Wie K. Itô den stochastischen Kalkül revolutionierte

W. Schachermayer*   J. Teichmann†

12. September 2007

Der Gaußpreis 2006 wurde am Internationalen Mathematikerkongress in Madrid an den 90-jährigen Kiyoshi Itô verliehen. Der Preis, der zum ersten Mal vergeben wurde, und der aus den Überschüssen des Berliner Kongress 1998 finanziert wird, würdigt

- “outstanding mathematical contributions that have found significant applications outside of mathematics, or
- achievements that made the application of mathematical methods to areas outside of mathematics possible in an innovative way”.

Zentrales Objekt der Arbeiten von K. Itô ist die Brownsche Bewegung. Wenn Lichtstrahlen in einen dunklen Raum fallen, kann man mit freiem Auge die Staubteilchen in der Luft bei ihrem erratischen Tanz beobachten. Der Grund für diese scheinbar regellose Bewegung liegt darin, dass die kleinen Staubteilchen ständig von noch viel kleineren – für das Auge nicht sichtbaren – Teilchen angestoßen werden, wobei die Richtung dieser Stöße zufällig ist (im gegenständlichen dreidimensionalen Fall gleichverteilt über die Kugeloberfläche).

Dieses Phänomen wurde aufgrund von intensiven Beobachtungen unter dem Mikroskop von dem bedeutenden britischen Botaniker Robert Brown 1827 beschrieben; allerdings ohne die oben angegebene und aus heutiger Sicht sehr einleuchtende Begründung, und auch ohne nur den Versuch einer mathematischen Modellierung zu unternehmen.

Dies wurde dann in meisterhafter Weise von Albert Einstein in seiner Dissertation im Jahr 1905 getan. Sein Artikel [2] ist eine der drei großen Arbeiten dieses „annus mirabilis“ 1905.

*Vienna University of Technology, Wiedner Hauptstrasse 8-10/105, A-1040 Wien, Austria. Financial support from the Austrian Science Fund (FWF) under the grant P19456, from Vienna Science and Technology Fund (WWTF) under Grant MA13 and by the Christian Doppler Research Association (CDG) is gratefully acknowledged.

†Vienna University of Technology, Wiedner Hauptstrasse 8-10/105, A-1040 Wien, Austria. Financial support from the Austrian Science Fund (FWF) under the grant Y328 and from Vienna Science and Technology Fund (WWTF) under Grant MA13 is gratefully acknowledged.

Unabhängig von A. Einstein hatte der in Vorderbrühl 1872 geborene polnische Physiker Marian Smoluchowski ähnliche Ideen, die er 1906 publizierte [5].

Beide hatten allerdings bereits einen mathematischen Vorläufer, von dem sie nichts wussten: Louis Bachelier hatte in seiner Dissertation "Théorie de la spéculation" im Jahr 1900 [1] die eindimensionale Version der Brownschen Bewegung zur Modellierung von Börsenkursen benutzt. Die ökonomische Motivation ist der physikalischen – zumindest formal – ähnlich: man stellt sich vor, dass an der Börse laufend Kauf- oder Verkaufsorder eintreffen, die dem Börsenkurs einen kleinen Stoß nach oben oder unten geben. Die Gleichverteilungsannahme lautet in diesem eindimensionalen Fall, dass die Wahrscheinlichkeit eines Stoßes nach oben bzw. unten jeweils $\frac{1}{2}$ beträgt.

Dieses Modell erlaubte L. Bachelier, die seinen Forschungen zugrundeliegende praktische Fragestellung in sehr eindrucksvoller Weise zu beantworten, nämlich einen kohärenten Kalkül zur Bewertung von Optionen zu entwickeln. Auf dem Weg dorthin hat er eine mathematische Modellierung der Brownschen Bewegung entwickelt und wichtige Eigenschaften abgeleitet (z.B. die Verteilung des Maximums eines Brownschen Pfads, die enge Verbindung mit der Wärmeleitungsgleichung etc.).

Die präzise mathematische Behandlung der Brownschen Bewegung und der darauf aufbauenden Konzepte ist – auch aus heutiger Sicht – durchaus heikel und anspruchsvoll. Am einfachsten beginnt man mit dem elementaren Konzept der Irrfahrt ("random walk"), die gerne als das Torkeln eines Betrunkenen (auf einer Linie) interpretiert wird. Fixieren wir die Raum- und Zeit-Diskretisierungen $\Delta x > 0$, $\Delta t > 0$, und eine Folge $(\varepsilon_n)_{n=1}^{\infty}$ von unabhängigen, identisch verteilten Zufallsvariablen mit $\mathbf{P}[\varepsilon_n = 1] = \mathbf{P}[\varepsilon_n = -1] = \frac{1}{2}$. Der stochastische Prozess X ist dann definiert als

$$X_{n\Delta t} = \Delta x \sum_{i=1}^{n} \varepsilon_i, \quad n \in \mathbb{N}_0. \tag{1}$$

Anders ausgedrückt bewegt sich X in einem Zeitintervall der Länge Δt um $\pm\Delta x$. Es liegt nahe, von diesem diskreten Prozess mit $\Delta X_t = X_{t+\Delta t} - X_t = \varepsilon_t \Delta x$ zum kontinuierlichen Limes überzugehen, indem man Δt und Δx gegen Null gehen lässt. Die Formel für die Varianz von Summen von unabhängigen Zufallsvariablen zeigt unmittelbar, dass bei diesem Limes-Übergang $\Delta x \approx \sqrt{\Delta t}$ gelten muss: um nämlich einen sinnvollen Limes zu erhalten, muss der Quotient $\frac{(\Delta x)^2}{\Delta t}$ bei den gewählten Verfeinerungen der Zeit- und Raum-Inkremente gegen einen endlichen Grenzwert streben. Wenn wir diesen Quotienten mit 1 normieren, erhalten wir im Limes die Brownsche Bewegung $B = (B_t)_{t \geq 0}$.

Aus dem zentralen Grenzwertsatz folgt unmittelbar, dass für jedes $t \geq 0$ die Zufallsvariable B_t normalverteilt ist, mit Mittelwert 0 und Varianz t. Etwas allgemeiner ist, für jedes $0 \leq t_1 < t_2 < \ldots < t_n$, der Zufallsvektor

$(B_{t_1}, \ldots, B_{t_n})$ normalverteilt, mit Mittelwert 0 und der offensichtlichen Kovarianzmatrix. Dies genügt bereits für eine präzise Definition von B. Wie nämlich A. Kolmogoroff in den 30er-Jahren noch gezeigt hat, lässt sich eine Theorie der stochastischen Prozesse $X = (X_t)_{t\geq 0}$ auf den endlichdimensionalen Verteilungen von $(X_{t_1}, \ldots, X_{t_n})$ aufbauen.

Diese Betonung der endlichdimensionalen Verteilungen erlaubt keinen direkten Zugang zum pfadweisen Verhalten des stochastischen Prozess $(B_t)_{t\geq 0}$, d.h. wie für ein festes Zufallselement ω der „Pfad“ oder die „Trajektorie“ $(B_t(\omega))_{t\geq 0}$ beschrieben werden kann. Diese dynamische Sichtweise aber war für L. Bachelier sehr naheliegend, da er an der Pariser Börse arbeitete und täglich die Kursentwicklungen „in stetiger Zeit“ beobachtete.

Es erweist sich als nützlich, formal zum infinitesimalen Analogon von (1) überzugehen, wobei wir Δt durch dt ersetzen, und Δx durch $dx = dt^{\frac{1}{2}}$:

$$dB_t = \varepsilon_t \sqrt{dt}, \quad t \geq 0. \tag{2}$$

Hier bezeichnet dB_t das infinitesimale Inkrement $dB_t = B_{t+dt} - B_t$ und $(\varepsilon_t)_{t\geq 0}$ ist eine Familie von unabhängigen Zufallsvariablen mit $\mathbf{P}[\varepsilon_t = 1] = \mathbf{P}[\varepsilon_t = -1] = \frac{1}{2}$.

Die intuitive Interpretation von (2) besagt, dass zu jedem Zeitpunkt t eine Münze aufgeworfen wird und dementsprechend der Prozess B sich während des Intervalls $[t, t+dt]$ um $\sqrt{dt}$ hinauf- oder hinunterbewegt. Selbstverständlich ist dies alles nur eine formale, heuristische Interpretation, die, wie in der Analysis üblich, durch epsilon-delta-Argumente (oder non-standard bzw. white noise analysis) gerechtfertigt werden muss.

Die Differenzialgleichung (2) benützt die Leibniz-Notation; dies ist die einzig mögliche Notation in diesem Kontext, da der Differenzialquotient $\frac{dB_t}{dt}$ keinen direkten Sinn ergibt. Er ist ja $\varepsilon_t dt^{-\frac{1}{2}}$. Eine intuitive Interpretation der Formel $\frac{dB_t}{dt} = \varepsilon_t dt^{-\frac{1}{2}}$ wäre, dass zu jedem Zeitpunkt t der Anstieg der Kurve $t \mapsto B_t(\omega)$ entweder $+\infty$ oder $-\infty$, wobei jede Möglichkeit die Wahrscheinlichkeit $\frac{1}{2}$ hat und für jeden Zeitpunkt t unabhängig ist. Aber da begeben wir uns schon auf sehr dünnes Eis.

Die präzise (und relativ leicht zu beweisende) mathematische Aussage dazu lautet: Für fast alle Zufallselemente ω ist der Pfad $t \mapsto B_t(\omega)$ in keinem Punkt differenzierbar. Ebenso kann man relativ leicht zeigen, dass für fast alle ω der Pfad $t \mapsto B_t(\omega)$ auf jedem Intervall $]a, b[$ mit $a < b$ unendliche Totalvariation hat.

Der Begriff der Brownschen Bewegung erlaubt es nunmehr, allgemeinere Konzepte zu definieren, indem man die Brownsche Bewegung als Baustein verwendet. Ein Itô-Prozess $X = (X_t)_{t\geq 0}$ kann durch einen stochastischen Differenzialausdruck der folgenden Bauart beschrieben werden. Der Einfachkeit halber beschränken wir uns auf den eindimensionalen Fall:

$$dX_t = \sigma_t dB_t + \mu_t dt. \tag{3}$$

Hier sind $(\sigma_t)_{t\geq 0}$ und $(\mu_t)_{t\geq 0}$ stochastische Prozesse, die zu jedem Zeitpunkt t nur von der Vergangenheit der Brownschen Bewegung $(B_u)_{u\geq 0}$ bis t abhängen. In vielen Anwendungen sind diese Prozesse von der Form $\sigma_t = \sigma(X_t)$ und $\mu_t = \mu(X_t)$ für deterministische Funktionen $\sigma : \mathbb{R} \to \mathbb{R}$ und $\mu : \mathbb{R} \to \mathbb{R}$. In diesem Fall beschreibt die Differenzialgleichung

$$dX_t = \sigma(X_t)dB_t + \mu(X_t)dt$$

einen Markov-Prozess, da die Information über die Diffusions- und Drift-Parameter σ_t und μ_t nur vom Wert der Zufallsvariablen X_t abhängt, und nicht von deren Vergangenheit. A. Kolmogoroff konstruierte solche Prozesse durch Berechnung der endlichdimensionalen Verteilungen aus der zugehörigen partiellen Differentialgleichung. Auch wenn man intuitiv gerne den obigen Differenzialausdruck zu Hilfe nahm, ein rigoroses Verständnis dieser Gleichung hatte man in den 30er-Jahren noch nicht. Folgende (heuristische) Überlegung und ihr bemerkenswertes Resultat unterstreichen die Notwendigkeit, dem Differenzialausdruck eine präzise mathematische Bedeutung zu geben.

Sei $f : \mathbb{R} \to \mathbb{R}$ eine (genügend glatte) Funktion und betrachten wir, für einen gegebenen Diffusionsprozess $(X_t)_{t\geq 0}$, der (3) erfüllt, den stochastischen Prozess $(Y_t)_{t\geq 0} := (f(X_t))_{t\geq 0}$. Die Frage ist, wie der zum Prozess $Y = (Y_t)_{t\geq 0}$ assoziierte stochastische Differenzialausdruck der Bauart (3) lautet. Mit anderen Worten fragen wir nach der Form der Kettenregel in dem gegenwärtigen stochastischen Kontext.

Die entscheidende Beobachtung, die durch die formale Gleichung (2) insinuiert wird, besteht darin, dass $(dB_t)^2$ von der Ordnung dt ist und daher (asymptotisch) berücksichtiget werden muss: formales Quadrieren von (2) liefert die bemerkenswerte Gleichung

$$(dB_t)^2 = dt. \tag{4}$$

Um die Differenzialgleichung für $Y = f(X)$ abzuleiten, entwickeln wir das infinitesimale Inkrement $dY_t = Y_{t+dt} - Y_t$ formal in eine Taylorreihe, wobei wir erst nach dem quadratischen Glied abbrechen:

$$\begin{aligned} dY_t &= f(X_{t+dt}) - f(X_t) \\ &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 + o(dX_t)^2. \end{aligned}$$

Unter Verwendung von (3) und (4) sowie Vernachlässigung aller Terme von kleinerer Ordnung als dt erhalten wir

$$\begin{aligned} dY_t &= f'(X_t)dX_t + f''(X_t)\frac{\sigma^2(X_t)}{2}dt \qquad (5)\\ &= f'(X_t)\sigma(X_t)dB_t + \left[f'(X_t)\mu(X_t) + f''(X_t)\frac{\sigma^2(X_t)}{2}\right] dt. \end{aligned}$$

Die zweite Zeile zeigt, dass der Differenzialausruck für Y wieder von derselben Bauart wie (3) ist, d.h. Itô-Prozesse sind abgeschlossen unter Komposition mit (hinreichend glatten) deterministischen Funktionen f. Die erste Zeile zeigt, dass die im deterministischen Fall wohlbekannte Kettenregel $df(X_t) = f'(X_t)dX_t$ im stochastischen Kontext durch den Term $f''(X_t)\frac{\sigma^2(X_t)}{2}dt$ ergänzt werden muss.

Die Formel (5) wird als Itô-Formel oder Itô-Lemma bezeichnet. Sie ist ein ebenso fundamentales Werkzeug für die stochastische Analysis wie die „gewöhnliche" Kettenregel für die klassische Analysis. Selbstverständlich folgen die weiteren Rechenregeln wie Produktformel, Quotientenformel etc. rasch aus der Kettenregel.

Bemerkenswert ist, dass die Berechnung des Differenzials erster Ordnung $dY_t = d(f(X_t))$ die zweite Ableitung von f ins Spiel bringt. Dies führt zu einer engen Verbindung zwischen Diffusionsprozessen einerseits und der Wärmeleitungsgleichung andererseits. Die gründliche Erforschung dieser wichtigen Querverbindung ist – neben Kiyoshi Itô – mit den Namen A. Kolmogoroff, S. Kakutani, R. Feynman, M. Kac verbunden; sie erwies sich als eine Goldmine für interessante Mathematik, z.B. in der Theorie der partiellen Differentialgleichungen oder der Differentialgeometrie, die bis heute noch keineswegs erschöpft ist.

Die Tatsache, dass die Übergangswahrscheinlichkeiten der Brownschen Bewegung die Wärmeleitungsgleichung erfüllen, war bereits 1900 von L. Bachelier erkannt worden, der dies in seiner Dissertation klar formulierte und bewies, aber nicht weiter ausführte. Henri Poincaré schrieb in seinem – ansonsten sehr positiven – Gutachten über diese Dissertation den geradezu prophetischen Satz: „Man könnte bedauern, dass Monsieur Bachelier diesen Teil seiner Dissertation nicht weiter ausgeführt hat."

Zurück zur Itô-Formel (5), die wir nur heuristisch motiviert haben, wobei wir durchaus verdächtige Argumente benützt haben, z.B. Größen der Ordnung $\sqrt{dt}$. Die Aufgabe besteht darin, den Differenzialausdrücken (3) einen präzisen mathematischen Sinn zu geben. Als ersten Schritt formen wir (3) in Integral-Form um:

$$X_t = X_0 + \int_0^t \sigma_u dB_u + \int_0^t \mu_u du.$$

Das zweite Integral macht keine Schwierigkeiten und kann mit der üblichen Lebesgueschen Integrationstheorie definiert werden. Heikel ist dagegen das erste Integral, da die Pfade $(B_u(\omega))_{0 \leq u \leq t}$ keine endliche Totalvariation besitzen.

Dieses Problem bereitete bereits Norbert Wiener Kopfzerbrechen. In einer Reihe von Arbeiten beweist er in den 20er-Jahren zum ersten Mal präzise, dass ein Modell für die Brownsche Bewegung existiert, in dem die Pfade $(B_t(\omega))_{t \geq 0}$ *stetig* sind (eine Tatsache, die wir in den heuristischen Argumenten oben wie selbstverständlich vorausgesetzt haben). Diese Arbeiten sind der Grund dafür, dass die Brownsche Bewegung auch Wiener-Prozess genannt wird. A. Kolmogoroff hatte übrigens die Bezeichnung Bachelier-Wiener Prozess vorgeschlagen, konnte sich damit aber nicht durchsetzen.

Was die stochastische Integration betrifft, konnte N. Wiener aber nur eine sehr eingeschränkte Antwort gegen. Er konnte zeigen, dass für einen Integranden $(\xi_t)_{t\geq 0}$, dessen Pfade $(\xi_t(\omega))_{t\geq 0}$ *von beschränkter Variation* sind, das stochastische Integral $\int \xi_t dB_t$ wohldefiniert werden kann. In diesem Fall ist die Situation nämlich sehr einfach: Mithilfe von partieller Integration kann das Integral $\int \xi_t dB_t$ auf das Integral $\int B_t d\xi_t$ zurückgeführt werden,

$$\int_0^T \xi_t dB_t = \xi_T B_T - \xi_0 B_0 - \int_0^T B_t d\xi_t,$$

und letzteres kann als Stieltjes-Integral interpretiert werden, wenn eben $(\xi_t)_{t\geq 0}$ von endlicher Variation ist.

Die crux ist, dass die in den Anwendungen vorkommenden Integranden typischerweise *nicht* von beschränkter Variation sind: wenn wir $\int \sigma(X_t)dB_t$ betrachten, wobei $(X_t)_{t\geq 0}$ ein Diffussionsprozess ist, dann ist der Integrand $(\sigma(X_t))_{t\geq 0}$ nicht von beschränkter Variation, falls die erste Ableitung von σ nicht identisch verschwindet. Die Situation scheint hoffnungslos, das Integral $\int \xi_t dB_t$ sinnvoll zu definieren, wenn für fast alle Zufallselemente ω sowohl die Funktion $t \mapsto \xi_t(\omega)$ als auch $t \mapsto B_t(\omega)$ von unbeschränkter Variation sind. Norbert Wiener hat, trotz langen und intensiven Nachdenkens, in den 20er-Jahren keine Lösung dazu gefunden.

Der Geniestreich gelang dann 1942 dem 27-jährigen K. Itô, der damals allein und isoliert im statistischen Büro der japanischen Regierung in Tokio saß. Seine Idee war, *nicht* zu versuchen, das Integral $\int \xi_t dB_t$ pfadweise für jedes Zufallselement ω und *jeden* Integranden ξ separat zu definieren, sondern es global für den gesamten Wahrscheinlichkeitsraum Ω und für eine genügend reichhaltige Klasse von Integranden zu konstruieren. Insbesondere sind die Limiten von Riemannschen Summen dann nicht mehr als „fast sicher“ zu verstehen, sondern nur mehr im L^2. Wie so oft in der Mathematik war dabei eine Invarianzeigenschaft des stochastischen Integrals von entscheidender Bedeutung.

Hier eine Skizze der Konstruktion: Man beginnt mit „simplen“ Integranden $(\xi_t)_{t\geq 0}$, für die das Integral $\int \xi_t dB_t$ in trivialer Weise durch endliche Riemannsche Summen definiert ist. Dieser Schritt entspricht der Integration von Treppenfunktionen im üblichen Aufbau der (nicht-stochastischen) Integrationstheorie. Bis hier erfolgt die Argumentation für jedes ω separat. Der entscheidende Schritt ist dann die Frage, wie man zu einem Limes dieser „simplen“ Integranden übergehen kann. Dabei bemerkte Itô, dass für bestimmte „simple“ Integranden $(\xi_t)_{t\geq 0}$, solche nämlich, wo ξ_t nur von der Vergangenheit der Brownschen Bewegung $(B_u)_{0\leq u\leq t}$ bis zum Zeitpunkt t abhängt, die Riemannschen Summen wieder ein Martingal definieren, so wie die Brownsche Bewegung selbst eines ist. Diese Integranden nennt man adaptiert und diese Invarianzeigenschaft gestattet es das stochastische Integral als Isometrie zwischen zwei Hilberträumen zu verstehen.

Itô fasste, für festes $t > 0$, einen simplen, adaptierten Integranden $(\xi_u)_{0 \le u \le t}$ als Funktion auf $\Omega \times [0, t]$ auf und betrachtet die Integration als linearen Operator, die jedem simplen, adaptierten Integranden $(\xi_u)_{0 \le u \le t}$ das Integral $\int_0^t \xi_u dB_u$ zuordnet. Dieser lineare Operator geht von einem Raum von adaptierten Funktionen auf $\Omega \times [0, t]$ in einen Raum von Funktionen auf Ω. Nun versieht man $\Omega \times [0, t]$ mit dem Maß $\mathbf{P} \otimes \lambda$, wobei $\mathbf{P}$ das gegebene Wahrscheinlichkeitsmaß auf Ω und λ das Lebesguemaß auf $[0, t]$ ist. Nun zeigt Itô, dass der Integrations-Operator eine Isometrie bezüglich der von $L^2(\mathbf{P} \otimes \lambda)$ bzw. $L^2(\mathbf{P})$ induzierten Hilbertnormen ist. Diese fundamentale Tatsache ist erstaunlich einfach zu beweisen und lag damals quasi in der Luft: zu etwa der gleichen Zeit zeigte A. Wald, der bis 1938 an der Universität Wien in der Gruppe rund um Karl Menger tätig war, die nach ihm benannten Identitäten, die einen analogen isometrischen Zusammenhang zwischen L^2-Normen formulieren. Mathematisch gesehen liegt der Grund für diese isometrischen Eigenschaften bezüglich geeigneter Hilbert-Normen darin, dass die Martingal-Eigenschaft bestimmten Orthogonalitäts-Relationen im Hilbertraum entspricht.

Wenn die Isometrie-Eigenschaft gezeigt ist, wird die Erweiterung des Integrals von „simplen, adaptierten" zu möglichst allgemeinen Integranden zu einer Formalität: es genügt, den auf den simplen Integranden definierten linearen Operator auf den Abschluss dieses Vektorraums im $L^2(\Omega \times [0, t])$ zu erweitern, man spricht dann einfach von adaptierten Integranden. Adaptierte Integranden σ_t müssen nicht von beschränkter Variation sein, es genügt, dass σ_t nur von der Brownschen Bewegung bis zum Zeitpunkt t abhängt und gewisse Integrabilitätsbedingungen erfüllt.

Aus heutiger Sicht scheint diese Konstruktion einfach und naheliegend. Wie immer in der Mathematik: wenn man weiß, wie es geht, ist alles ganz einfach! Aus historischer Sicht ist der Zugang von Itô aber als revolutionär zu bezeichnen (wir wählen hier bewusst dieselben Worte wie M. Yor in [7]). Jedenfalls schloss er damit das Tor zu einem neuen Gebiet der Mathematik auf: der stochastischen Analysis.

Wir möchten an dieser Stelle nicht versäumen auch auf die tragische Geschichte von Wolfgang Döblin, einem Sohn des Schriftstellers Alfred Döblin (Autor des Romans „Berlin Alexanderplatz") hinzuweisen, der erst im Jahr 2000 zu der Bekanntheit gelangte, die er verdient.

Wolfgang Döblin, der ebenso wie Itô im Jahr 1915 geboren wurde, musste mit seiner Familie 1933 aus dem Dritten Reich über Zürich nach Paris fliehen. In Paris begann er dann Mathematik am Institut Henri Poincaré bei P. Lévy und M. Fréchet zu studieren. Bis 1939 hatte er bereits ein Dutzend kleinere Arbeiten – vor allem über Markov-Prozesse – veröffentlicht, als er in die französische Armee einberufen wurde. Während K. Itô seine Arbeiten 1942 im statistischen Büro der japanischen Regierung entwickelte, hatte Wolfgang Döblin auf der anderen Seite des Globus bereits zwei Jahre davor sehr ähnliche Ideen entwickelt, während er im Winter 1939/40 als Telephonist in einer Kaserne in den Ardennen stationiert war. Er hielt sie in einem umfangreichen handgeschriebenen Manuskript "Sur l'équation de Kolmogoroff" fest. Diese zeitliche Koinzidenz lässt sich auf eine Gemeinsamkeit von Wolfgang Döblin und Kiyoshi Itô zurückführen: das intensive Studium der Werke von Paul Lévy.

Das Manuskript "Sur l'équation de Kolmogoroff" hinterlegte W. Döblin am 26. Februar 1940 in einem verschlossenen Kuvert ("pli cacheté") bei der Akademie der Wissenschaften in Paris. Am 21. Juni 1940 erschoss er sich, eingekesselt von deutschen Truppen, als jüdischer Emigrant in französischer Uniform, der sich über seine Situation keine Illusionen machte.

Das "pli cacheté Nr. 11668" ruhte bis zum 18. Mai 2000 verschlossen in den Archiven der Pariser Akademie der Wissenschaften, bis nach langen Verhandlungen von Akademiemitgliedern, insbesondere durch das Engagement von Jean-Pierre Kahane, erreicht wurde, dass die überlebenden Brüder von Wolfgang Döblin die Einwilligung zur Öffnung des Kuverts gaben. Die Sensation war perfekt: In dem Artikel hatte Döblin einen Satz bewiesen, aus dem die Itô-Formel leicht ableitbar ist (zumindestens im eindimensionalen Fall). Er verwendete dabei aber völlig andere Argumente als K. Itô.

Das Manuskript von Döblin wurde von B. Bru und M. Yor in einem 2002 bei "Finance and Stochastics" erschienenen Artikel liebevoll editiert und mit vielen Kommentaren versehen [6]. Wir verweisen auf diesen sehr lesenswerten Artikel für detaillierte Informationen und wollen hier nur den Versuch wagen, den Zugang von W. Döblin mit wenigen Worten zu skizzieren. Ausgangspunkt ist wieder die Gleichung (3)

$$dX_t = \sigma(X_t)dB_t + \mu(X_t)dt,$$

die Döblin "l'équation de Kolmogoroff" nennt. Die Idee ist, den Prozess X dieser Gleichung durch geschickte Transformationen in den Fall $\mu_t \equiv 0$ und $\sigma_t \equiv 1$, also $X_t = B_t$, überzuführen. Dass eine erfolgreiche Durchführung dieser Transformation, die den Fall eines allgemeineren Diffusionsprozesses X auf den Fall einer Brownschen Bewegung B zurückführt, "en passant" eine Kettenregel liefert, also die Itôsche Formel zeigt, ist ziemlich klar.

Döblin begann damit, den Driftterm μ_t zu eliminieren, was sich als der relativ einfachere Teil des Programms erwies. Hier konnte er mit klassischem Lebesgue-Integralen $\int \mu_t dt$ operieren und musste nicht – wie Itô zwei Jahre später – durch Einführung eines allgemeineren Integralbegriffs mathematisches Neuland betreten.

Nach diesem ersten Schritt reduziert sich das Problem auf die Betrachtung von Prozessen der Form

$$dX_t = \sigma_t dB_t. \tag{6}$$

Die entscheidende Idee Döblins, um in (6) die Situation auf den Fall $\sigma_t \equiv 1$ zurückführen, besteht darin, eine Transformation der Zeitvariablen t vorzunehmen.

Die Reparametrisierung der Zeit ist allerdings nicht so einfach, wie man meinen könnte, da σ_t nicht nur von der Zeit, sondern auch vom Zufallselement ω abhängt. Diese Technik wurde unter den Begriffen "stochastic time change" oder "sub-ordination" in den 60er- und 70er-Jahren durch eine Reihe von Mathematikern und Mathematikerinnen entwickelt. Diese hatten allerdings keine Ahnung vom Inhalt des "pli cacheté Nr. 11668". W. Döblin hat im Winter 1939/40 in einem beeindruckenden Kraftakt diese Technik im Alleingang entwickelt und konnte – so wie K. Itô zwei Jahre später – einen Differenzialkalkül für Diffusionsprozesse entwickeln.

Durch seinen Zugang der Zeittransformation konnte Wolfgang Döblin das oben skizzierte Konzept des Itô-Integrals umgehen. Für ihn ist die Lösung X der "Kolmogoroff-Gleichung" (3) im Fall $\mu_t \equiv 0$ eine Zeit-Transformierte einer Brownschen Bewegung, was aus heutiger Sicht völlig korrekt ist.

Das Itô-Integral seinerseits musste auch etwa 15 bis 20 Jahre warten, bis es von breiteren Mathematikerkreisen wahrgenommen wurde. Dann aber wurde es rasch zum zentralen Werkzeug der stochstischen Integral- und Differenzialrechnung und deren Anwendungen. Wir möchten das an einem Beispiel illustrieren:

Wir haben oben skizziert, dass für jeden adaptierten Integranden $(\xi_u)_{0\leq u\leq t}$ in $L^2(\Omega \times [0,t])$ das Integral $\int_0^t \xi_u dB_u$ ein wohldefiniertes Element von $L^2(\Omega)$ ist. Der Martingal-Darstellungs-Satz gibt eine Art Umkehrung an. Unter geeigneten Messbarkeitsvoraussetzungen gilt, dass für jedes $f \in L^2(\Omega)$ mit Erwartungswert $\mathbf{E}[f] = 0$ ein eindeutiger adaptierter Integrand $(\xi_u)_{0\leq u\leq t} \in L^2(\Omega \times [0,t])$ existiert, sodass

$$f = \int_0^t \xi_u dB_u.$$

Das diskrete Analogon, d.h. der entsprechende Satz für den Fall der Irrfahrt mit T Schritten, läuft auf die simple Tatsache hinaus, dass das Haarsche System eine Orthonormal-Basis im $\mathbb{R}^{2^T}$ bildet.

Der Martingale-Darstellungs-Satz ist zum Angelpunkt der Anwendungen der stochastischen Analysis im Finanzbereich geworden, da er die Idee der „Replikation" eines derivativen Finanztitels (z.B. einer Option) f durch dynamisches Handeln (dies enspricht dem Integranden $(\xi_u)_{0\leq u\leq t}$) im zugrundeliegenden Basistitel formalisiert. Man weiß nicht nur, dass eine Handelsstrategie existiert, sondern man kann diese auch konkret mithilfe der Itô-Formel berechnen.

K. Itô hatte nicht im Traum an diese Anwendungen gedacht, als er vor über 50 Jahren diesen Satz bewies, wie er einmal lachend dem erstgenannten Autor versicherte. Er fügte aber hinzu, dass er sich außerordentlich darüber freut, dass seine aus rein mathematischer Neugier entwickelten Arbeiten so breite Anwendung in den verschiedensten Disziplinen finden.

Das Werk von K. Itô beschränkt sich selbstverständlich nicht nur auf die Entwicklung des stochastischen Kalküls in seinen jungen Jahren, sondern er blieb bis ins hohe Alter außerordentlich aktiv. Wir möchten hier noch zwei bemerkenswerte Themen aus seinem großen Werk skizzieren (siehe dazu den ausgezeichneten Artikel [7], wo wir auch den Titel dieses Aufsatzes entlehnt haben).

Das erste Thema stellt eine schöne Verbindung zwischen der Stochastik und der Funktionentheorie her. Wir betrachten einen $\mathbb{C}$-wertigen Prozess $Z_t = X_t + iY_t$, wobei X und Y zwei unabhängige reellwertige Brownsche Bewegungen sind. Sei $f : \mathbb{C} \setminus \{z_1, \ldots, z_n\} \to \mathbb{C}$ eine meromorphe Funktion, wobei der Startpunkt Z_0 der Brownschen Bewegung $(Z_t)_{t\geq 0}$ nicht in der Menge $\{z_1, \ldots, z_n\}$ enthalten ist. Dann trifft der Brownsche Pfad $(Z_t)_{t\geq 0}$ fast sicher nie auf die Punkte $\{z_1, \ldots, z_n\}$, sodass es möglich ist, das komplexe Itô-Integral $\int f(Z_t)dZ_t$ zu definieren.

Ein konkretes Beispiel: Mit Hilfe dieses Integrals kann man eine stochastische Darstellung des Arguments ϑ_t der Brownschen Bewegung Z_t geben:

$$\vartheta_t - \vartheta_0 = \operatorname{Im} \int_0^t \frac{dZ_u}{Z_u} = \int_0^t \frac{X_u dY_u - Y_u dX_u}{X_u^2 + Y_u^2}.$$

Dies gestattet beispielsweise einen raschen und eleganten Beweis (mit funktionentheoretischen Argumenten) des Satzes von Spitzer über die Windungszahlen eines komplexen Brownschen Pfads,

$$\frac{2}{\log t}\vartheta_t \overset{t\to\infty}{\longrightarrow} C_1,$$

d.h. die Verteilung der Zufallsvariablen $\frac{2\vartheta_t}{\log t}$ konvergiert, für $t \to \infty$, gegen die Cauchy-Verteilung C_1.

Als zweites Beispiel wollen wir die Theorie der Exkursionen von Markov-Prozessen erwähnen, die Itô im Jahr 1970 entwickelt hat. Wir wollen sie für den Fall einer reell-wertigen Brownschen Bewegung $B = (B_t)_{t\geq 0}$ skizzieren. Die Idee besteht darin, die Brownschen Pfade $(B_t(\omega))_{t\geq 0}$ als die Gesamtheit der „Exkursionen“, d.h. ihres Verhaltens zwischen je zwei aufeinanderfolgenden Nullstellen, aufzufassen. Die Idee ist: wenn man die Exkursionen kennt, kann man daraus den Brownschen Pfad zusammensetzen. Etwas präziser: wir assoziieren zu jedem reell-wertigen Brownschen Pfad $(B_t(\omega))_{t\geq 0}$ die Menge $\mathcal{Z}(\omega)$ der Nullstellen

$$\mathcal{Z}(\omega) = \{t \geq 0 \mid B_t(\omega) = 0\}.$$

Die Menge $\mathcal{Z}(\omega)$ hat, für fast alle ω, Lebesguemaß null und hat eine ähnliche Struktur wie die Cantor-Menge. Man kann ein sinnvolles (zufälliges) Maß auf $\mathcal{Z}(\omega)$ definieren, indem man die „lokale Zeit“ L_t einführt, die formal gegeben ist durch

$$dL_t(\omega) = \lim_{\varepsilon\to 0}\left(\frac{1}{2\varepsilon}\mathbf{1}_{\{|B_t(\omega)|<\varepsilon\}}dt\right).$$

Dieses Konzept wurde bereits in den 40er-Jahren von P. Lévy studiert. Die inverse Funktion von L_t

$$\tau_l = \inf\{t \mid L_t > l\}, \quad l \geq 0$$

erlaubt es, die Menge der Exkursionen zu „etikettieren“:

$$e_l(\omega): \; t \mapsto B_{\tau_{l\_}(\omega)+t} \mathbf{1}_{\{t<\tau_l(\omega)-\tau_{l\_}(\omega)\}}.$$

Die Kenntnis der Trajektorie $(e_l(\omega))_{l\geq 0}$ mit Werten in den Exkursionen ist äquivalent zur Kenntnis des Brownschen Pfads $(B_t(\omega))_{t\geq 0}$.

Der in diesem Kontext fundamentale Satz von Itô besagt, dass der Prozess $(e_l)_{l\geq 0}$ ein Poisson-Prozess ist, allerdings mit Werten in der Menge der Exkursionen! Damit konnte Itô eine tiefe Verbindung zwischen den beiden Archetypen der stochastischen Prozesse herstellen, nämlich der Brownschen Bewegung und dem Poisson-Prozess.

Es gäbe noch viele interessante Beispiele zu nennen und viele bedeutende Gebiete anzuführen, die auf den Entdeckungen von Kiyoshi Itô aufbauten. Wir verweisen dazu auf Hans Föllmers Laudatio am ICM in Madrid [3] und natürlich auf Kiyoshi Itôs ausgewählte Werke [4], wo D. Stroock und S. R. S. Varadhan (er ist Abel-Preisträger 2007) in der Einleitung schreiben: “Everyone who is likely to pick up this book has at least heard that there is a subject called the theory of stochastic integration and that K. Itô is the Lebesgue of this branch of integration theory (Paley and Wiener were its Riemann).”

Literatur

[1] L. Bachelier, Théorie de la spéculation. *Ann. Sci. Ecole Norm. Sup.* 17 (1900).

[2] A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. der Phys.* (4) 17, 549–560 (1905).

[3] H. Föllmer, On Kiyoshi Itô’s Work and its Impact. Proceedings of the International Congress of Mathematicians, Madrid 2006, Vol. I, 109–124.

[4] K. Itô, Selected Papers, edited by D.W. Stroock and S.R.S. Varadhan, Springer-Verlag, 1986.

[5] M. Smoluchowski, Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. der Phys.* (4) 21, 756–780 (1906).

[6] B. Bru, M. Yor, Comments on the life and mathematical legacy of Wolfgang Doeblin. *Finance Stoch.* 6 (2002), no. 1, 3–47.

[7] M. Yor, Comment K. Itô a révolutinné l’étude des processus stochastiques, *Gazette des SMF* 111, Janvier 2007.