be continuous semimartingales. Then

$$
\begin{equation*}
X Y=X_{0} Y_{0}+X \circ Y+Y \circ X+\langle X, Y\rangle \tag{4.13}
\end{equation*}
$$

Itô formula. Let $F: \mathbb{R} \rightarrow \mathbb{R}$, and $F \in C_{2}$; put $F_{x}=\frac{\partial F}{\partial x}$ and $F_{x x}=\frac{\partial^{2} F}{\partial x^{2}}$. Further, if $G: \mathbb{R}_{+} \times \mathbb{R} \rightarrow \mathbb{R}$, then we put $G_{t}=\frac{\partial F}{\partial t}, G_{x}=\frac{\partial G}{\partial x}$ and $G_{x x}=\frac{\partial^{2} G}{\partial x^{2}}$.
We start with the following Itô formula. The bautiful proof is due to H.P. McKean.

Theorem 4.4. Let $X$ be a continuous semimartingale and $F \in C_{2}$. Then

$$
\begin{equation*}
F\left(X_{t}\right)=F\left(X_{0}\right)+\int_{0}^{t} F_{x}\left(X_{s}\right) d X_{s}+\frac{1}{2} \int_{0}^{t} F_{x x}\left(X_{s}\right) d\langle X, X\rangle_{s} \tag{4.14}
\end{equation*}
$$

Proof Let $X=X_{0}+M+A$ be the semimartingale decomposition of $X$.. Let $\mathcal{I}$ be the family of functions in $F \in C_{2}$ satisfying the Itô formula.
It is easy to check that the functions $F(x)=1, F(x)=x$ belong to the family $\mathcal{I}$. Moreover, if $F, G \in \mathcal{I}$ and $a, b \in \mathbb{R}$. Then we have $(a F+b G)_{x}=a F_{x}+b G_{x}$ and $(a F+b G)_{x x}=a F_{x x}+b G_{x x}$, and so $a F+b G \in \mathcal{I}$. This means that $\mathcal{I}$ is a linear vector space.
Let now $F, G \in \mathcal{I}$ : then we have

$$
F\left(X_{t}\right)=F\left(X_{0}\right)+\int_{0}^{t} F_{x}\left(X_{s}\right) d X_{s}+\frac{1}{2} \int_{0}^{t} F_{x x}\left(X_{s}\right) d\langle X, X\rangle_{s}
$$

and

$$
G\left(X_{t}\right)=G\left(X_{0}\right)+\int_{0}^{t} G_{x}\left(X_{s}\right) d X_{s}+\frac{1}{2} \int_{0}^{t} G_{x x}\left(X_{s}\right) d\langle X, X\rangle_{s}
$$

Observe that $F(X)$ is a continuous semimartingale with the decomposition

$$
\int_{0}^{t} F_{x}\left(X_{s}\right) d M_{s}
$$

as the local martingale part and the process

$$
\int_{0}^{t} F_{x}\left(X_{s}\right) d A_{s}+\frac{1}{2} \int_{0}^{t} F_{x x}\left(X_{s}\right) d\langle X, X\rangle_{s}
$$

as the local martingale part.
We can now compute using the integration by parts formula (4.13):

$$
\begin{aligned}
F\left(X_{t}\right) G\left(X_{t}\right)= & F\left(X_{0}\right) G\left(X_{0}\right)+(F(X) \circ G(X))_{t}+(G(X) \circ F(X))_{t}+\langle F(X), G(X)\rangle_{t} \\
= & F\left(X_{0}\right) G\left(X_{0}\right)+\int_{0}^{t}\left(F_{x}\left(X_{s}\right) G\left(X_{s}\right)+G_{x}\left(X_{s}\right) F\left(X_{s}\right)\right) d X_{s} \\
& +\frac{1}{2} \int_{0}^{t}\left(2 F_{x}\left(X_{s}\right) G_{x}\left(X_{s}\right)+F\left(X_{s}\right) G_{x x}\left(X_{s}\right)+G\left(X_{s}\right) F_{x x}\left(X_{s}\right)\right) d\langle X, X\rangle_{s}
\end{aligned}
$$

where we have used the fact that the martingale part of the semimartingale $F(X)$ is $F_{x}(X) \circ M$ (resp. the martingale part of the semimartingale $G(X)$ is $G_{x}(X) \circ M$ and we also have $\langle M, M\rangle=\langle X, X\rangle$ by definition. Moreover, we have used the following property of stochastic integrals: $\langle H \circ X, K \circ X\rangle=$ $H K \cdot\langle X, X\rangle$.
Recall the differentiation rules for the product: $(F G)_{x}=F_{x} G+G_{x} G$ and $(F G)_{x x}=F_{x x} G+G_{x x} F+2 F_{x} G_{x}$. This means that we can write
$F\left(X_{t}\right) G\left(X_{t}\right)=F\left(X_{0}\right) G\left(X_{0}\right)+\int_{0}^{t}(F G)_{x}\left(X_{s}\right) d X_{s}+\frac{1}{2} \int_{0}^{t}(F G)_{x x}\left(X_{s}\right) d\langle X, X\rangle_{s}$.

Hence $F G \in \mathcal{I}$. This means that our set $\mathcal{I}$ is a vector space and algebra. Since $1 \in \mathcal{I}$ and the identity map is in $\mathcal{I}$. Since $\mathcal{I}$ is a vector space and an algebra, this means that Itô formula is valid for all polynomials.
The next step is to use Weierstrass approximation theorem, which tells that every continuous function can be approximated by polynomials. So we have $X$ a continuous semimartingale with $X=X_{0}+M+A$. We also know that equality (4.14) is valid for polynomials $p(x)$.
By the Weierstrass approximation theorem there exists a sequence of polynomials $p_{1}, p_{2}, \ldots$ such that $\sup _{|x| \leq c}\left|p_{n}(x)-F_{x x}(x)\right| \rightarrow 0$, and $n \rightarrow \infty$ and $c>0$. By integrating polynomials $p_{n}$ twice we obtain a sequence $f^{n}$ of polynomials, which satisfy

$$
\sup _{|x| \leq c}\left(\left|f^{n}(x)-F(x)\right| \vee\left|f_{x}^{n}(x)-F_{x}(x)\right| \vee\left|f_{x x}^{n}(x)-F_{x x}(x)\right|\right) \rightarrow 0
$$

when $n \rightarrow \infty$ and $c>0$.
Let us now assume that $X_{t}^{*} \leq c$ and $\langle M, M\rangle_{t} \leq c$ for some $c>0$. Then $f^{n}\left(X_{t}\right) \rightarrow f\left(X_{t}\right)$ for all $t>0$. Because

$$
\sup _{|y| \leq c}\left(\left|F_{x}(y)\right|+\left|f_{x}^{n}(y)\right|+\left|F_{x x}(y)\right|+f_{x x}^{n}(y) \mid\right) \leq K<\infty
$$

then by DCT for Riemann-Stieltjes- integrals we obtain that

$$
\left(f_{x}^{n}(X) \cdot A\right)_{t} \rightarrow\left(F_{x}(X) \cdot A\right)_{t}
$$

and

$$
\left(f_{x x}^{n}(X) \cdot\langle X, X\rangle\right)_{t} \rightarrow\left(F_{x x}(X) \cdot\langle X, X\rangle\right)_{t}
$$

as $n \rightarrow \infty$.
Moreover, for the same reason we have

$$
\left(f_{x}^{n}(X)-F_{x}(X)\right)^{2} \cdot\langle M, M\rangle_{t} \rightarrow 0
$$

and so

$$
\left(\left(f_{x}^{n}(X)-F_{x}(X)\right) \circ M\right)_{t}^{*} \xrightarrow{L^{2}(\mathbb{P})} 0
$$

as $n \rightarrow 0$. Hence the Itô formula is valid in the case of $X_{t}^{*} \leq c$ and $\langle M, M\rangle_{t} \leq$ $c$ for any $c>0$.
Finally we can prove the general case by stopping. Indeed, define stopping time $t_{c}$ by

$$
\tau_{c}=\inf \left\{u:\left|X_{u}\right| \geq c\right\} \wedge \inf \left\{u:\langle M, M\rangle_{u} \geq c\right\}
$$

Then the Itô formula is valid on the set $\left\{\tau_{c} \geq t\right\}$. But the processes $X$ and $\langle M, M\rangle$ are continuous, and so $\mathbb{P}\left(\tau_{c} \geq t\right) \rightarrow 1$ as $c \rightarrow \infty$, and we have proved the Itô formula.

Variants of Itô formula. It is useful to stop here and consider some special cases:

- If the semimartingale $X$ is continuous and has bounded variation on compacts, then $M \equiv 0, X=X_{0}+A$ and we obtain the change of variables formula

$$
F\left(X_{t}\right)=F\left(X_{0}\right)+\int_{0}^{t} F_{x}\left(X_{s}\right) d X_{s}=F\left(X_{0}\right)+\int_{0}^{t} F_{x}\left(X_{s}\right) d A_{s}
$$

- If the semimartingale $X$ is the Brownian motion $W$, then $\langle W, W\rangle_{t}=$ $t$, and we can write the Itô formula as follows

$$
F\left(W_{t}\right)=F(0)+\int_{0}^{t} F_{x}\left(W_{s}\right) d W_{s}+\frac{1}{2} \int_{0}^{t} F_{x x}\left(W_{s}\right) d s .
$$

A sufficient condition for the martingale property of the stochastic integral is $\mathbb{E} \int_{0}^{t}\left(F_{x}\left(W_{s}\right)\right)^{2} d s<\infty$ for all $t \geq 0$. Note also that for Brownian motion we have that $\mathbb{E} W_{t}^{2}=t<\infty$, but $W$ is not in the space $\mathcal{M}^{2}$, because $\sup _{t} \mathbb{E} W_{t}^{2}=\infty$.

- Let $X \in \mathbb{R}^{d}$ be a vector values continuous semimartingale: $X=$ $\left(X^{1}, \ldots, X^{d}\right)^{t}$, where $X^{k}=X_{0}^{k}+A^{k}+M^{k}, k=1, \ldots, d$. If $F$ : $\mathbb{R}^{d} \rightarrow \mathbb{R}$ and $F \in C_{2}\left(\mathbb{R}^{d}\right)$, then

$$
F\left(X_{t}\right)=F\left(X_{0}\right)+\sum_{k=1}^{d} \int_{0}^{t} F_{x_{k}}\left(X_{s}\right) d X_{s}^{k}+\frac{1}{2} \sum_{j, k=1}^{d} \int_{0}^{t} F_{x_{k} x_{j}}\left(X_{s}\right) d\left\langle X^{k}, X^{j}\right\rangle_{s} ;
$$

here $\left\langle X^{k}, X^{j}\right\rangle=\left\langle M^{k}, M^{j}\right\rangle$.

- We can apply the previous result to the special case of $X \in \mathbb{R}^{2}$ with $X_{t}^{1}=t$ and $X_{t}^{2}=W_{t}$. If $G \in C_{1,2}\left(\mathbb{R}_{+}, \mathbb{R}\right)$, then one can show that we have the following variant of the Itô formula:
$G\left(t, W_{t}\right)=G(0,0)+\int_{0}^{t} G_{t}\left(u, W_{u}\right) d u+\int_{0}^{t} G_{x}\left(u, W_{u}\right) d W_{u}+\frac{1}{2} \int_{0}^{t} G_{x x}\left(u, W_{u}\right) d u$.
- In the previous case we can find explicit solutions of certain stochastic integrals. Arrange the terms in (4.15) to obtain
$\int_{0}^{t} G_{x}\left(u, W_{u}\right) d W_{u}=G\left(t, W_{t}\right)-G(0,0)-\int_{0}^{t}\left(G_{t}\left(u, W_{u}\right)+G_{x x}\left(u, W_{u}\right)\right) d u$.
Remark 4.4. Let $X$ be a continuous semimartingale and $F \in C_{2}$. Then by the proof of theorem 4.4 we obtain that $F(X)$ is also a continuous semimartingale.


## 5. Applications of the Itô formula

5.1. A characterization of Brownian motion by Lévy. Recall that the characteristic function $\psi_{X}(\lambda)$ of a random variable $X$ is

$$
\psi_{X}(\lambda)=\mathbb{E} e^{i \lambda X},
$$

where $\lambda \in \mathbb{R}, i=\sqrt{-1}$ and $e^{i y}=i \sin (y)+\cos (y)$, when $y \in \mathbb{R}$. We recall that the distribution of a random variable is uniquely defined by its characteristic function.

Theorem 5.1 (Lévy). Let $X$ be a continuous process with $X_{0}=0$ and $\mathbb{E} X_{t}=0$. The process $X$ is a Brownian motion if and only if the processes $X$ and $X_{t}^{2}-t, t \geq 0$, are martingales with respect to the history $F_{t}^{X}, t \geq 0$.

Proof Let $X$ be a Brownian motion. By theorem 3.4 the processes $X$ and $X_{t}^{2}-t$ are martingales with respect to the history $F_{t}^{X}$.
Conversely, assume that $X$ and $X_{t}^{2}-t$ are martingales. Then the angle bracket of the martingale $X$ is $\langle X, X\rangle_{t}=t$, because the angle bracket process
is unique. Let us apply Itô formula (4.14) to the martingale $X$ and function $F(x)=e^{i \lambda x}$. Then $F_{x}(y)=i \lambda F(y)$ and $F_{x x}(y)=-\lambda^{2} F(y)$.
By considering real and imaginary parts separately we obtain:

$$
\begin{aligned}
e^{i \lambda X_{t}} & =1+\int_{0}^{t} e^{i \lambda\left(X_{u}-X_{s}\right)} d u-\frac{\lambda^{2}}{2} \int_{0}^{t} e^{i \lambda X_{u}} d u \\
& =e^{i \lambda X_{s}}+i \lambda \int_{s}^{t} e^{i \lambda\left(X_{u}-X_{s}\right)} d u-\frac{\lambda^{2}}{2} \int_{s}^{t} e^{i \lambda X_{u}} d u
\end{aligned}
$$

where the last line is obtained, since the representation is also valid at time $s<t$. Let us divide the above equality by $e^{i \lambda X_{s}}$ :

$$
\begin{equation*}
e^{i \lambda\left(X_{t}-X_{s}\right)}=1+i \lambda \int_{s}^{t} e^{i \lambda\left(X_{u}-X_{s}\right)} d X_{u}-\frac{\lambda^{2}}{2} \int_{s}^{t} e^{i \lambda\left(X_{u}-X_{s}\right)} d u \tag{5.1}
\end{equation*}
$$

Next, take conditional expectations on the both sides of (5.1) with respect to sigma-algebra $F_{s}^{X}$ and we obtain

$$
\begin{align*}
\mathbb{E}\left[e^{i \lambda\left(X_{t}-X_{s}\right)} \mid F_{s}^{X}\right]= & 1+i \lambda \mathbb{E}\left[\int_{s}^{t} e^{i \lambda\left(X_{u}-X_{s}\right)} d X_{u} \mid F_{s}^{X}\right] \\
& -\frac{\lambda^{2}}{2} \mathbb{E}\left[\int_{s}^{t} e^{i \lambda\left(X_{u}-X_{s}\right)} d u \mid F_{s}^{X}\right] \\
= & 1-\frac{\lambda^{2}}{2} \int_{s}^{t} \mathbb{E}\left[e^{i \lambda\left(X_{u}-X_{s}\right)} d u \mid F_{s}^{X}\right] d u \tag{5.2}
\end{align*}
$$

where we have used the fact that the stochastic integral is a true martingale, and the order of taking the conditional expectation and integral can be changed by the ordinary Fubini theorem.
Put $g_{s}(u)=\mathbb{E}\left[e^{i \lambda\left(X_{u}-X_{s}\right)} \mid F_{s}^{X}\right]$; now we can rewrite the equality (5.2) as:

$$
g_{s}(t)=1-\frac{\lambda^{2}}{2} \int_{s}^{t} g_{s}(u) d u
$$

We get that $g_{s}(t)=e^{-\frac{\lambda^{2}}{2}(t-s)}$, since $g_{s}(t)$ satisfies the following ordinary differential equation $g_{s}^{\prime}=-\frac{\lambda^{2}}{2} g_{s}$ with initial value $g_{s}(s)=1$. Hence the expression $g_{s}(t)$ is deterministic.
This means that $\mathbb{E}\left[e^{i \lambda\left(X_{t}-X_{s}\right)} \mid F_{s}^{X}\right]=e^{-\frac{\lambda^{2}}{2}(t-s)}$. In other words, for all $A \in F_{s}^{X}$ we have the equality

$$
\begin{aligned}
\int_{A} e^{i \lambda\left(X_{t}-X_{s}\right)} d \mathbb{P} & =\int_{A} e^{-\frac{\lambda^{2}}{2}(t-s)} d \mathbb{P} \\
& =\mathbb{P}(A) e^{-\frac{\lambda^{2}}{2}(t-s)}
\end{aligned}
$$

We have shown that the conditional characteristic function of the increment $X_{t}-X_{s}$ is independent from the condition $F_{s}^{X}$.
Because the characteristic function determines the distribution, this means that the conditional distribution of the increment is independent of $F_{s}^{X}$ : $X_{t}-X_{s} \Perp F_{s}^{X}$ Finally, the characteristic function of the increment $X_{t}-X_{s}$ is $e^{-\frac{\lambda^{2}}{2}(t-s)}$, and so $X_{t}-X_{s} \sim N(0, t-s)$. Hence the process $X$ is a Brownian motion.

