11 Preliminaries for MDPs in continuous time

11.1 Markov processes in continuous time

A stochastic process in continuous time is a collection of random variables \( \{X_t\}_{t \geq 0} \), where each \( X_t \) takes values in some state space \( \mathcal{X} \), and the time variable \( t \) takes values in \( \mathbb{R}_+ := [0, \infty) \). The Markov property in continuous time is described similar to the discrete-time case, by sampling the continuous-time process at an arbitrary finite set of time instants:

**Definition 11.1** A stochastic process \( \{X_t\}_{t \geq 0} \) is said to have the Markov property if, for any sequence of times \( 0 \leq t_1 < \cdots < t_m < t \) and for any \( A \in \mathcal{B}(\mathcal{X}) \),

\[
P[X_t \in A | X_{t_1}, \ldots, X_{t_m}] = P[X_t \in A | X_{t_m}].
\]

We also define transition kernels of a Markov process: For \( 0 \leq s < t \) and \( A \in \mathcal{B}(\mathcal{X}) \),

\[
P_{s,t}(x,A) := P[X_t \in A | X_s = x]
\]

A transition kernel of a Markov process has the following properties:

1. For \( s, t, A \) fixed, \( x \mapsto P_{s,t}(x,A) \) is measurable.
2. For \( s, t, x \) fixed, \( A \mapsto P_{s,t}(x,A) \) is a probability measure on \( (\mathcal{X}, \mathcal{B}(\mathcal{X})) \).
3. (Chapman-Kolmogorov equation) For all \( 0 \leq r < s < t \) and for all \( x, A \),

\[
P_{r,t}(x,A) = \int_{\mathcal{X}} P_{s,t}(x',A)P_{r,s}(x, \mathrm{d}x').
\]

In this course, we will consider two cases: \( \mathcal{X} = \{1, \ldots, n\} \) and \( \mathcal{X} = \mathbb{R}^n \).

11.2 Finite-state case

Consider finite state-space \( \mathcal{X} = \{1, 2, \ldots, n\} \) and a Markov process \( \{X_t\}_{t \geq 0} \) on \( \mathcal{X} \). Let \( \mu_t(\cdot) := P[X_t = \cdot] \) be the probability mass function of \( X_t \). In finite-state case, the transition probabilities are given by a matrix defined for all \( s, t \) with \( 0 \leq s < t \):

\[
P_{s,t} = [P_{s,t}(a,b)]_{a,b \in \mathcal{X}}
\]

For any \( 0 \leq t_1 < t_2 \ldots < t_m \), we have

\[
P[X_{t_1} = a_1, \ldots, X_{t_m} = a_m] = P[X_{t_1} = a_1]P[X_{t_2} = a_2 | X_{t_1} = a_1] \cdots P[X_{t_m} = a_2 | X_{t_{m-1}} = a_{m-1}]
\]

\[
= \mu_{t_1}(a_1) \cdot \prod_{j=1}^{m-1} P_{t_j, t_{j+1}}(a_j, a_{j+1}).
\]

Therefore \( \mu_0, \{P_{s,t}\}_{0 \leq s < t} \) completely characterize the MP. Since the map \( \mu_s \mapsto \mu_t \) is defined by \( \mu_t(i) = \sum_{j \in \mathcal{X}} \mu_s(j)P_{s,t}(j, i) \), we can express this in a matrix form:

\[
\mu_t = \mu_s P_{s,t}, \quad \forall 0 \leq s < t
\]
for row vector $\mu$. The C-K equation is also represented using matrix multiplication:

$$P_{r,t} = P_{r,s}P_{s,t}, \quad \forall \ 0 \leq r < s < t$$

**Definition 11.2** A Markov process $\{X_t\}_{t \geq 0}$ on $\mathcal{X}$ is time-homogeneous if for any $0 \leq s < t$, $i,j \in \mathcal{X}$:

$$P[X_t = j | X_s = i] = P[X_{t-s} = j | X_0 = i]$$

With slightly abusive notation, we denote

$$P_{t-s} := P_{0,t} - P_{0,s} = P_{s,t}$$

Now it seems $\mu_0, \{P_t\}_{t \geq 0}$ can completely characterize a time-homogeneous Markov process in finite-state case. However, it turns out that we only need a single matrix called transition intensity.

### 11.2.1 Transition intensity matrix

We define Q-matrices as following:

$$\Lambda = [\lambda_{i,j}]_{i,j \in \mathcal{X}} \quad \text{s.t.} \quad \begin{cases} 
\lambda_{i,j} \geq 0, & \text{for } i \neq j \\
\sum_j \lambda_{i,j} = 0, & \forall i \in \mathcal{X}
\end{cases}$$

Given a finite-state Markov process characterized by $\mu_0, \{P_t\}_{t \geq 0}$, consider the derivative of $P_t$:

$$\left. \frac{d}{dt} P_t \right|_{t=0} = \lim_{h \downarrow 0} \frac{P_h - P_0}{h} = \lim_{h \downarrow 0} \frac{P_h - I_n}{h} =: \Lambda$$

Without proof we claim that the limit exists, and moreover, it is a Q-matrix. We call $\Lambda$ the transition intensity of the finite-state Markov process. Since

$$\mu_h = \mu_0 P_h = \mu_0 (I_n + h \Lambda + o(h))$$

$$\Rightarrow \mu_h = \mu_0 + h \mu_0 \Lambda + o(h) \Rightarrow \frac{\mu_h - \mu_0}{h} = \mu_0 \Lambda + o(1)$$

$$\Rightarrow \frac{\mu_{t+h} - \mu_t}{h} = \mu_t \Lambda + o(1)$$

we conclude the forward Kolmogorov equation:

$$\frac{d\mu_t}{dt} = \mu_t \Lambda$$

Let us investigate the other direction. Given a Q-matrix $\Lambda$, can a Markov process be constructed? We first prove the following statement. For fixed $\mu \in \mathcal{P}(\mathcal{X})$, consider an ordinary differential equation:

$$\frac{d\mu_t}{dt} = \mu_t \Lambda, \quad \mu_0 = \mu$$

if $\Lambda$ is a Q-matrix, then $\mu_t \in \mathcal{P}(\mathcal{X})$ for all $t \geq 0$.

**Proof:** The solution to the ODE is given by $\mu_t = \mu e^{t\Lambda}$. We first prove that the matrix exponential
$e^{t\Lambda}$ has non-negative entities if $\Lambda$ is a Q-matrix. Let $\lambda := \max_i \sum_{j \neq i} \lambda_{i,j}$, and define $R := I_n + \frac{1}{\lambda} \Lambda$ where $I_n$ is $n$-dimensional identity matrix. It is easy to see $R$ has non-negative entities, and therefore

$$e^{\lambda t R} = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} R^k$$

has non-negative entities. Now it is straightforward that

$$e^{t\Lambda} = e^{\lambda t R - \lambda t I_n} = e^{-\lambda t} e^{\lambda t R}$$

also has non-negative entities. Upon using this fact, $\mu e^{t\Lambda}$ is element-wise non-negative. Let $\mathbf{1}$ be $n$-dimensional column vector with 1 for all its elements. Then

$$\frac{d}{dt}(\mu t \mathbf{1}) = \mu_t \Lambda \mathbf{1} = 0$$

Therefore $\mu_t \mathbf{1} = \mu \mathbf{1} = 1$, since $\mu \in \mathcal{P}(X)$. Since $\mu_t(\cdot) \geq 0$ and $\mu_t \mathbf{1} = 1$, $\mu_t \in \mathcal{P}(X)$.

Now we claim the following proposition.

**Proposition 11.1** For a given Q-matrix $\Lambda$ and an arbitrary initial distribution, a stochastic process on $X$ is well-defined via forward Kolmogorov equation. Moreover, it is a time-homogeneous Markov process with transition probability matrix $P_t$ given by the following ODE:

$$\frac{dP_t}{dt} = P_t \Lambda \quad P_0 = I_n$$

*Proof:* The first statement is natural consequence of the previous claim. The Markov property is justified by choosing initial condition by $\mu_t_m(i) = 1_{i=a_m}$. From the forward equation,

$$\mu_h = \mu_0 P_h = \mu_0 (I_n + h\Lambda + o(h))$$

Since $\mu_0$ may be arbitrarily chosen in probability simplex, we conclude

$$P_h = I_n + h\Lambda + o(h)$$

$$\Rightarrow \frac{P_h - I_n}{h} = \Lambda + o(1)$$

$$\Rightarrow \frac{P_h P_t - P_t}{h} = \Lambda P_t + o(1)$$

By C-K equation, $P_h P_t = P_{0,h} P_{h,t+h} = P_{t+h}$, we conclude

$$\frac{dP_t}{dt} = \Lambda P_t \quad P_0 = I_n$$

Also note that $P_t P_h = P_{0,t} P_{t,t+h} = P_{t+h}$, and hence $\frac{dP_t}{dt} = P_t \Lambda$, that is, $P_t$ and $\Lambda$ commute.
11.2.2 Uniformization trick

Given \((\mu_0, \Lambda)\), recall the matrix \(R = I_n + \frac{1}{\lambda} \Lambda\). It is in fact a row-stochastic matrix, that is, every entity is non-negative and \(\sum_j R_{i,j} = 1\). Therefore, we can define a discrete-time Markov process \(\{\xi_k\}_{k=0}^\infty\) with \(\xi_0 \sim \mu_0\) and one-step transition probability matrix is \(R\). We also define \(\{N_t\}_{t \geq 0}\) be an independent Poisson process with rate \(\lambda\), then

\[
P[N_t = k] = e^{-\lambda t} \frac{(\lambda t)^k}{k!}
\]

Now we claim the following: Let \(\hat{X}_t := \xi_{N_t}\) then \(\hat{X}_t\) and \(X_t\) have the same probability distribution at each \(t \geq 0\).

**Proof:** We already showed that

\[
\mu_t = \mu_0 e^{t\Lambda} = \mu_0 e^{-\lambda t} e^{tR}
\]

and therefore,

\[
\mu_t(\cdot) = \sum_{k=0}^\infty P[N_t = k] P[\xi_k = \cdot]
\]

\[
= \sum_{k=0}^\infty P[N_t = k, \xi_{N_t} = \cdot]
\]

\[
= P[\hat{X}_t = \cdot]
\]

where we use independence between \(\{N_t\}\) and \(\{\xi_k\}\), and the law of total probability. \(\blacksquare\)

12 Diffusion processes

Now we consider another class of continuous-time Markov processes, diffusion processes. These have the state space \(\mathcal{X} = \mathbb{R}^n\). Let \(\{X_t\}_{t \geq 0}\) be a time-homogeneous Markov process with state space \(\mathcal{X} = \mathbb{R}^n\). We will denote by \(P_t(\cdot, \cdot)\) its transition kernel: for any \(x \in \mathbb{R}^n\), Borel set \(B \subseteq \mathbb{R}^n\), \(0 \leq s < t\),

\[
P_t(x, B) := P[X_t \in B | X_0 = x] = P[X_{s+t} \in B | X_s = x].
\]

We say it is a diffusion process if:

- for any \(x \in \mathbb{R}^n\) and \(r > 0\),

\[
\lim_{h \downarrow 0} \frac{1}{h} P[||X_h - X_0|| > r | X_0 = x] = 0,
\]
or,
\[
\lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}_h(X, (\mathbb{B}^n(x, r))^c) = 0,
\]
where \(\mathbb{B}^n(x, r)^c\) is the complement of the ball in \((\mathbb{R}^n, \| \cdot \|)\) with center \(x\), radius \(r\).

- There exist functions: \(b_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, n\) and \(A : \mathbb{R}^n \to \mathbb{R}^{n \times n}\), such that \(A(x) := (a_{ij}(x))_{i,j=1}^n\) is a symmetric positive-semidefinite \(n \times n\) matrix for any \(x \in \mathbb{R}^n\), and for any \(x \in \mathbb{R}^n\) and \(r > 0\)

\[
\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[(X_{h,i} - X_{0,i})_1 \{\|X_h - X_0\| \leq r\} | X_0 = x] = b_i(x), \forall i.
\]

and

\[
\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[(X_{h,i} - X_{0,i})(X_{h,j} - X_{0,j})_1 \{\|X_h - X_0\| \leq r\} | X_0 = x] = a_{ij}(x), \forall i, j,
\]

i.e.,

\[
\mathbb{E}[X_h | X_0] = b(x) h + o(h)
\]

\[
\text{Cov}(X_h | X_0) = A(x) h + o(h)
\]

The vector-valued function \(b : \mathbb{R}^n \to \mathbb{R}^n\) given by \(b(x) := (b_1(x), \ldots, b_n(x))^T\) is called the (local) drift of the diffusion process, while the matrix-valued function \(A : \mathbb{R}^n \to \mathbb{R}^{n \times n}\) is the (local) covariance matrix.

As an example, consider the standard Brownian motion (or Wiener process) in \(\mathbb{R}^n\). Recall that it is a random process \(\{W_t\}_{t \geq 0}\) on \(\mathbb{R}^n\) satisfying the following:

- \(W_0 = 0\).
- \(W_t - W_s \sim \mathcal{N}(0, (t - s)I_n), \ 0 \leq s < t\).
- \((W_t - W_s) \perp (W_r - W_s), \ 0 \leq r < s < t\).

Then it can be shown that \(\{W_t\}_{t \geq 0}\) is a diffusion process with \(b(x) \equiv 0\) and \(A(x) \equiv I_n\) for all \(x \in \mathbb{R}^n\).

### 12.1 Forward Kolmogorov equation

Now we will show how one can construct the forward Kolmogorov equation for a diffusion process specified by the pair \((b(x), A(x))\). Recall the finite-state case: If \(\{X_t\}_{t \geq 0}\) is a time-homogeneous Markov process with finite state space \(\mathcal{X} = \{1, \ldots, n\}\) and the transition intensities matrix \(\Lambda \in \mathbb{R}^{n \times n}\), then the probability law \(\mu_t(\cdot) := \mathbb{P}[X_t = \cdot]\) evolves according to

\[
\frac{d\mu_t}{dt} = \mu_t \Lambda,
\]
with a given initial condition \( \mu_0 \). Explicitly, \( \mu_t = \mu_0 e^{At} \); in particular, using this with \( \mu_0(x') := 1_{\{x'=x\}} \) for some fixed \( x \) gives the formula for the transition probabilities: \( P_t(x, x') = \mathbb{P}[X_t = x'|X_0 = x] = e^{At}(x, x') \) — that is, the entry in row \( x \), column \( x' \) of \( e^{At} \). We can also use the forward Kolmogorov equation to keep track of the evolution of expected values. Let a function \( f : \mathcal{X} \rightarrow \mathbb{R}^n \) and an initial state \( x \in \mathcal{X} \) be given. Just as we represented probability distributions on \( \mathcal{X} \) by row vectors, we can represent such an \( f \) by a column vector \( f = (f(1), \ldots, f(n))^\top \). Then 

\[
\mathbb{E}[f(X_t)|X_0 = x] = \sum_{x' \in \mathcal{X}} P_t(x, x')f(x') = P_t f(x),
\]

where the quantity on the right-hand side is obtained by multiplying the column vector \( f \) on the left by the matrix \( P_t \). Since 

\[
\frac{dP_t}{dt} = P_t \Lambda = \Lambda P_t, \quad P_0 = I_n
\]

it follows that 

\[
\frac{d}{dt} P_t f = P_t \Lambda f = \Lambda P_t f.
\]

It follows directly from definitions that, for each \( t \geq 0 \), the map \( f \mapsto P_t f \) is linear, and it also has the following two crucial properties: if \( f \geq 0 \) everywhere, then \( P_t f \geq 0 \) everywhere (that is, \( P_t \) preserves positivity) and \( P_1 \mathbf{1} = \mathbf{1} \), where \( \mathbf{1} = (1, 1, \ldots, 1)^\top \) is the constant function \( x \mapsto 1 \).

We can develop a similar formalism for diffusion processes. Let \( \{X_t\}_{t \geq 0} \) be a time-homogeneous Markov process on \( \mathbb{R}^n \) with transition kernel \( P_t(x, B) \). Let a bounded measurable function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be given; for each \( t \geq 0 \) and each \( x \in \mathbb{R}^n \), define the function \( P_t f \) by 

\[
P_t f(x) := \int_{\mathbb{R}^n} f(x') P_t(x, dx')
\]

— recall that, for each \( x \) and \( t, P_t(x, \cdot) \) is a Borel probability measure on \( \mathbb{R}^n \), and, in particular, taking \( f(x) = 1_B(x) \) for any Borel set \( B \subseteq \mathbb{R}^n \), we see that 

\[
P_t(x, B) = \int_B P_t(x, dx') = \int_{\mathbb{R}^n} 1_B(x') P_t(x, dx') = P_t 1_B(x).
\]

It is also easy to see that \( f \mapsto P_t f \) is linear, preserves positivity and constants. Moreover, it follows from the Chapman–Kolmogorov equation that, for any \( s, t \geq 0 \), \( P_{s+t} = P_s \circ P_t = P_t \circ P_s \) and \( P_0 \) is the identity operator: \( P_0 f = f \) for any \( f \) (we say that \( P_t \) has the semigroup property).

Now we take the cue from our construction of the transition intensity matrix \( \Lambda \) and consider the limit 

\[
\lim_{h \downarrow 0} \frac{P_h f - f}{h} =: Af,
\]

provided the limit exists. In other words, for any \( x \in \mathbb{R}^n \), any sufficiently regular \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), and any small enough \( h > 0 \),

\[
P_h f(x) = f(x) + h \cdot Af(x) + o(h).
\]
Moreover, by time homogeneity and by the semigroup property \( P_{s+t} = P_t \circ P_s \),
\[
P_{t+h}f(x) = P_tf(x) + h \cdot AP_tf(x) + o(h)
\]
or, equivalently,
\[
E[f(X_{t+h})|X_0 = x] = E[f(X_t)|X_0 = x] + h \cdot E[Af(X_t)|X_0 = x] + o(h)
\]
which leads to the forward Kolmogorov equation
\[
\frac{dP_tf}{dt} = AP_tf = P_tAf, \quad P_0f = f.
\]
The linear operator \( A \) is called the generator of the Markov process; in general, it will be well-defined only for \( f \) in some set \( D(A) \), which is called the domain of \( A \).

Now, it can be shown that, if \( \{X_t\}_{t \geq 0} \) is a diffusion process with drift \( b(x) \) and diffusion matrix \( A(x) \), then the generator \( A \) acts on any \( f : \mathbb{R}^n \to \mathbb{R} \) which is twice continuously differentiable as follows:
\[
Af = b^T \nabla f + \frac{1}{2} \text{tr}(A \nabla^2 f).
\]
Coming back to the example of the standard Brownian motion with \( b = 0 \) and \( A = I_n \), we see that it has the generator
\[
Af = b^T \nabla f + \frac{1}{2} \text{tr}(A \nabla^2 f)
\]
\[
= \frac{1}{2} \text{tr} \nabla^2 f
\]
\[
= \frac{1}{2} \Delta f,
\]
where \( \Delta \) is the Laplace operator. Using this, it is not hard to derive the transition kernel of the Brownian motion: let \( f : \mathbb{R}^n \to \mathbb{R} \) be a bounded measurable function. Then it is readily verified by direct calculation that the forward Kolmogorov equation
\[
\frac{d}{dt}P_tf = \frac{1}{2} \Delta P_tf, \quad P_0f = f
\]
has the solution
\[
P_tf(x) = \frac{1}{(2\pi t)^{n/2}} \int_{\mathbb{R}^n} f(x') \exp \left(-\frac{1}{2t} ||x' - x||^2 \right) dx'
\]
\[
= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x + \sqrt{t}Z) \exp \left(-\frac{1}{2} ||x'\|^2 \right) dx'
\]
\[
\equiv E[f(x + \sqrt{t}Z)],
\]
where \( Z \sim N(0, I_n) \).

In general, one can extract the drift \( b \) and a diffusion matrix \( A \) from the knowledge of the generator \( A \), provided it satisfies some regularity conditions. In particular, suppose that \( Af(x) := \lim_{h \to 0} \frac{1}{h}(P_hf(x) - f(x)) \) has the following properties:
Then, for any Borel probability measure

Then, for each

Then it can be shown that

is well-defined. In fact, an explicit construction relies on the following scheme: Choose any factoriza-

that the probability laws

Given an arbitrary

One way to construct

and

A is as follows: Choose any function

η : \mathbb{R}^n \rightarrow \mathbb{R}

with the following properties:

• \eta \in C^\infty, i.e., \eta is infinitely continuously differentiable.

• \eta(z) = 1 for z \in \mathbb{B}^n(0, 1).

• \eta(z) = 0 for z \notin \mathbb{B}^n(0, 2).

Given an arbitrary \( x \in \mathbb{R}^n \), for \( i, j \in [n] \) consider the functions

\[
\begin{aligned}
f_i(z) &:= \eta(z - x)(z_i - x_i), \\
f_{ij}(z) &:= f_i(z)f_j(z) = \eta^2(z - x)(z_i - x_i)(z_j - x_j).
\end{aligned}
\]

Then it can be shown that \( \mathcal{A}f(x) \) has the form (3) with \( b_i(x) = \mathcal{A}f_i(x) \) and \( a_{ij}(x) = \mathcal{A}f_{ij}(x) \).

### 12.2 Constructing diffusion processes

Now we are ready to state the following key result: Let a drift \( b(x) \) and a diffusion matrix \( A(x) \) be given. Suppose there exists some constant \( 0 < c < \infty \), such that

\[
\text{tr} A(x) + \max\{b(x)^T x, 0\} \leq c(1 + \|x\|^2), \quad \forall x \in \mathbb{R}^n.
\]

Then, for any Borel probability measure \( \mu_0 \) on \( \mathbb{R}^n \), there exists a diffusion process \( \{X_t\}_{t \geq 0} \), such that the probability laws \( \mu_t(\cdot) := P[X_t \in \cdot] \) satisfy

\[
\int f \, d\mu_t = \int f \, d\mu_0 + \int_0^t \left( \int \mathcal{A}f \, d\mu_s \right) \, ds
\]

for any \( f \) such that

\[
\mathcal{A}f = b^T \nabla f + \frac{1}{2} \text{tr}(A \nabla^2 f)
\]

is well-defined. In fact, an explicit construction relies on the following scheme: Choose any factorization

A(x) = \sigma(x)\sigma(x)^T

for some \( \sigma(x) \in \mathbb{R}^{n \times r} \) and, for a fixed \( h > 0 \), consider the process \( \{X_t^{(h)}\}_{t \geq 0} \)

defined by

\[
X_t^{(h)} = X_{kh}^{(h)} + hb(X_{kh}^{(h)}) + \sigma(X_{kh}^{(h)})(W_t - W_{kh}), \quad t \in [kh, (k+1)h),
\]

where

\[
W_t := \int_0^t \sigma(W_s) \, dB_s,
\]

and \( B_t \) is a standard one-dimensional Brownian motion.

Moreover, the diffusion matrix of this process satisfies

\[
A(X^{(h)}) = \sigma(X^{(h)})\sigma(X^{(h)})^T
\]

and the drift of this process satisfies

\[
b(X^{(h)}) = \sigma(X^{(h)})^T h^2 B
\]

for some symmetric \( 2 \times 2 \) matrix \( B = (b_{ij}) \).
where \( \{W_t\} \) is the standard Wiener process on \( \mathbb{R}^r \). Then one takes the limit as \( h \downarrow 0 \) and obtains a limiting process \( \{X_t\}_{t \geq 0} \) whose marginal laws \( \mu_t \) obey (4). In fact, if \( A(x) \equiv 0 \) everywhere, then this is just the familiar Peano–Euler scheme for approximately solving the deterministic ODE

\[
dX_t = b(X_t) \, dt
\]

with the random initial condition \( X_0 \sim \mu_0 \). Here, \( dX_t \) is shorthand for \( X_{t+dt} - X_t \). In the next lecture, we will show that we can interpret the diffusion process \( \{X_t\}_{t \geq 0} \) as the solution of a stochastic differential equation

\[
dX_t = b(X_t) \, dt + \sigma(X_t) \, dW_t, \quad X_0 \sim \mu_0
\]

where \( dW_t \) is shorthand for infinitesimal increments of the Wiener process: \( dW_t = W_{t+dt} - W_t \). In order to do this properly, we will introduce the notion of stochastic integral due to K. Itô.