LIFE and MATHEMATICAL LEGACY of ITÔ SENSEI

Masatoshi Fukushima (Osaka University)

November 26, 2015

Kiyosi Itô's Legacy from a Franco-Japanese Perspective

Cournot Centre and IHES

France Embassy in Tokyo

Kiyosi Itô was born on September 7, 1915, in Mie prefecture (middle south of Japan)

He entered Tokyo University, Department of Mathematics, in 1935 and graduated from it in 1938.

After graduation, he worked in the Statistical Bureau of the Goverment in Tokyo until he became an Associate Professor of Nagoya University in 1943.

In 1942, Itô published two fundamental papers:

[I.1(1942)] On stochastic processes (infinitely divisible laws of probability)(Doctoral Thesis), *Japan. Journ. Math.* **XVIII**(1942), 261-301

[I.2(1942)] Differential equations determining a Markoff process (in Japanese), *Journ. Pan-Japan Math. Coll.* No. 1077(1942);
(in English) in *Kiyosi Itô Selected Papers*, 42-75, Springer-Verlag, 1986

The following (displayed with blue color) are some excerpts from the Video

Lecture by Kiyosi Itô at Probability Seminar, Kyoto University, 1985

recorded in the year of his retirement from Gakushuin-University, Tokyo.

In 1935, I visited with my classmates Kodaira and Kawada a bookstore, a foreign book dealer at Tokyo, Kanda , and encountered

[K1(1933)] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung,* Erg. der Math., Berlin, 1933.

Kodaira told me that this is said to be a Probability Theory. I did not take it seriously at that time.

But after 1937, I got gradually interested in Probability theory and realized that this book by Kolmogorov is what I have been truely looking for.

Since then, I have laid it as the firm cornerstone in my mathematical thinking.

When I came across with the book

[L(1937)] P. Lévy, *Théorie de l'Addition des Variables Aléatoires,* Gauthier-Villars, Paris, 1937

where the process of independent increments were studied
in the sample path level

and the Lévy-Khinchin formula was then derived by taking expectation,

I was extremely impressed by the presentations in it and
I really thought that Probability Theory must be developed in this way.

I felt that here is a new science, a new culture in mathematics called Probability Theory bearing a new type of interest of its own.

In order to study it, we apply other technique like Fourier analysis, differential equations.

In [I.1(1942)], I made the presentation in [L(1937)] more rigorous with a help of the idea in

[D(1937)] J.L. Doob, Stochastic processes depending on a continuous parameter, *Trans. Amer. Math. Soc.* **42**(1937), 107-140

on the concept of the càdlàg version of the path.

However, for P. Lévy, this might not be so original.

In fact, he mentioned in the first part of his book the already well known compound Poisson processes
that arose practically as an expenditure process of an insurance company.

As their possible limits, he eventually attained
all the stochastic continuous processes of independent increments
quite constructively with no big surprize.

Nevertheless, I got interested in probability theory by Lévy's book [L(1937)].

With this experience, I turned to a construction of a Markov process by interpreting Kolmogorov's analytic approach in

[K.2(1931)]A. Kolmogorov, Über die analytischen Methoden in der Wahrscheilichkeitsrechnung, *Math. Ann.* **104**(1931), 415-458

in the following way:

the sample path $\{X_t,\ t \geq 0\}$ of a Markov process (in the diffusion context)
is a curve possessing as its tangent at each instant t the infinitesimal Gaussian process

$$\sqrt{v(t, X_t)}dB_t + m(t, X_t)dt$$

with mean $m(t, X_t)$ and variance $v(t, X_t)$

I had to spend almost one year

by drawing such pictures of solution curves repetedly

before arriving at a right formulation of SDE.

My first duaghter was 2 years old, and I remember that she said

Daddy is always drawing pictures of kites !

Such a repetition of drawing curves eventually convinced myself that

a general coefficient of the equation ought to be a functional of the whole past events $\{X_s; s \leq t\}$,

a process adapted to a filtration $\{\mathcal{F}_t\}$ in the modern term.

In this sense, Probability Theory is truely

an infinite dimensional and a non-linear analysis.

I clearly realized this when I gave later an new proof of Maruyama's criterion on the ergodicity of a stationary Gaussian process

by making use of multiple Wiener integrals.

An extended English version of the Japanese paper [I.2(1942)] was eventually published in

[I.3(1951)] K. Itô, On stochastic differential equations, *Mem. Amer. Math. Soc.* **4**(1951), 1-51

by a kind arrangment of J. L. Doob.

SDE theory developed slowly when Itô devoted himself to the study of the one-dimensional diffusion theory.

In 1960's, many researchers paid their attentions on general theory of Markov process, martingales and their relationship to potential theory.

In particular, Minoru Motoo and Shinzo Watanabe gave a profound analysis on square integrable additive functionals of Markov processes accompanied by a special case of a new notion of stochastic integrals.

At the same time, the Doob-Meyer decomposition theorem of sub-martingales was completed
by Paul-André Meyer.

These works merged into Kunita-Watanabe-Meyer's formulation of Itô integral and Itô formula
in the magnificient framework of semimartingales.

From 1954 to 1956, Itô was a Fellow of the Institute for Advanced Study at Princeton University.

To emphasize the advantage of his sample path-level approach,

Itô quoted in the Introduction of his AMS-Memoirs [I.3(1951)] on SDE

a 1936 paper by W. Feller where the fundamental solution of the one-dimensional Kolmogorov differential equation

$$u_t(t,x) = \frac{1}{2}a(x)u_{xx}(t,x) + b(x)u_x(t,x), \quad a(x) > 0, \quad x \in \mathbb{R}, \tag{1}$$

was constructed together with a certain local property.

Even the path continuity of the associated Markov process was not shown at that time however.

But, after moving to Princeton in 1950, W. Feller resumed his study of the Kolmogorov equation (1) with an entirely renewed approach :

To rewrite the right hand side of (1) in terms of quantities that are intrinsic for the one-dimensional diffusion process X behind it.

When I was in Princeton, Feller calculated repeatedly for one dimensional diffusion with a simple generator $\mathcal{G}u = \frac{a}{2}u'' + bu'$ for constants $a,\ b$, the quantities like

$$s(x) = P_x(\sigma_\alpha < \sigma_\beta), \quad m(x) = -\frac{dE.[\sigma_\alpha \wedge \sigma_b]}{ds}, \quad \alpha < x < \beta.$$

In the beginning, I wondered why he repeated such computations so simple as exercises, not for objects in higher dimensions.

However, in this way, he was bringing out intrinsic topological invariants that do not depend on the differential structure.
I understood that the one dimensional diffusion is a topological concept but I was not so throughgoing as Feller.

When Feller told about this, he said he once heard from Hilbert saying that

Study a very simple case far more profoundly than others,

then you will truely understand a general case.

Today I like to convey this message to you.

Later on, Itô and McKean represented the C_b-generator of a general minimal diffusion X^0 on a regular interval $I = (r_1, r_2)$ as

$$\mathcal{G}u = \left[d\frac{du}{ds} - udk \right] \; / \; dm$$

by a strictly increasing continuous function s called a canonocal scale, a positive Radon measure k called a killing measure and a positive Radon measure m of full support called a speed measure.

In my student days of Itô Sensei, I often heard of a famous phrase by Feller :

X^0 travels following the road map indicated by (s, k)

and with speed indicated by m

A best way to appreciate Feller's phrase would be to look at, instead of the generator $\mathcal{G}u$, the assoicated bilinear form

$$\mathcal{E}(u,v) \ \left(= \ -\int_I \mathcal{G}uvdm\right) \ = \ \int_I \frac{du}{ds}\frac{dv}{ds}ds + \int_I uvdk$$

defined on the function space

$$\mathcal{F}_e = \{u: \ \text{absolutely continuous in } s, \ \mathcal{E}(u,u) < \infty,$$
$$u(r_i) = 0 \text{ whenever } |s(r_i)| < \infty\}.$$

Then $(\mathcal{F}_e \cap L^2(I;m), \mathcal{E})$ is a regular local Dirichlet form on $L^2(I;m)$ associated with X^0.

Making a time change of X^0 amounts to a change of m, while keeping the space $(\mathcal{F}_e, \mathcal{E})$ invariant.

Notice that $(\mathcal{F}_e, \mathcal{E})$ (called the extended Dirichlet space of X^0) is independent of m
and may well be considered as the road map for X^0.

Nowadays we know that Feller's phrase well applies to a general symmetric Markov process X by employing

as a speed measure

a symmetrizing smooth measure (a certain Borel measure charging no set of zero capacity) of full quasi-support

and as a road map

the extended Dirichlet space $\mathcal{F}_e$ of X consisting of quasi-continuous functions equipped with the Beurling-Deny formula of the form $\mathcal{E}$.

The Dirichlet space theory due to Arne Beuring and Jacques Deny appeared in 1959.

In the beginning of his address at ICM 1958 Edinburgh, Feller pointed out

an intimate relationship between the Beurling-Deny theory of general potentials, $\cdots$, and the Hunt's basic results concerning potentials and Markov processes despite diversity of formal appearances and methods.

The notion of the extended Dirichlet space was introduced later in 1974 by Martin Silverstein a PhD student of Feller.

The Abel Symposium 2005 'Stochastic analysis and its applications' in honor of 90-th birthday of Kiyosi Itô took place in Oslo.
Itô was not able to attend because he was hospitalized, but he sent Memoirs of My Reseach on Stochastic Analysis to be read in the opening of the Symposium. Here are two excerpts.

When I was in Princeton, I learned about Feller's ongoing work from Henry McKean, a graduate student of Feller, while I explained my previous work to McKean.

There was once an occasion when McKean tried to explain to Feller my work on the stochastic differential equations along with the above mentioned idea of tangent.

It seemed to me that Feller did not fully understand its significance, but when I explained Lévy's local time to Feller, he immediately appreiciated its relevance to the study of the one-dimensional diffusion.

Indeed, Feller later gave us a cojecture that the elastic reflecting Brownian motion on the half line could be constructed directly from the reflecting Brownian motion B^r by using Lévy's local time at $\{0\}$,

that was eventually substantiated in my work with McKean in 1963.

Itô and McKean went on further to construct probabilistically from the reflecting Brownian motion B^r

all possible Markovian extensions of the absorbing Brownian motion B^0 on $(0, \infty)$ to $[0, \infty)$ that had been analytically characterized by Feller in terms of boundary conditions at the origin 0.

Itô and McKean called such extensions Feller's Brownian motions.

However the construction was not so straightforward.

If a point a of the state space S of a general standard Markov process X is regular for itself,

then the local time of X at a is well defined

and its inverse is an increasing Lévy process.

In 1970, Itô was bold enough to replace the jump size at the jump time of this increasing Lévy process by the excursion of X around a, namely, the portion of the sample path $X_t(\omega)$ starting at a until returning back to a.

A Poisson point process taking value in the space U of excursions around point a was then associated,
and its characteristic measure $\mathbf{n}$ (a σ-finite measure on U) together with the minimal process X^0 obtained from X by killing upon hitting a
was shown to determine the law of X uniquely.
This approach may be considered as an infinite dimensional analogue to a part of the decomposition of the Lévy process I studied in 1942,
and may have revealed a new aspect in the study of Markov processes.

The excursion law $\mathbf{n}$ is also determined by the minimal process X^0 coupled with an X^0-entrance law $\{\mu_t\}$ (a σ-finite measure on $S \setminus \{a\}$).

When $X = B^r$ and $X^0 = B^0$, μ_t is known to be

$$\mu_t(dx) = 2(2\pi t^3)^{-1/2} x \exp(-x^2/(2t))dx, \quad x > 0.$$

Itô's Excursion Theory are being used and developed extensively in diverse directions including:

- It enables us to construct Feller's Brownian motion from the absorbing Brownian motion B^0 at once by a suitable employment of the excursion path space U and the excursion law $\mathbf{n}$ in accordance with Feller's boundary condition at $\{0\}$.

  Shinzo Watanabe has developed this idea to give a probabilistic construction of the diffusion process on the half space satisfying Wentell's boundary condition.

- It make the study of distributions of variety of Brownian functionals considerably straightforward and transparent

  as can be seen in the works by D. Williams, J. Pitman, M. Yor.

  Marc Yor examined and emphasized
  great efficiency of Itô's excursion theory
  in parallel with that of Itô's stochastic calculus on semimartingales

  in his article as an invited editor of
  the special issue a tribute to Kiyosi Itô of
  Stochastic processes and their applications, Vol.120, 2010.

- Itô's 1970 work is a wonderful decomposition theorem of Markov processes
  but the following basic question was left unanswered in it:

  Is the X^0-entrance law μ_t involved in the excursion law $\mathbf{n}$ uniquely determined by the minimal process X^0 ?

  My joint work with Hiroshi Tanaka in 2005 gave an affirmative answer for a general m-symmetric diffusion X as

  $$\int_0^\infty \mu_t \, dt \; = \; \mathbf{P}^0_x(X^0_{\zeta^0-} = a \, ; \; \zeta^0 < \infty) \, m(dx)$$

  by invoking a unique integral representation theorem of an X^0-purely excessive measure due to P.J. Fitzsimmons.

This makes it possible to construct a unique m-symmetric extension of X^0 from $S \setminus \{a\}$ to S admitting no killing nor sojourn at a.

For instance, the celebrated Walsh's Brownian motion on the plane can be conceived most naturally this way.

Furthermore, such a construction is robust enough to be carried out

by replacing a point a of S with a compact subset $A \subset S$ and

by regarding A as a one point a^*.

This idea is being effectively used

to extend the SLE (stochastic Loewner evolution) theory

from a simply connected planar domain to multiply connected ones.